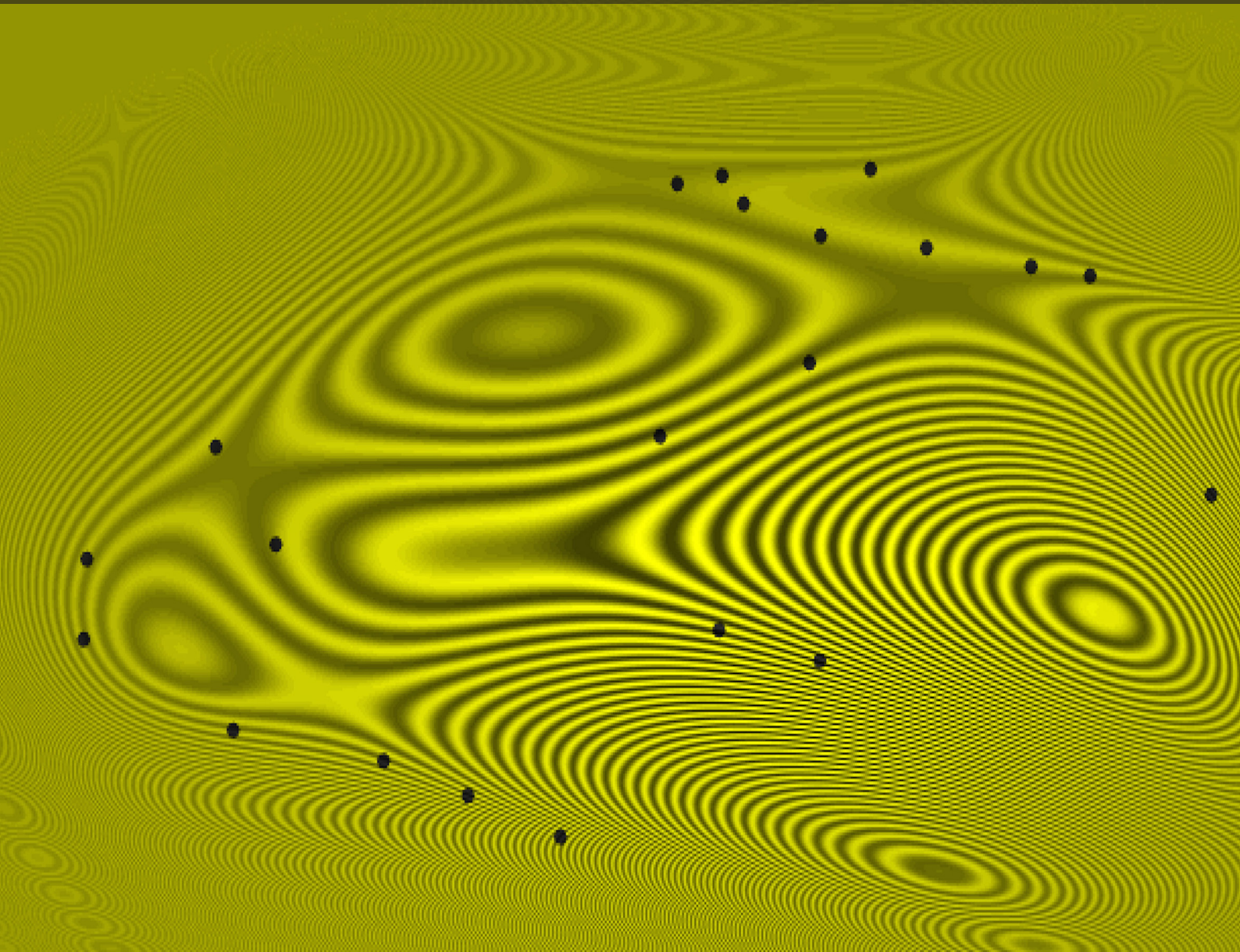# QUANTUM STRUCTURES IN COGNITIVE AND SOCIAL SCIENCE

EDITED BY : Diederik Aerts, Jan Broekaert, Liane Gabora and Sandro Sozzo
PUBLISHED IN : Frontiers in Psychology

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**
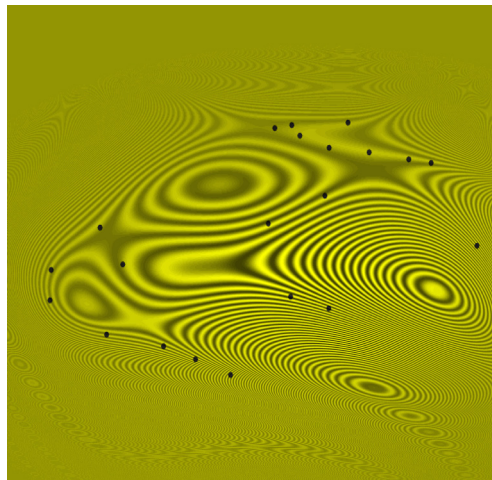
QUANTUM STRUCTURES IN COGNITIVE AND SOCIAL SCIENCE

Topic Editors:
Diederik Aerts, Free University of Brussels, Belgium
Jan Broekaert, Free University of Brussels, Belgium
Liane Gabora, University of British Columbia, Canada
Sandro Sozzo, University of Leicester, UK

Double slit-like pattern created by the concept Fruits interfering with the concept Vegetables in the disjunction Fruits or Vegetables.
Image by Diederik Aerts

Traditional approaches to cognitive psychology correspond with a classical view of logic and probability theory. More specifically, one typically assumes that cognitive processes of human thought are founded on the Boolean structures of classical logic, while the probabilistic aspects of these processes are based on the Kolmogorovian structures of classical probability theory. However, growing experimental evidence indicates that the models founded on classical structures systematically fail when human decisions are at stake. These experimental deviations from classical behavior have been called `paradoxes', `fallacies', `effects' or `contradictions', depending on the specific situation where they appear. But, they involve a broad spectrum of cognitive and social science domains, ranging from conceptual combination to decision making under uncertainty, behavioral economics, and linguistics. This situation has constituted a serious drawback to the development of various disciplines, like cognitive science, linguistics, artificial intelligence, economic modeling and behavioral finance.

A different approach to cognitive psychology, initiated two decades ago, has meanwhile matured into a new domain of research, called 'quantum cognition'. Its main feature is the use of the mathematical formalism of quantum theory as modeling tool for these cognitive situations where traditional classically based approaches fail. Quantum cognition has recently attracted the interest of important journals and editing houses, academic and funding institutions, popular science and media. Specifically, within a quantum cognition approach, one assumes that human decisions do not necessarily obey the rules of Boolean logic and Kolmogorovian probability, and can on the contrary be modeled by the quantum-mechanical formalism. Different concrete quantum-theoretic models have meanwhile been developed that successfully represent the

cognitive situations that are classically problematical, by explaining observed deviations from classicality in terms of genuine quantum effects, such as `contextuality', `emergence', `interference', `superposition', `entanglement' and `indistinguishability'. In addition, the validity of these quantum models is convincingly confirmed by new experimental tests. We also stress that, since the use of a quantum-theoretic framework is mainly for modeling purposes, the identification of quantum structures in cognitive processes does not presuppose (without being incompatible with it) the existence of microscopic quantum processes in the human brain.

In this Research Topic, we review the major achievements that have been obtained in quantum cognition, by providing an accurate picture of the state-of-the-art of this emerging discipline. Our overview does not pretend to be either complete or exhaustive. But, we aim to introduce psychologists and social scientists to this challenging new research area, encouraging them, at the same time, to consider its promising results. It is our opinion that, if continuous progress in this domain can be realized, quantum cognition can constitute an important breakthrough in cognitive psychology, and potentially open the way towards a new scientific paradigm in social science.

Table of Contents

Editorial: Quantum Structures in Cognitive and Social Science

Diederik Aerts[1], Jan Broekaert[1], Liane Gabora[2] and Sandro Sozzo[3]*

[1] *Center Leo Apostel for Interdisciplinary Studies, Free University of Brussels, Brussels, Belgium,* [2] *Department of Psychology, University of British Columbia, Kelowna, BC, Canada,* [3] *School of Management, Institute for Quantum Social and Cognitive Science, University of Leicester, Leicester, UK*

The Editorial on the Research Topic

Quantum Structures in Cognitive and Social Science

A fundamental problem in cognitive and social science concerns the identification of the principles guiding human cognitive acts such as decision-making, categorization, and behavior under uncertainty. Identifying these mechanisms would have manifold implications for fields ranging from psychology to economics, finance, politics, computer science, and artificial intelligence. The predominant theoretical paradigm rests on a classical conception of logic and probability theory. According to this paradigm people make decisions by following the rules of Boole's logic, while the probabilistic aspects of these decisions can be formalized by Kolmogorov's probability theory. This classical approach was believed to provide a quite complete and accurate account of human decision-making at both a normative level (describing what people should do) and a descriptive level (describing what people actually do). However, starting from the seventies, experimental studies of conceptual categorization, human judgment and perception, and behavioral economics have revealed that this classical conception is fundamentally problematical, in the sense that the cognitive models based on these mathematical structures are not capable of capturing how humans make decisions in situations involving uncertainty. In the last decade, an alternative scientific paradigm has arisen that employs a different and more general modeling scheme; it uses the mathematical formalism of quantum theory to model situations and processes in cognitive and social science. This new approach has not only met with considerable success but is becoming increasingly accepted in the scientific community, having attracted interest from important scientists, top journals, funding institutions, and media. Prisoners' dilemmas, conjunction and disjunction fallacies, disjunction effects, violations of the Sure-Thing principle, Allais, Ellsberg and Machina paradoxes, are only some of the examples where the application of the quantum mechanical formalism has shown significant effectiveness over traditional modeling schemes of a classical type.

The Frontiers Research Topic "Quantum Structures in Cognitive and Social Science" present an overview of current research that applies the formalism of quantum theory to cognitive and socio-economic domains. The term "quantum" may be misleading. The aim here is not to investigate the microphysical processes occurring in the human brain and, as a consequence, driving human judgments. Rather, we inquire into the validity of quantum theory as a general, coherent, and unitary paradigm for human cognition. In this respect, this research benefits from studies into the axiomatic and operational foundations of quantum physics. The scope of this bold approach to human cognition is discovering general rules that associate the empirical phenomenology in these domains with states, measurements, and probabilities of outcomes in such a way that these entities are represented exactly as quantum theory in Hilbert space represents states, measurements, and probabilities of outcomes in the phenomenology of microphysics. The ensuing modeling is

theory-based, not experiment-based; that is, the models are not built around a specific effect or experiment, although they are sometimes used in conjunction with empirical data to build a stronger case. The models are constructed following the general epistemological and technical constraints of quantum theory; hence the successes of this quantum theory-based modeling suggest that it might provide a general theory for human cognition.

This Research Topic develops around three main directions of research, as follows.

(i)   The deep reasons underlying the success of the quantum paradigm in cognitive and socio-economic domains are investigated.

(ii) Further empirical situations are identified in these domains where the quantum formalism presents advantages with respect to traditional modeling schemes, and new genuine quantum structures appear.

(iii) The application of this quantum paradigm is extended to novel and barely explored domains.

The first set of results concern knowledge representation and conceptual categorization. Aerts et al. analyze the results of a cognitive test on conjunctions and negations of natural concepts, showing that a quantum-theoretic probabilistic model in Hilbert space faithfully represents the collected data, at variance with a set-theoretic Kolmogorovian model. This result is explained by assuming the existence of two types of reasoning in human cognition, a dominant emergent reasoning, and a secondary logical reasoning. Some mathematical aspects of this quantum-theoretic model on conceptual conjunctions and negations are developed in Veloz and Desjardins through the introduction of unitary operators in Hilbert space. Aerts et al. show instead that the quantum-theoretic approach Aerts et al. can be interpreted as a suitable generalization of Rosch's prototype theory, where prototypes are context-dependent and may interfere when concepts combine.

The second set of results concern the modeling of human decision-making. Moreira and Wichert explore an alternative quantum-theoretic approach, the quantum-like Bayesian network, to describe the paradoxes related to the violation of the Sure-Thing Principle in experiments on human judgments. Their model is in a good agreement with different sets of empirical data. The opinion paper in Pothos et al. reviews some current progress on the quantum similarity model in Hilbert space recently proposed by Pothos et al. which correctly represents human similarity judgments. Decision-making errors and preference reversal are also investigated in Yukalov and Sornette within their quantum decision theory. Wang and Busemeyer analyze the notion of complementarity in human cognition, and claim that the way in which it is used in quantum physics can also be helpful in cognitive science. Human perception is the object of the study in Khrennikov, where the author develops a quantum model of the sensation-perception dynamics, illustrating it

by means of the model of bistable perception of a specific ambiguous figure, the Schröder stair. Finally, Tressoldi et al. identify a significant violation of temporal Bell inequalities in a set of cognitive tests. The violation indicates, according to the authors, the presence of temporal entanglement between binary human behavioral unconscious choices at a given time and binary random outcomes at a different time. In all these approaches, the presence of quantum structures in cognition is determined by the fact that the cognitive systems under investigation share a common feature, namely contextuality. A different position with respect to the presence of contextuality in cognition is assumed in Zhang and Dzhafarov, where the authors apply a theory of (non)contextuality to analyze series-parallel (SP) mental architectures.

The third set of results concern advanced applications of the quantum-mathematical formalism to wider ranges of social science. Bisconti et al. propose an inverse Potts model, typically used in statistical quantum field theory, to reconstruct the node states in a real-world social network. Haven explores the properties of two types of potential functions, inspired by classical and quantum physics that can be potentially employed to model financial information, including preferences toward risk and uncertainty. Finally, Dalla Chiara et al. investigate different, but mutually related, aspects of parallelism within the framework of quantum computation, cognition and music, and study potential applications of quantum computational semantics in both natural and musical language.

Leaving aside the specific differences between the approaches above, most of them agree in claiming that quantum structures are systematically present in cognitive and social science phenomena, and that quantum-inspired models are more efficient than traditional set-theoretic models of probability. Is "quantum" the end of the story? Is Hilbert space really the place where all these phenomena can be modeled? Is there any empirical deviation from quantum predictions? We do not have yet an answer to these questions. This is why we believe that the road that will lead further to possibly a generally accepted quantum theory of human decision-making will still be full of fascinating surprises.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct, and intellectual contribution to the work, and approved it for publication.

Generalizing Prototype Theory: A Formal Quantum Framework

Diederik Aerts[1], Jan Broekaert[1], Liane Gabora[2] and Sandro Sozzo[3]*

[1] Center Leo Apostel for Interdisciplinary Studies, Free University of Brussels, Brussels, Belgium, [2] Department of Psychology, University of British Columbia, Kelowna, BC, Canada, [3] School of Management, Institute for Quantum Social and Cognitive Science, University of Leicester, Leicester, UK

Theories of natural language and concepts have been unable to model the flexibility, creativity, context-dependence, and emergence, exhibited by words, concepts and their combinations. The mathematical formalism of quantum theory has instead been successful in capturing these phenomena such as graded membership, situational meaning, composition of categories, and also more complex decision making situations, which cannot be modeled in traditional probabilistic approaches. We show how a formal quantum approach to concepts and their combinations can provide a powerful extension of prototype theory. We explain how prototypes can interfere in conceptual combinations as a consequence of their contextual interactions, and provide an illustration of this using an intuitive wave-like diagram. This quantum-conceptual approach gives new life to original prototype theory, without however making it a privileged concept theory, as we explain at the end of our paper.

Keywords: cognition, concept theory, prototype theory, contextuality, interference, quantum modeling

1. INTRODUCTION

Theories of concepts struggle to capture the creative flexibility with which concepts are used in natural language, and combined into larger complexes with emergent meaning, as well as the context-dependent manner in which concepts are understood (Geeraerts, 1989). In this paper, we present some recent advances in our quantum approach to concepts. More specifically, we follow the general lines illustrated in Gabora and Aerts (2002), Aerts and Gabora (2005a,b), and Gabora et al. (2008), and generalize the quantum-theoretic model elaborated in Aerts (2009b) and Aerts et al. (2013a).

According to the "classical," or "rule-based" view of concepts, which can be traced back to Aristotle, all instances of a concept share a common set of necessary and sufficient defining properties. Wittgenstein pointed out that: (i) in some cases it is not possible to give a set of characteristics or rules defining a concept; (ii) it is often unclear whether an object is a member of a particular category; (iii) conceptual membership of an instance strongly depends on the context.

A major blow to the classical view came from Rosch's work on color. This work showed that colors do not have any particular criterial attributes or definite boundaries, and instances differ with respect to how typical they are of a concept (Rosch, 1973, 1978, 1983). This led to formulation of "prototype theory," according to which concepts are organized around family resemblances, and consist of characteristic, rather than defining, features. These features are weighted in the definition of the "prototype." Rosch showed that subjects rate conceptual membership as "graded," with degree of membership of an instance corresponding to conceptual distance from the prototype. Moreover, the prototype appears to be particularly resistant to forgetting. Prototype theory also

has the strength that it can be mathematically formulated and empirically tested. By calculating the similarity between the prototype of a concept and a possible instance of it, across all salient features, one arrives at a measure of the "conceptual distance" between the instance and the prototype. Another means of calculating conceptual distance comes out of "exemplar theory" (Nosofsky, 1988, 1992), according to which a concept is represented by, not a set of defining or characteristic features, but a set of salient "instances" of it stored in memory. Exemplar theory has met with considerable success at predicting empirical results. Moreover, there is evidence of preservation of specific training exemplars in memory. Classical, prototype, and exemplar theories are sometimes referred to as "similarity based" approaches, because they assume that categorization relies on data-driven statistical evidence. They have been contrasted with "explanation based" approaches, according to which categorization relies on a rich body of knowledge about the world. For example, according to "theory theory" concepts take the form of "mini-theories" (Murphy and Medin, 1985) or schemata (Rumelhart and Norman, 1988), in which the causal relationships among properties are identified.

Although these theories do well at modeling empirical data when only one concept is concerned, they perform poorly at modeling combinations of two concepts. As a consequence, cognitive psychologists are still looking for a satisfactory and generally accepted model of how concepts combine.

The inadequacy of fuzzy set models of conceptual conjunctions (Zadeh, 1982) to resolve the "Pet-Fish problem" identified by Osherson and Smith (1981) highlighted the severity of the combination problem. People rate the item *Guppy* as a very typical example of the conjunction *Pet-Fish*, without rating *Guppy* as a typical example neither of *Pet* nor of *Fish* ("Guppy effect") (Osherson and Smith, 1981, 1982). Studies by Hampton on concept conjunctions (Hampton, 1988a), disjunctions (Hampton, 1988b) and negations (Hampton, 1997) confirmed that traditional fuzzy set and Boolean logical rules are violated whenever people combine concepts, as one usually finds "overextension" and "underextension" in the membership weights of items with respect to concepts and their combinations. It has been shown that people estimate a sentence like "*x* is tall and *x* is not tall" as true, in particular when *x* is a "borderline case" ("borderline contradictions") (Bonini et al., 1999; Alxatib and Pelletier, 2011), again violating the rules of set-theoretic Boolean logic. The seriousness of the combination problem was pointed out by various scholars (Komatsu, 1992; Fodor, 1994; Kamp and Partee, 1995; Rips, 1995; Osherson and Smith, 1997). More recently, other theories of concepts have been developed, such as "Costello and Keane's constraint theory" (Costello and Keane, 2000), "Dantzig, Raffone, and Hommel's connectionist CONCAT model of concepts" (Van Dantzig et al., 2011), "Thagard and Stewart's emergent binding model" (Thagard and Stewart, 2011), and "Gagne and Spalding's morphological approach" (Gagne and Spalding, 2009). However, none of these theories has a strong track record of modeling the emergence and non-compositionality of concept combinations.

The approach to concepts presented in this paper grew out earlier work on the application to concept theory on the axiomatic and operational foundations of quantum theory and quantum probability (Aerts, 1986; Pitowsky, 1989; Aerts, 1999). A major theoretical insight was to shift the perspective from viewing a concept as a "container" to viewing it as "an entity in a specific state that is changing under the influence of a context" (Gabora and Aerts, 2002). This allowed us to provide a solution to the Guppy effect and to successfully represent the data collected on *Pet*, *Fish* and *Pet-Fish* by using the mathematical formalism of quantum theory (Aerts and Gabora, 2005a,b). Then, we proved that none of the above experiments in concept theory can be represented in a single probability space satisfying the axioms of Kolmogorov (1933). We developed a general quantum framework to represent conjunctions, disjunctions and negations of two concepts, which has been successfully tested several times (Aerts, 2009a,b; Sozzo, 2014, 2015; Aerts et al., 2015a), and we put forward an explanatory hypothesis for the observed deviations from traditional logical and probabilistic structures and for the occurrence of quantum effects in cognition (Aerts et al., 2015b). We recently identified a strong and systematic non-classical phenomenon effect, which is deeper than the ones typically detected in concept combinations and directly connected with the mechanisms of concept formation (Aerts et al., 2015c). This work is part of a growing domain of cognitive psychology that uses the mathematical formalism of quantum theory and quantum structures to model empirical situations where the application of traditional probabilistic approaches is problematical (probability judgments errors, decision-making errors, violations of expected utility theory, etc.; Aerts and Aerts, 1995; Aerts et al., 2000, 2013a,b, 2014, 2015; Aerts and Sozzo, 2011, 2014; Busemeyer and Bruza, 2012; Haven and Khrennikov, 2013; Pothos and Busemeyer, 2013; Khrennikov et al., 2014; Wang et al., 2014).

This paper outlines recent progress in the development of a quantum-theoretic framework for concepts and their dynamics. Section 2 explains how the "SCoP formalism" can be interpreted as a "contextual and interfering prototype theory that is a generalization of standard prototype theory" in which prototypes are not fixed, but change under the influence of a context, and interfere as a consequence of their contextual interactions (see also Gabora et al., 2008; Aerts et al., 2013a). Section 3 presents an amended explanatory version of the quantum-mechanical model in complex Hilbert space worked out in Aerts (2009b) and Aerts et al. (2013a) for the typicality of items with respect to the concepts *Fruits* and *Vegetables*, and their disjunction *Fruits or Vegetables*. This improved quantum model illustrates how the prototype of *Fruits* (*Vegetables*) changes under the influence of the context *Vegetables* (*Fruits*) in the combination *Fruits or Vegetables*. The latter combination is represented using the quantum-mathematical notion of linear superposition in a complex Hilbert space, which entails the genuine quantum effect of "interference." Hence, our model shows that the prototypes of *Fruits* and *Vegetables* interfere in the disjunction *Fruits or Vegetables*. Sections 2, 3 also justify the fact that our quantum-theoretic framework for concepts can be considered as a "contextual and interfering generalization of original prototype theory." The presence of linear superposition and interference could suggest that concepts combine and interact like waves do.

In Section 4 we develop this intuition in detail and propose an intuitive wave-like illustration of the disjunction *Fruits or Vegetables*. Finally, Section 5 discusses connections between the quantum-theoretic approach to concepts presented here, and other theories of concepts. Although this approach can be interpreted as a specific generalization of prototype theory, it is compatible with insights from other theories of concepts.

We stress that our investigation does not deal with the elaboration of a "specific typicality model" that represents a given set of data on the concepts *Fruits*, *Vegetables*, and their disjunction *Fruits or Vegetables*. We inquire into the mathematical formalism of quantum theory as a general, unitary and coherent formalism to model natural concepts. Our quantum-theoretic model in Section 3 has been derived from this general quantum theory, hence it satisfies specific technical and general epistemological constraints of quantum theory. As such, it does not apply to any arbitrary set of experimental data. Our formalism exactly applies to those data that exhibit a peculiar deviation from classical set-theoretic modeling; such deviations are taken in our framework as indicative of interference and emergence. Data collected on combinations of two concepts systematically exhibit deviations from classical set-theoretical modeling, and traditional probabilistic approaches have difficulty coping with this. In this sense, the success of the quantum-theoretic modeling can be interpreted as a confirmation of the effectiveness of quantum theory to model conceptual combinations. We should also mention that our quantum-theoretic approach has recently produced new predictions, allowing us to identify entanglement in concept combinations (Aerts and Sozzo, 2011, 2014), and systematic deviations from the marginal law, deeply connected to the mechanisms of concept formation (Aerts et al., 2015a,c). These effects would not have been identified in a more traditional investigation of overextension and underextension.

It follows from the above analysis that our quantum-theoretic modeling rests on a "theory based approach," as it straightforwardly derives from quantum theory as "a theory to represent natural concepts." Hence, it should be distinguished from an "*ad-hoc* modeling based approach," only devised to fit data. One should be suspicious of models in which free parameters are added after the fact on an *ad-hoc* basis to fit the data more closely. In our opinion, the fact that our "theory derived model" reproduces different sets of experimental data is a convincing argument to support its advantage over traditional modeling approaches and to extend its use to more complex combinations of concepts.

2. THE SCoP FORMALISM AS A CONTEXTUAL INTERFERING PROTOTYPE THEORY

This section summarizes the SCoP approach to concepts by providing new insights to the research in Aerts and Gabora (2005a,b) and Gabora et al. (2008).

We mentioned in Section 1 that, according to prototype theory, concepts are associated with a set of characteristic, rather than defining, features (or properties), that are weighted in the definition of the prototype. A new item is categorized as an instance of the concept if it is sufficiently similar to this prototype (Rosch, 1973, 1978, 1983). The original prototype theory was subsequently put into mathematical form as follows. The prototype consists of a set of features $\{a_1, a_2, \ldots, a_M\}$, with associated "weights" (or "application values") $\{x_{p1}, x_{p2}, \ldots, x_{pM}\}$, where M is the number of features that are considered. A new item k is also associated with a set $\{x_{k1}, x_{k2}, \ldots, x_{kM}\}$, where the number x_{km} refers to the applicability of the m-th feature to the item k (for a given stimulus). Then, the conceptual distance between the item k and the prototype, defined as

$$d_k = \sqrt{\sum_{m=1}^{M} (x_{km} - x_{pm})^2} \qquad (1)$$

is a measure of the similarity between item and prototype. The smaller the distance d_k for the item k, the more representative k is of the given concept.

Prototype theory was developed in response to findings that people rate conceptual membership as graded (or fuzzy), with the degree of membership of an instance corresponding to the conceptual distance from the prototype. A second fundamental element of prototype theory is that it can in principle be confronted with empirical data, e.g., membership or typicality measurements.

A fundamental challenge to prototype theory (but also to any other theory of concepts) has become known as the "Pet-Fish problem." The problem can be summarized as follows. We denote by *Pet-Fish* the conjunction of the concepts *Pet* and *Fish*. It has been shown that people rate *Guppy* neither as a typical *Pet* nor as a typical *Fish*, they do rate it as a highly typical *Pet-Fish* (Osherson and Smith, 1981). This phenomenon of the typicality of a conjunctive concept being greater than—or overextends—that of either of its constituent concepts has also been called the "Guppy effect." Using classical logic, or even fuzzy logic, there is no specification of a prototype for *Pet-Fish* starting from the prototypes of *Pet* and *Fish* that is consistent with empirical data (Osherson and Smith, 1981, 1982; Zadeh, 1982). Fuzzy set theory falls short because standard connectives for conceptual conjunction involve typicality values that are less than or equal to each of the typicality values of the conceptual components, i.e., the typicality of an item such as *Guppy* is not higher for *Pet-Fish* than for either *Pet* or *Fish*.

Similar effects occur for membership weights of items with respect to concepts and their combinations. Hampton's experiments indicated that people estimate membership in such a way that the membership weight of an item for the conjunction (disjunction) of two concepts, calculated as the large number limit of relative frequency of membership estimates, is higher (lower) than the membership weight of this item for at least one constituent concept (Hampton, 1988a,b). This phenomenon is referred to as "overextension" ("underextension"). "Double overextension" ("double underextension") is also an experimentally established phenomenon, when the membership

weight with respect to the conjunction (disjunction) of two concepts is higher (lower) than the membership weights with respect to both constituent concepts (Hampton, 1988a,b). Furthermore, conceptual negation does not satisfy the rules of classical Boolean logic (Hampton, 1997). More, Bonini et al. (1999), and Alxatib and Pelletier (2011), identified the presence of "borderline contradictions," directly connected with overextension, namely, a sentence like "John is tall and John is not tall" is estimated as true by a significant number of participants, again violating basic rules of classical Boolean logic. More generally, for each of these experimental data, a single classical probability framework satisfying the axioms of Kolmogorov does not exist (Aerts, 2009a,b; Aerts et al., 2013a,b, 2015a; Sozzo, 2014, 2015). To clarify the latter sentence no single probability space can be constructed for an item whose membership weight with respect to the conjunction of two concepts is overextended with respect to both constituent concepts.

These problems—compositionality, the graded nature of typicality, and the probabilistic nature of membership weights—present a serious challenge to any theory of concepts.

We have developed a novel theoretical model of concepts and their combinations (Gabora and Aerts, 2002; Aerts and Gabora, 2005a,b), conjunction (Aerts, 2009a; Aerts et al., 2013a, 2015a; Sozzo, 2014, 2015), disjunction (Aerts, 2009a; Aerts et al., 2013a), conjunction and negation (Aerts et al., 2015a; Sozzo, 2015). It uses the mathematical formalism of quantum theory in Hilbert space to represent data on conceptual combinations, which has been successfully tested several times. This quantum-conceptual approach enables us to model the above-mentioned deviations from classicality in terms of genuine quantum phenomena (contextuality, emergence, entanglement, interference, and superposition), thus capturing fundamental aspects of how concepts combine. More importantly, we have recently identified stronger deviations from classicality than overextension and underextension, which unveil, in our opinion, deep non-classical aspects of concept formation (Aerts et al., 2015c).

The approach was inspired by similarity based theories, such as prototype theory, in several respects:

(i)   a fundamentally probabilistic formalism is needed to represent concepts and their dynamics;
(ii)   the typicality of different items with respect to a concept is context-dependent;
(iii) features (or properties) of a concept vary in their applicability.

A key insight underlying our approach is considering a concept as, not a "container of instantiations" but, rather, an "entity in a specific state," which changes under the influence of a context. In our quantum-conceptual approach, a context is mathematically modeled as quantum physics models of a measurement on a quantum particle. The (cognitive) context changes the state of a concept in the way a measurement in quantum theory changes the state of a quantum particle (Aerts and Gabora, 2005a,b). For example, in our modeling of the concept *Pet*, we considered the context *e* expressed by *Did you see the type of pet he has?*

This explains that he is a weird person, and found that when participants in an experiment were asked to rate different items of *Pet*, the scores for *Snake* and *Spider* were very high in this context. In this approach, this is explained by introducing different states for the concept *Pet*. We call "the state of *Pet* when no specific context is present" its ground state $\hat{p}$. The context *e* changes the ground state $\hat{p}$ into a new state $p_{weird\ person\ pet}$. Typicality here is an observable semantic quantity, which means that it takes different values in different states of the concept. As a consequence, a substantial part of the typicality variations that are encountered in the Guppy effect are due to, e.g., changes of state of the concept *Pet* under the influence of a context. More specifically, the typicality variations for the conjunction *Pet-Fish* are in great part similar to the typicality variations for *Pet* under the context *Fish* (and also for *Fish* under the context *Pet*). Not only does context play a role in shaping the typicality variations for *Pet-Fish*, but also interference between *Pet* and *Fish* contributes, as we will analyze in detail in Section 3.

In general, whenever someone is asked to estimate the typicality of *Guppy* with respect to the concept *Pet* in the absence of any context, it is the typicality in the ground state $\hat{p}_{Pet}$ that is obtained, and whenever the typicality of *Guppy* is estimated with respect to the concept *Fish* in the absence of any context, it is the typicality in the ground state $\hat{p}_{Fish}$ that is obtained. But, whenever someone is asked to estimate the typicality of *Guppy* with respect to the conjunction *Pet-Fish*, it is the typicality in a new ground state $\hat{p}_{Pet-Fish}$ that is obtained. This new ground state $\hat{p}_{Pet-Fish}$ is different from $\hat{p}_{Pet}$ as well as from $\hat{p}_{Fish}$. It is close but not equal to the changed state of the ground state $\hat{p}_{Pet}$ under the context e_{Fish}, and close but not equal to the changed state of the ground state $\hat{p}_{Fish}$ under the context e_{Pet}, the difference being due to interference taking place between *Pet* and *Fish* when they combine into *Pet-Fish* (see Section 3). The "changes of state under the influence of a context" and corresponding typicalities behave like the changes of state and corresponding probabilities behave in quantum theory, giving rise to a violation of corresponding fuzzy set and/or classical probability rules. This partly explains the high typicality of *Guppy* in the conjunction *Pet-Fish*, and its normal typicality in *Pet* and *Fish*, and the reason why we identify the Guppy effect as an effect at least partly due to context. There is also an interference effect, as we will see later.

We developed this approach in a formal way, and called the underlying mathematical structure a "State Context Property (SCoP) formalism" (Aerts and Gabora, 2005a). Let *A* denote a concept. In SCoP, *A* is associated with a triple of sets, namely the set Σ of states—we denote states by $p, q, \ldots$, the set $\mathcal{M}$ of contexts, we denote contexts by $e, f, \ldots$, and the set $\mathcal{L}$ of properties—we denote properties by $a, b, \ldots$. The "ground state" $\hat{p}$ of the concept *A* is the state where *A* is not under the influence of any particular context. Whenever *A* is under the influence of a specific context *e*, a change of the state of *A* occurs. In case *A* was in its ground state $\hat{p}$, the ground state changes to a state *p*. The difference between states $\hat{p}$ and *p* is manifested, for example, by the typicality values of different items of the concept, as we have seen in the case of the Guppy effect, and the applicability values of different properties being different in the two states $\hat{p}$ and *p*. Hence, to complete the mathematical

construction of SCoP, also two functions μ and ν are needed. The function $\mu : \Sigma \times \mathcal{M} \times \Sigma \longrightarrow [0, 1]$ is defined such that $\mu(q, e, p)$ is the probability that state p of concept A under the influence of context e changes to state q of concept A. The function $\nu : \Sigma \times \mathcal{L} \longrightarrow [0, 1]$ is defined such that $\nu(p, a)$ is the weight, or normalization of applicability, of property a in state p of concept S. The function μ mainly accounts for typicality measurements, the function ν mainly accounts for applicability measurements. Through these mathematical structures the SCoP formalism captures both "contextual typicality" and "contextual applicability" (Aerts and Gabora, 2005a).

A quantum representation in a complex Hilbert space of data on *Pet* and *Fish* and different states of *Pet* and *Fish* in different contexts was developed (Aerts and Gabora, 2005a), as well as of the concept *Pet-Fish* (Aerts and Gabora, 2005b). Let us deepen the connections between the quantum-theoretic approach to concepts and prototype theory (see also Gabora et al., 2008). This approach can be interpreted in a rather straightforward way as a generalization of prototype theory which mathematically integrates context and formalizes its effects, unlike standard prototype theory. What we call the ground state of a concept can be seen as the prototype of this concept. The conceptual distance of an item from the prototype can be reconstructed from the functions μ and ν in the SCoP formalism. Thus, as long as individual concepts are considered and in the absence of any context, prototype theory can be embodied into the SCoP formalism, and the prototype of a concept A can be represented as its ground state $\hat{p}_A$. However, any context will change this ground state into a new state. An important consequence of this is that when the concept is in this new state, the prototype changes. An intuitive way of understanding this is to consider this new state a new "contextualized prototype." More concretely, the concept *Pet*, when combined with *Fish* in the conjunction *Pet-Fish*, has a new contextualized prototype, which could be called "*Pet* contextualized by Fish." The new state can be thought of as a "contextualized prototype." Hence, this is a prototype-like theory that is capable of mathematically describing the presence and influence of context. From the point of view of conceptual distance, this contextualized prototype will be close to, e.g., *Guppy*.

The interpretation of the SCoP formalism as a contextual prototype theory can be applied not just to conjunctions and disjunctions of two concepts, but also to abstract categories such as *Fruits*. It is very likely that the prototype of *Fruits* is close to, e.g., *Apple*, or *Orange*. But let us now consider the combination *Tropical Fruits*, that is, *Fruits* under the context *Tropical*. It is then reasonable to maintain that the new contextualized prototype of *Tropical Fruits* is closer to, e.g., *Pineapple*, or *Mango*, than to *Apple*, or *Orange*. The introduction of contextualized prototypes within the SCoP formalism enables us to incorporate abstract categories as well as deviations of typicality from fuzzy set behavior.

Another interesting aspect of this approach to prototype theory comes to light if we consider again the conceptual combination *Pet-Fish*. It is reasonable that the prototypes of *Pet* and *Fish*—ground states $\hat{p}_{Pet}$ and $\hat{p}_{Fish}$—interfere in *Pet-Fish* whenever the typicality of an item, e.g., *Guppy*, is measured

with respect to *Pet-Fish*. This sentence cannot, however, be made more explicit in the absence of a concrete quantum-theoretic representation of typicality measurements of items with respect to concepts and their combinations. Indeed, interference and superposition effects can be precisely formalized in such quantum representation. This will be the content of Sections 3 and 4.

3. A HILBERT SPACE MODELING OF MEMBERSHIP MEASUREMENTS

One can gain insight into how people combine concepts by gathering data on "membership weights" and "typicalities." To obtain data on "typicalities," participants are given a concept, and a list of instances or items, and asked to estimate their typicality on a Likert scale. In other experiments participants are asked to choose which instance they consider most typical of the concept. Averages of these estimates or relative frequencies of the picked items give rise to values representing the typicalities of the respective items. A membership weight is obtained by asking participants to estimate the membership of specific items with respect to a concept. This estimation can be quantified using the 7-point Likert scale and then converted into a relative frequency, and then into a probability called the "membership weight."

Hampton used membership weights instead of typicalities (Hampton, 1988b), because all you can do with typicalities is fuzzy set type calculations: the minimum rule of fuzzy sets for conjunction or the maximum rule for disjunction. This approach has many serious shortcomings; indeed the Pet-Fish problem could not be addressed by it. More serious failures are revealed by membership weight data. Hampton measured "membership weights" and "degrees of non-membership or membership," making these two measurements in one experiment. More specifically, Hampton's experiment generates magnitude data, measuring the "degree of membership or non-membership" using a 7-point Likert scale providing $-1, -2, -3$ for degrees of non-membership, $1, 2, 3$ for degrees of membership and, 0 for borderline cases. From the same experiment membership weight data are obtained, with 8 possible triplets $[\pm, \pm, \pm]$ per item. Each triplet indicating with a $+$ whether the participant considered item k to be a member of the first category (A), the second category (B) and the third disjunction category (A or B), and with a $-$ respectively otherwise. In the present Hilbert space model we use the "degree of membership or non-membership" values obtained by Hampton, add $+3$ to them to make them all non-negative, sum them, and divide each one by this sum. Since there are 24 items in total, in this way we get a set of 24 values in the interval $[0,1]$, that sum up to 1. We will use these values as a substitute for membership collapse probabilities.

Let us first explain how we arrive at the membership collapse probabilities as a consequence of a measurement, and why we can use the above-mentioned calculated values of "degree of membership or non-membership" as substitutes. Suppose that instead of using the data obtained by Hampton, we performed the following experiment. For each pair of concepts and their combination we ask the participant to select one and only one

item that they consider the best choice for membership. Then we calculate for each of the 24 items the relative frequency of its appearance. These relative frequencies are 24 values in the interval [0,1] summing up to 1, and their limits for increasing numbers of participants represent the probabilities for each item to be chosen as the best member. These probabilities are what in a quantum model are called the "membership collapse probabilities." Of course, the above described experiment to determine the membership collapse probabilities has not been performed. However, the values calculated from Hampton's measurement of "degree of membership or non-membership," after renormalization as explained above, are expected to correlate with what these membership collapse probabilities would be if they were measured. This is why we use them as substitutes for the membership collapse probabilities in our quantum model. As we will see when we construct the quantum model, the exact values of the substitutes for the membership collapse probabilities are not critical. Thus, if we can model the substitutes for the membership collapse probabilities calculated from Hampton's data, we can also model the actual membership collapse probabilities (the data we would have if the experiment had been done).

So, we repeat, in **Table 1**, Hampton's experimental data (Hampton, 1988b) have been converted into relative frequencies. The "degrees of non-membership and degrees of membership" give rise to $\mu_k(X)$ and now stand for the probability of concepts *Fruits* ($X = A$), *Vegetables* ($X = B$) and *Fruits or Vegetables* ($X =$ "A or B") to collapse to the item k, and thus add up to 1, that is,

$$\sum_{k=1}^{24} \mu_k(A) = \sum_{k=1}^{24} \mu_k(B) = \sum_{k=1}^{24} \mu_k(A \text{ or } B) = 1 \qquad (2)$$

for the 24 items. The quantum model for concepts and their disjunction in complex Hilbert space is developed by building appropriate state vectors and projection operators for a given ontology of 24 items of two more abstract "container" concepts.

In our model, the Hilbert space is a complex n-dimensional $\mathbb{C}^n$, in which state vectors are n-dimensional complex numbered vectors. We use the "bra-ket" notation – respectively $\langle \cdot |$ and $| \cdot \rangle$— for vector states (see the Appendix for further explanation). The complex conjugate transpose of the $| \cdot \rangle$ ket-vector (nx1 dim.) is the $\langle \cdot |$ bra-vector (1xn dim.). Projectors and operators are then combined as matrices $| \cdot \rangle \langle \cdot |$, while scalars are obtained by inner product $\langle \cdot | \cdot \rangle$. We represent the measurement, consisting in the question "Is item k a good example of concept X?," by means of an orthogonal projection operator M_k. Each self-adjoint operator in the Hilbert space $\mathcal{H}$ has a spectral decomposition on $\{M_k | k = 1, \ldots, 24\}$, where each M_k is the projector corresponding to item k from the list of 24 items in **Table 1**. A priori we set no restrictions to the dimension of the complex Hilbert space, and thus neither to the projection space of the operators M_k. Each separate concept *Fruits* and *Vegetables* is now represented by its proper state vector $|A\rangle$ and $|B\rangle$ respectively, while their disjunction *Fruits or Vegetables* is realized by their equiponderous superposition $\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$. It is precisely this feature of the

model—its ability to represent combined concepts as superposed states—that provides the interferential composition of what could not be classically composed using sets.

Following the standard rule of average outcome values of quantum theory, the probabilities, $\mu_k(A)$, $\mu_k(B)$ and $\mu_k(A \text{ or } B)$ are given by:

$$\mu_k(A) = \langle A | M_k | A \rangle \qquad (3)$$

$$\mu_k(B) = \langle B | M_k | B \rangle \qquad (4)$$

$$\mu_k(A \text{ or } B) = \frac{\langle A | + \langle B |}{\sqrt{2}} M_k \frac{|A\rangle + |B\rangle}{\sqrt{2}} \qquad (5)$$

After a straightforward calculation, the membership probability expression $\mu_k(A \text{ or } B)$ becomes:

$$\begin{aligned}
\mu_k(A \text{ or } B) &= \frac{1}{2} \left(\langle A | M_k | A \rangle + \langle A | M_k | B \rangle + \langle B | M_k | A \rangle \right. \\
&\quad \left. + \langle B | M_k | B \rangle \right) \\
&= \frac{1}{2} \left(\mu_k(A) + \mu_k(B) \right) + \Re \langle A | M_k | B \rangle \qquad (6)
\end{aligned}$$

where $\Re$ takes the real part of $\langle A | M_k | B \rangle$. This expression shows the contribution of the interference term $\Re \langle A | M_k | B \rangle$ in $\mu_k(A \text{ or } B)$ with respect to the "classical average" term $\frac{1}{2} \left(\mu_k(A) + \mu_k(B) \right)$. It consists of the real part of the complex probability amplitude of the k-th item in *Vegetables* (concept $|B\rangle$) to be the one in *Fruits* (concept $|A\rangle$).

The quantum concept model imposes the orthogonality of the state vectors corresponding to different concepts. Therefore, we have for the states of *Fruits* and *Vegetables*,

$$\langle A | B \rangle = 0. \qquad (7)$$

Each different item of the projector M_k also provides an orthogonal projection space. Since the conceptual disjunction *Fruits or Vegetables* spans a subspace of 2 dimensions in the complex Hilbert space (along the rays of $|A\rangle$ and $|B\rangle$), we set forth the possibility for a complex 2-dimensional subspace for each item. This brings the dimension of the complex Hilbert space to 48. However, we will choose the unit vectors of these subspaces in such a way as to eliminate redundant dimensions whenever possible. Each category vector is built on orthogonal unit vectors, defined by the projection operators M_k. i.e., we define $|e_k\rangle$ the unit vector on $M_k|A\rangle$, and define $|f_k\rangle$ the unit vector on $M_k|B\rangle$. Thus, each item is now represented by a vector spanned by $|e_k\rangle$ and $|f_k\rangle$. Due the orthogonality of the projectors M_k, we have

$$\langle e_k | f_l \rangle = \delta_{kl} c_k e^{i\gamma_k} \qquad (8)$$

where the Kronecker $\delta_{kl} = 1$ for same indices and zero otherwise, i.e., different item states are orthogonal as well. And c_k expresses the angle between the two unit vectors $|e_k\rangle$ and $|f_k\rangle$ of each 2-dimensional subspace of item k. Notice that should some c_k be 1, then the required dimension of the complex Hilbert space diminishes by 1, since the vectors $|e_k\rangle$ and $|f_k\rangle$ then coincide— a property that we will use to minimize the size of the required

TABLE 1 | Membership collapse probability values $\mu_k(X)$ of 24 items for the categories _Fruits_, _Vegetables_, and _Fruits or Vegetables_ (Hampton, 1988b).

		A = FRUITS	B = VEGETABLES					
k	Item	$\mu_k(A)$	$\mu_k(B)$	$\mu_k(A$ or $B)$	λ_k	λ-rank	ϵ_k	ϕ_k
1	Almond	0.0359	0.0133	0.0269	0.0217	16	+1	84.0°
2	Acorn	0.0425	0.0108	0.0249	0.0214	17	−1	−94.5°
3	Peanut	0.0372	0.0220	0.0269	0.0285	10	−1	−95.4°
4	Olive	0.0586	0.0269	0.0415	0.0397	9	+1	91.9°
5	Coconut	0.0755	0.0125	0.0604	0.0260	12	+1	57.7°
6	Raisin	0.1026	0.0170	0.0555	0.0415	7	+1	95.9°
7	Elderberry	0.1138	0.0170	0.0480	0.0404	8	−1	−113.3°
8	Apple	0.1184	0.0155	0.0688	0.0428	5	+1	87.6°
9	Mustard	0.0149	0.0250	0.0146	0.0186	19	−1	−105.9°
10	Wheat	0.0136	0.0255	0.0165	0.0184	20	+1	99.3°
11	Root Ginger	0.0157	0.0323	0.0385	0.0172	22	+1	49.9°
12	Chili Pepper	0.0167	0.0446	0.0323	0.0272	11	−1	−86.4°
13	Garlic	0.0100	0.0301	0.0293	0.0146	23	−1	−57.6°
14	Mushroom	0.0140	0.0545	0.0604	0.0087	24	+1	18.5°
15	Watercress	0.0112	0.0658	0.0482	0.0253	13	−1	−69.1°
16	Lentils	0.0095	0.0713	0.0338	0.0252	14	+1	104.7°
17	Green Pepper	0.0324	0.0788	0.0506	0.0503	4	−1	−95.7°
18	Yam	0.0533	0.0724	0.0541	0.0615	3	+1	98.1°
19	Tomato	0.0881	0.0679	0.0688	0.0768	1	+1	98.5°
20	Pumpkin	0.0797	0.0713	0.0579	0.0733	2	−1	−103.5°
21	Broccoli	0.0143	0.1284	0.0642	0.0423	6	−1	−99.5°
22	Rice	0.0140	0.0412	0.0248	0.0238	15	−1	−96.7°
23	Parsley	0.0155	0.0266	0.0308	0.0178	21	−1	−61.1°
24	Black Pepper	0.0127	0.0294	0.0222	0.01929	18	+1	86.7°

Notice also the membership collapse probabilities for Mustard _and_ Pumpkin _still show the mark of double underextension of the disjunction. Membership collapse probability data with_ $\delta\mu \approx 10^{-4}$ _entail phase data_ $\delta\phi \approx 2 \cdot 10^{-1}$ _and lambda data_ $\delta\lambda \approx 4 \cdot 10^{-4}$.

Hilbert space. Should c_k be different from 1, then $|e_k\rangle$ and $|f_k\rangle$ span a subspace of 2 dimensions. The state vectors $|A\rangle$ and $|B\rangle$ of the concepts can then be expressed as a superposition of the vectors $|e_k\rangle$ and $|f_k\rangle$ for the items:

$$|A\rangle = \sum_{k=1}^{24} a_k e^{i\alpha_k} |e_k\rangle, \quad |B\rangle = \sum_{k=1}^{24} b_k e^{i\beta_k} |f_k\rangle \quad (9)$$

where $a_k, b_k, \alpha_k, \beta_k \in \mathbb{R}$.

We can express their inner product as follows:

$$\langle A|B\rangle = \left(\sum_{k=1}^{24} a_k e^{-i\alpha_k} \langle e_k| \right) \left(\sum_{l=1}^{24} b_l e^{i\beta_l} |f_l\rangle \right)$$

$$= \sum_{k=1}^{24} a_k b_k c_k e^{i(\beta_k - \alpha_k + \gamma_k)} = \sum_{k=1}^{24} a_k b_k c_k e^{i\phi_k}$$

where we have defined phase ϕ_k as $\phi_k := \beta_k - \alpha_k + \gamma_k$ in the last step. The membership probabilities given in Equations (3 and 4) and the interference terms in Equation (6) can be expanded on the projection spaces of the items:

$$\mu_k(A) = \left(\sum_{l=1}^{24} a_l e^{-i\alpha_l} \langle e_l| \right) (a_k e^{i\alpha_k} |e_k\rangle) = a_k^2 \quad (10)$$

$$\mu_k(B) = \left(\sum_{l=1}^{24} b_l e^{-i\beta_l} \langle f_l| \right) (b_k e^{i\beta_k} |f_k\rangle) = b_k^2 \quad (11)$$

$$\langle A|M_k|B\rangle = \left(\sum_{l=1}^{24} a_l e^{-i\alpha_l} \langle e_l| \right) M_k| \left(\sum_{m=1}^{24} b_m e^{i\beta_m} |f_m\rangle \right)$$

$$= a_k b_k e^{i(\beta_k - \alpha_k)} \langle e_k|f_k\rangle = a_k b_k c_k e^{i\phi_k} \quad (12)$$

Notice that the phase of the k-th component of the conceptual disjunction is not at play in the interference term $\langle A|M_k|B\rangle$ (Equation 6). Taking the real part of the interference term in Equation (12), we can rewrite the membership probability of the disjunction in Equation (6) as follows:

$$\mu_k(A \text{ or } B) = \frac{1}{2} \left(\mu_k(A) + \mu_k(B) \right) + c_k \sqrt{\mu_k(A)\mu_k(B)} \cos\phi_k \quad (13)$$

Rearranging this equation we now choose ϕ_k must satisfy

$$\cos\phi_k = \frac{\mu_k(A \text{ or } B) - \frac{1}{2}(\mu_k(A) + \mu_k(B))}{c_k\sqrt{\mu_k(A)\mu_k(B)}} \qquad (14)$$

Since all the membership probabilities on the right side are fixed, the only remaining free parameters are the coefficients c_k. These parameters must now be tuned in order to satisfy the orthogonality of $|A\rangle$ and $|B\rangle$. Using the expansion on the unit vector sets $\{|e_k\rangle\}$, $\{|f_k\rangle\}$ we obtain for their orthogonality (Equation 7):

$$\sum_{k=1}^{24} c_k\sqrt{\mu(A)_k\mu(B)_k}\cos\phi_k = 0, \qquad (15)$$

$$\sum_{k=1}^{24} c_k\sqrt{\mu(A)_k\mu(B)_k}\sin\phi_k = 0. \qquad (16)$$

The "cosine sum" (Equation 15) is automatically satisfied due to the definition of $\cos\phi_k$ and the normalization of membership probabilities (Equation 2). This can be seen by substituting the expression of $\cos\phi_k$ in Equation (14) and then applying the normalization condition of the membership probabilities (Equation 2). The "sine sum" equation still needs to be satisfied. With the defining relation (Equation 14) of ϕ_k, and $\sin\phi_k = \epsilon_k\sqrt{1-\cos^2\phi_k}$, where $\epsilon_k = \pm$ provides the sign, this becomes[1]

$$\sum_{k=1}^{24} \epsilon_k\sqrt{c_k^2\mu_k(A)\mu_k(B) - (\mu_k(A \text{ or } B) - \frac{1}{2}(\mu_k(A) + \mu_k(B)))^2}$$
$$= 0. \qquad (17)$$

In order to satisfy this equation a simple algorithm was devised (Aerts, 2009a). For convenience of notation we denote the square root expression, with $c_k = 1$, by a separate symbol:

$$\lambda_k := \sqrt{\mu_k(A)\mu_k(B) - (\mu_k(A \text{ or } B) - \frac{1}{2}(\mu_k(A) + \mu_k(B)))^2}. \qquad (18)$$

First, we order the values λ_k from large to small and then assign a sign ϵ_k to each of them in such a way that each next partial sum (increasing index) remains smallest. The λ-ranking with corresponding values have been tabulated in **Table 1**. We assign index m to the item with the largest λ-value. In the present case, the item *Tomato* has the largest value, 0.07679.

We now adopt an optimized complex Hilbert space for our model in which $c_k = 1$ ($k \neq m$), which reduces the space to 25 complex dimensions. We again note that all items except *Tomato* receive a 1-dimensional complex subspace, while *Tomato* is represented by a 2-dimensional subspace. The "sine sum" equation in Equation (17) can be written as

$$\sum_{k=1,k\neq m}^{24} \epsilon_k\lambda_k + \epsilon_m\sqrt{\begin{array}{c}c_m^2\mu_m(A)\mu_m(B) - (\mu_m(A \text{ or } B) \\ -\frac{1}{2}(\mu_m(A) + \mu_m(B)))^2\end{array}} = 0. \qquad (19)$$

---
[1]The cosine value only defines the phase up to its absolute value $|\phi_k|$. Thus, the sign of the sine value is undefined. If $\epsilon_k = -1$, then $\phi_k = -|\phi_k|$.

Next, we define the partial sum of the λ_k according a scheme of signs ϵ_k such that from large to small the next $\epsilon_k\lambda_k$ is added to make the sum smaller but not negative.

$$S_j = \sum_{size\ ordered\ \lambda_i}^{j} \epsilon_i\lambda_i \qquad (20)$$

$$S_{j+1} = S_j - \lambda_{j+1} \text{ and } \epsilon_{j+1} = -1, \text{ if } S_j - \lambda_{j+1} \geq 0 \quad (21)$$
$$= S_j + \lambda_{j+1} \text{ and } \epsilon_{j+1} = +1, \text{ if } S_j - \lambda_{j+1} < 0 \quad (22)$$

The first summand is thus λ_m, with $\epsilon_m = +1$. Finally this procedure leads to

$$S_{24} = \sum_{k=1}^{24} \epsilon_k\lambda_k \geq 0$$

In the *Fruits* and *Vegetables* example with membership probability data in **Table 1**, this procedure gives:

$$S_{24} = 0.0154 \qquad (23)$$

In general the "sine sum" equation then becomes

$$S_{24} - \lambda_m + \sqrt{\begin{array}{c}c_m^2\mu_m(A)\mu_m(B) - (\mu_m(A \text{ or } B) \\ -\frac{1}{2}(\mu_m(A) + \mu_m(B)))^2\end{array}} = 0. \qquad (24)$$

From which we can fix c_m, the remaining c_k not equal to 1:

$$c_m = \sqrt{\frac{(S_{24} - \lambda_m)^2 + (\mu_m(A \text{ or } B) - \frac{1}{2}(\mu_m(A) + \mu_m(B)))^2}{\mu_m(A)\mu_m(B)}} \qquad (25)$$

In the present example we obtain the value $c_m = 0.8032$. We thus have fixed the inner product—or "angle"—of the vectors $|e_m\rangle$ and $|f_m\rangle$, and can now write an explicit representation in the canonical 25 dimensional complex Hilbert space $\mathbb{C}^{25}$. We can take $M_k(\mathcal{H})$ to be rays of dimension 1 for $k \neq m$, and $M_m(\mathcal{H})$ to be a 2-dimensional plane spanned by the vectors $|e_m\rangle$ and $|f_m\rangle$.

We let the space $\mathbb{C}^{25}$ be spanned on the canonical base $\{\mathbf{1}_i\}$, $i \in [1\dots25]$. All items $k \neq m$ are represented by the respective $\mathbf{1}_i$. While for $k = m$ we express the projections of $|A\rangle$ and $|B\rangle$ by $M_m(\mathcal{H})$ accordingly on $\mathbf{1}_m$ and $\mathbf{1}_{25}$

$$a_me^{i\alpha_m}|e_m\rangle = \tilde{a}_me^{i\alpha_{m_1}}\mathbf{1}_m + \tilde{a}_{25}e^{i\alpha_{m_2}}\mathbf{1}_{25} \qquad (26)$$
$$b_me^{i\beta_m}|f_m\rangle = \tilde{b}_me^{i\beta_{m_1}}\mathbf{1}_m + \tilde{b}_{25}e^{i\beta_{m_2}}\mathbf{1}_{25} \qquad (27)$$

with $\tilde{a}_m, \tilde{b}_m, \tilde{a}_{25}, \tilde{b}_{25} \in \mathbb{R}$ to be specified. The parameters in these expressions should satisfy the inner product (Equation 8) for $k, l = m$

$$a_mb_m\langle e_m|f_m\rangle = \tilde{a}_m\tilde{b}_me^{-i(\alpha_{m_1}-\beta_{m_1})} + \tilde{a}_{25}\tilde{b}_{25}e^{-i(\alpha_{m_2}-\beta_{m_2})}, (28)$$
$$= a_mb_mc_me^{i(\gamma_m-\alpha_m+\beta_m)} \qquad (29)$$

and the probability weights for $k = m$

$$a_m^2 = \tilde{a}_m^2 + \tilde{a}_{25}^2, \qquad (30)$$
$$b_m^2 = \tilde{b}_m^2 + \tilde{b}_{25}^2. \qquad (31)$$

Finally, the representation of all vectors of the items can now be rendered explicit by simply choosing $\alpha_k = \gamma_k = 0$, and thus $\beta_k = \phi_k$, $\forall k$. A further simplification for *Tomato* is done by setting $\tilde{a}_{25} = 0$, which also allows free choice of $\beta_{m_2} = 0$. Then $\tilde{a}_m = a_m$ and $\tilde{b}_m = b_m c_m$, and $\tilde{b}_{25} = b_m\sqrt{(1 - c_m^2)}$. We have rendered explicit these membership probabilities and phases in **Table 1**. Thus we can write the vectors $|A\rangle$ and $|B\rangle$ in $\mathbb{C}^{25}$ Hilbert space corresponding to the categories *Fruits* and *Vegetables* respectively.

$$|A\rangle = (0.1895, 0.2062, 0.1929, 0.2421, 0.2748, 0.3203, 0.3373,$$
$$0.3441, 0.1221, 0.1166, 0.1253, 0.1292, 0.1000, 0.1183,$$
$$0.1058, 0.0975, 0.1800, 0.2309, 0.2968, 0.2823, 0.1196,$$
$$0.1183, 0.1245, 0.1127, 0.0000) \qquad (32)$$

$$|B\rangle = (0.1153e^{i84.0°}, 0.1039e^{-i94.5°}, 0.1483e^{-i95.4°}, 0.1640e^{i91.9°},$$
$$0.1118e^{i57.7°}, 0.1304e^{i95.9°}, 0.1304e^{-i113.3°}, 0.1245e^{i87.6°},$$
$$0.1581e^{-i105.9°}, 0.1597e^{i99.3°}, 0.1797e^{i49.9°}, 0.2112e^{-i86.4°},$$
$$0.1735e^{-i57.6°}, 0.2335e^{i18.5°}, 0.2565e^{-i69.1°}, 0.2670e^{i104.7°},$$
$$0.2807e^{-i95.7°}, 0.2691e^{i98.0°}, 0.2606e^{i96.8°}, 0.2670e^{-i103.5°},$$
$$0.3583e^{-i99.5°}, 0.2030e^{-i96.7°}, 0.1631e^{-i61.1°}, 0.1715e^{i86.7°},$$
$$0.1552). \qquad (33)$$

This completes the quantum model for the membership probability of items with respect to *Fruits*, *Vegetables* and *Fruits or Vegetables*. It captures the enigmatic aspects of conceptual overextension and underextension identified in Hampton (1988b), explaining them in terms of genuine quantum phenomena.

Recalling the terminology adopted in Section 2, the unit vectors $|A\rangle$ and $|B\rangle$ in Equations (32) and (33) represent the ground states of the concepts *Fruits* and *Vegetables*, respectively. Equivalently, these unit vectors represent the prototypes of the concepts *Fruits* and *Vegetables* in prototype theory. The unit vector $\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$ instead represents the "contextualized prototype" obtained by combining the prototypes of *Fruits* and *Vegetables* in the disjunction *Fruits or Vegetables*. If one now looks at Equation (6), one sees that the prototypes *Fruits* and *Vegetables* interfere in the disjunction *Fruits or Vegetables*, and the term $\Re\langle A|M_k|B\rangle$ in Equation (6) specifies how much interference is present when the membership probability of k is measured.

4. AN ILLUSTRATION OF INTERFERING PROTOTYPES

In this section we provide an illustration of contextual interfering prototypes. It is not a complete mathematical representation as presented in Section 3 but, rather, an illustration that can help the reader with a non-technical background to have an intuitive picture of what a contextual prototype is and how contextual prototypes interfere. Consider the concepts *Fruits*, *Vegetables* and their disjunction *Fruits or Vegetables*. The contextual prototype of *Fruits* can be represented by the x-axis of a plane surrounded by a cloud containing items, features, etc.—all the contextual elements

connected with the prototype of *Fruits*. Similarly, the contextual prototype of *Vegetables* can be represented by the y-axis of the same plane surrounded by a cloud containing items, features, etc.—all the contextual elements connected with the prototype of *Vegetables*. How can we represent the contextual prototype of the disjunction *Fruits or Vegetables*? Although as we have seen it cannot be represented in traditional fuzzy set theory, it can be represented in terms of waves, with peaks and troughs. Indeed, waves can be summed up in such a way that peaks and troughs of the combined wave reproduce overextension and underextension of the data. In other words, waves provide an intuitive geometric illustration of the interference taking place when contextual prototypes are combined in concept combination as discussed in Section 3. For example, let us demonstrate the interference of the item *Almond* when its membership probability with respect to the disjunction *Fruits or Vegetables* is calculated based on its membership probabilities for *Fruits* and for *Vegetables*. The membership probabilities for the categories *Fruits*, *Vegetables* and *Fruits or Vegetables* have been calculated from the Hampton's data and are reported in **Table 1**.

The idea of an illustration would be to show that in addition to "fuzziness" (as modeled using a fuzzy set-theoretic approach) there is a "wave structure." How can we graphically represent this "wave structure" of a concept? We start from the standard interference formula of quantum theory, which is the following. For an arbitrary item k we have

$$\mu_k(A \text{ or } B) = \frac{1}{2}(\mu_k(A) + \mu_k(B)) + c_k\sqrt{\mu_k(A)\mu_k(B)}\cos\phi_k. \qquad (34)$$

Now, we have

$$\phi_k = \beta_k - \alpha_k + \gamma_k \qquad (35)$$

where α_k is the phase angle connected with $\mu_k(A)$, β_k the phase angle connected to $\mu_k(B)$, and γ_k the phase angle connected to $\langle A|M_k|B\rangle$. This has not yet been emphasized but if one analyses the rest of the construction in Hilbert space, it is possible to see that one can always choose $\gamma_k = 0$, which means that, with this choice, ϕ_k becomes the difference in phases β_k and α_k.

This is all we need to represent the "wave" nature of a concept in a manner analogous to that of quantum theory. Indeed, it is the "phase difference" between the waves—their phases being α_k and β_k respectively – that we attach to $\mu(A)_k$ and $\mu(B)_k$. They determine, together with the membership probabilities $\mu(A)_k$ and $\mu(B)_k$ the interference that gives rise to the measured data for $\mu(A \text{ or } B)_k$.

The choice of the c_k is such that only for the biggest value of λ_k, which in this case of *Tomato*, the c_k is chosen different from 1. The only choice different from 1, for *Tomato*, still does not influence the fact that ϕ_k is the difference between β_k and α_k, when we decide to choose $\gamma_k = 0$. Let us consider for example the first item *Almond* of the list of 24 in **Table 1**. We have

$$\mu(A)_1 = 0.0359 \qquad (36)$$
$$\mu(B)_1 = 0.0133 \qquad (37)$$
$$\mu(A \text{ or } B)_1 = 0.0269 \qquad (38)$$

These are the data measured by Hampton, and also what exists for the concepts *Fruits*, *Vegetables* and their combination *Fruits or Vegetables* with respect to membership probability of the item *Almond* in the realm where fuzzy set probability appears. These are the values that do not fit into a model in this realm, and for which a wave-like realm underneath is necessary. Calculating the angle ϕ_1 we get

$$\phi_1 = 84.0^\circ \tag{39}$$

(see **Table 1**). This angle is the result of a wave being present underneath the fuzzy, probability realm for $\mu(A)_1$ and $\mu(B)_1$, such that both waves give rise to a difference in phase—where the crests of one wave meet the troughs of the other—which is equal to $\beta_1 - \alpha_1$, and is the value of ϕ_1. This can be represented graphically by attaching a wave pattern to $\mu(A)_1$ and another one to $\mu(B)_1$, such that both have a phase difference of 84.0°—see also **Figure 1**.

Let us apply quantum theory to each of the items apart. Each item k has a Schrödinger wave function vibrating in the neighborhood of A, another one vibrating in the neighborhood of B and a third vibrating in the neighborhood of "A or B," and they are related by superposition. We have:

$$\psi_k^A = \sqrt{\mu_k(A)}e^{i\alpha_k} \tag{40}$$

$$\psi_k^B = \sqrt{\mu_k(B)}e^{i\beta_k} \tag{41}$$

$$\psi_k^{AorB} = \sqrt{\mu_k(A \text{ or } B)}e^{i\delta_k} \tag{42}$$

In each case, this gives us the membership probabilities. Squaring (multiplying by its complex conjugate), we have

$$\langle\psi_k^A|\psi_k^A\rangle = (\psi_k^A)^*(\psi_k^A) = \left(\sqrt{\mu_k(A)}e^{i\alpha_k}\right)^*\left(\sqrt{\mu_k(A)}e^{i\alpha_k}\right)$$

$$= \left(\sqrt{\mu_k(A)}e^{-i\alpha_k}\right)\left(\sqrt{\mu_k(A)}e^{i\alpha_k}\right) = \mu_k(A)e^{i(\alpha-\alpha)} = \mu_k(A) \tag{43}$$

$$\langle\psi_k^B|\psi_k^B\rangle = (\psi_k^B)^*(\psi_k^B) = \left(\sqrt{\mu_k(B)}e^{i\beta_k}\right)^*\left(\sqrt{\mu_k(B)}e^{i\beta_k}\right)$$

$$= \left(\sqrt{\mu_k(B)}e^{-i\beta_k}\right)\left(\sqrt{\mu_k(B)}e^{i\beta_k}\right) = \mu_k(B)e^{i(\beta-\beta)} = \mu_k(B) \tag{44}$$

$$\langle\psi_k^{AorB}|\psi_k^{AorB}\rangle = (\psi_k^{AorB})^*(\psi_k^{AorB})$$

$$= \left(\sqrt{\mu_k(A \text{ or } B)}e^{i\delta_k}\right)^*\left(\sqrt{\mu_k(A \text{ or } B)}e^{i\delta_k}\right)$$

$$= \left(\sqrt{\mu_k(A \text{ or } B)}e^{-i\delta_k}\right)\left(\sqrt{\mu_k(A \text{ or } B)}e^{i\delta_k}\right) = \mu_k(A \text{ or } B)e^{i(\delta-\delta)}$$

$$= \mu_k(A \text{ or } B) \tag{45}$$

If we write the quantum superposition equation for each item we get

$$\frac{1}{\sqrt{2}}(\psi_k^A + \psi_k^B) = \psi_k^{AorB} \tag{46}$$

$$\Leftrightarrow \frac{1}{\sqrt{2}}\left(\sqrt{\mu(A)_k}e^{i\alpha_k} + \sqrt{\mu(B)_k}e^{i\beta_k}\right) = \sqrt{\mu(A \text{ or } B)_k}e^{i\delta}_k \tag{47}$$

where $\frac{1}{\sqrt{2}}$ is a normalization factor. It is the squaring (i.e., multiplying each with its complex conjugate) that gives rise to the interference equation. Let us do this explicitly to see it.

FIGURE 1 | Interference of items *Almond*, *Acorn* and *Coconut* in the concept *Fruits or Vegetables*. Elementary oscillatory waves $\sqrt{\mu_k(A)}\cos(x)$ and $\sqrt{\mu_k(B)}\cos(x + \phi_k)$ are associated to the components of each given item in *Fruits* and *Vegetables* respectively. The weight amplitude of the item in the disjunction *Fruits or Vegetables* emerges at the origin of $\left(\sqrt{\mu_k(A)}\cos(x) + \sqrt{\mu_k(B)}\cos(x + \phi_k)\right)/\sqrt{2}$.
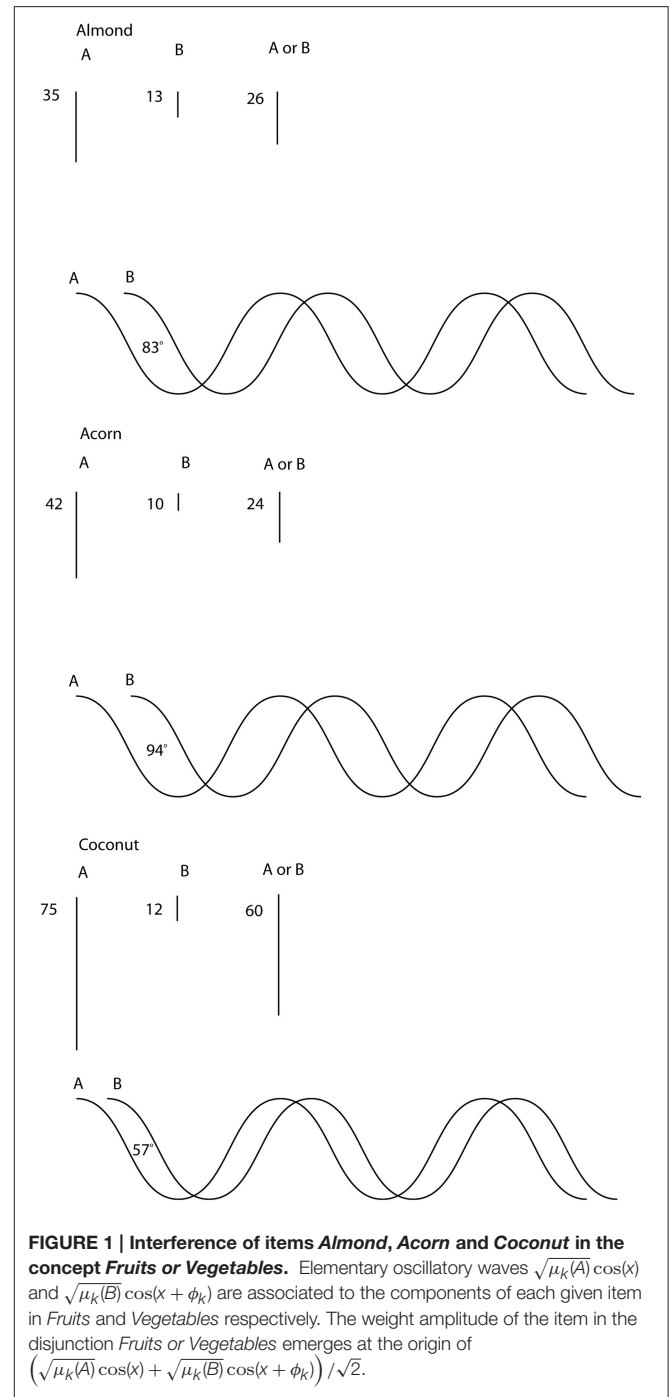
First we multiply the left hand side with its complex conjugate. We do the multiplication explicitly writing each step of it, to see well how the interference formula appears. Hence, we have

$$\left(\frac{1}{\sqrt{2}}\left(\sqrt{\mu_k(A)}e^{i\alpha_k} + \sqrt{\mu_k(B)}e^{i\beta_k}\right)\right)^*\left(\frac{1}{\sqrt{2}}\left(\sqrt{\mu_k(A)}e^{i\alpha_k}\right.\right.$$

$$\left.\left. + \sqrt{\mu_k(B)}e^{i\beta_k}\right)\right)$$

$$= \left(\frac{1}{\sqrt{2}} \left(\sqrt{\mu_k(A)} e^{-i\alpha_k} + \sqrt{\mu_k(B)} e^{-i\beta_k} \right) \right) \left(\frac{1}{\sqrt{2}} \left(\sqrt{\mu_k(A)} e^{i\alpha_k} \right.\right.$$
$$\left.\left. + \sqrt{\mu_k(B)} e^{i\beta_k} \right) \right)$$

$$= \frac{1}{2} \left(\sqrt{\mu_k(A)} e^{-i\alpha_k} + \sqrt{\mu_k(B)} e^{-i\beta_k} \right) \left(\sqrt{\mu_k(A)} e^{i\alpha_k} + \sqrt{\mu_k(B)} e^{i\beta_k} \right)$$

$$= \frac{1}{2} \left(\sqrt{\mu_k(A)} e^{-i\alpha_k} \cdot \sqrt{\mu_k(A)} e^{i\alpha_k} + \sqrt{\mu_k(A)} e^{-i\alpha_k} \cdot \sqrt{\mu_k(B)} e^{i\beta_k} \right.$$
$$\left. + \sqrt{\mu_k(B)} e^{-i\beta_k} \cdot \sqrt{\mu_k(A)} e^{i\alpha_k} + \sqrt{\mu_k(B)} e^{-i\beta_k} \cdot \sqrt{\mu_k(B)} e^{i\beta_k} \right)$$

$$= \frac{1}{2} (\mu_k(A) e^{i(\alpha_k - \alpha_k)} + \sqrt{\mu_k(A)\mu_k(B)} e^{i(\beta_k - \alpha_k)}$$
$$+ \sqrt{\mu_k(A)\mu_k(B)} e^{-i(\beta_k - \alpha_k)} + \mu_k(B) e^{i(\beta_k - \beta_k)})$$

we use now that $e^{i(\alpha_k - \alpha_k)} = e^0 = 1$, $e^{i(\beta_k - \beta_k)} = e^0 = 1$, $e^{i(\beta_k - \alpha_k)} = \cos(\beta_k - \alpha_k) + i\sin(\beta_k - \alpha_k)$ and $e^{-i(\beta_k - \alpha_k)} = \cos(\beta_k - \alpha_k) - i\sin(\beta_k - \alpha_k)$, to get to the following

$$= \frac{1}{2} (\mu_k(A) + \sqrt{\mu_k(A)\mu_k(B)}(\cos(\beta_k - \alpha_k) + i\sin(\beta_k - \alpha_k))$$
$$+ \sqrt{\mu_k(A)\mu_k(B)}(\cos(\beta_k - \alpha_k) - i\sin(\beta_k - \alpha_k)) + \mu_k(B))$$

see that the term in $i\sin(\beta_k - \alpha_k)$ cancels, to get

$$= \frac{1}{2} (\mu_k(A) + 2\sqrt{\mu_k(A)\mu_k(B)}\cos(\beta_k - \alpha_k) + \mu_k(B))$$
$$= \frac{1}{2} (\mu_k(A) + \mu_k(B)) + \sqrt{\mu_k(A)\mu_k(B)}\cos(\beta_k - \alpha_k) \quad (48)$$

Let is multiply now the right hand sight of Equation (46) with its complex conjugate. This gives

$$= \left(\sqrt{\mu_k(A \text{ or } B)} e_k^{i\delta} \right)^* \left(\sqrt{\mu_k(A \text{ or } B)} e_k^{i\delta} \right) = \mu_k(A \text{ or } B) \quad (49)$$

Hence, we get, as a consequence of squaring (Equation 46), exactly our interference formula

$$\frac{1}{2} (\mu_k(A) + \mu_k(B)) + \sqrt{\mu_k(A)\mu_k(B)}\cos(\beta_k - \alpha_k) = \mu_k(A \text{ or } B) \quad (50)$$

Note that the difference in phase $\beta_k - \alpha_k$ between the waves connected with the item k and A and the item k and B is what generates the interference. The new wave connected to the item k and A or B, of which the phase is δ is not influenced by it, is the amplitude of this new wave which is affected. This is the reason that interference is visible in the realm where the fuzzy nature appears, while it is provoked by the realm where the waves occur.

We put forward this "wave nature" aspect of concepts not just as an illustration, but to help the reader understand the manner in which such an underlying wave structure increases substantially the possible ways in which concepts can interact, as compared to the interaction possibilities in a modeling with fuzzy set structures. Of course the notion of a "wave" only adds clarification if we can imagine it to exist in some space-like realm. This is the case for the type of waves we all know from our daily physical environment, such as water waves, sound waves or

light waves. The quantum waves of physical quantum particles can also be made visible in general by looking at probabilistic detection patterns of these quantum particles on a physical screen, and noting the typical interference patterns when the waves interact and the particles are detected on the screen. One might believe that an analogous situation is not possible for concepts, because intuitively concepts, unlike quantum particles, do not exist "inside" space. If we look at things is an operational way, however, an analysis can be put forward for the quantum model of the combination of the two concepts, and the graded structure of collapse probability weights of the 24 items, which does illustrate the presence of an interference pattern, and as a consequence reveals the underlying wave structure of concepts and their interactions. Let us explain how we can proceed to accomplish such an analysis.

We start by considering **Figure 2**. We see there the 24 different items of **Table 1** represented by numbered spots in a plane where a graded pattern, starting with the lightest region around the spot number 8, which is *Apple*, systematically becomes darker. Different numbers of items are situated in spots in regions of different darkness, for example, number 16, *Lentils*, is situated in a spot in the darkest region. Let us explain how the figure is constructed. The "intensity of light" of a specific region corresponds to the "weights of the items" with respect to the concept *Fruits* in **Table 1**. Looking at **Table 1**, it is indeed *Apple*, which has the highest weight, equal to 0.1184, and hence is represented by spot number 8 on **Figure 2**, in the lightest region. Next comes *Elderberry* with weight equal to 0.1134, represented by spot number 7 on **Figure 2**, on the border of the lightest and second lightest region. Next comes *Raisin*, with weight equal to 0.1026, represented by spot number 6 on **Figure 2**, on the border of the third and the fourth lightest region. Next comes *Tomato*, with weight equal to 0.0881, represented by spot number 19 on **Figure 2**, in the seventh lightest region, etc. last is *Lentils*, with weight equal to 0.0095, represented by spot number 16 on **Figure 2**, in the one to darkest region. Hence **Figure 2** contains a representation of the values of the collapse probability weights of the 24 items with respect to the concept *Fruits*. There is however more; we can, for example, wonder what the reason is to choose a representation in a plane? To explain this, turn to **Figure 3**. Let us first note with respect to the two figures, although it might not seem the case at first sight, all the numbered regions are located at exactly the same spots in both **Figures 2, 3**, with respect to the two orthogonal axes that coordinate the plane. What is different in both figures are the graded structures of lighter to darker regions, while they are centered around the spot number 8, representing the item *Apple*, in **Figure 2** they are centered around the spot number 21, representing the item *Broccoli*, in **Figure 3**. And, effectively, **Figure 3** represents analogous to **Figure 2** of the same 24 items, their collapse probability weights, but this time with respect to the concept *Vegetables*. This explains why in **Figure 3** the lightest region is the one centered around spot number 21, representing *Broccoli*, while the lightest region in **Figure 2** is the one centered around spot number 8, representing *Apple*. Indeed, *Broccoli* is the most characteristic vegetable of the considered items, while *Apple* is the most characteristic fruit, if "characteristic" is measured by the size of the respective collapse
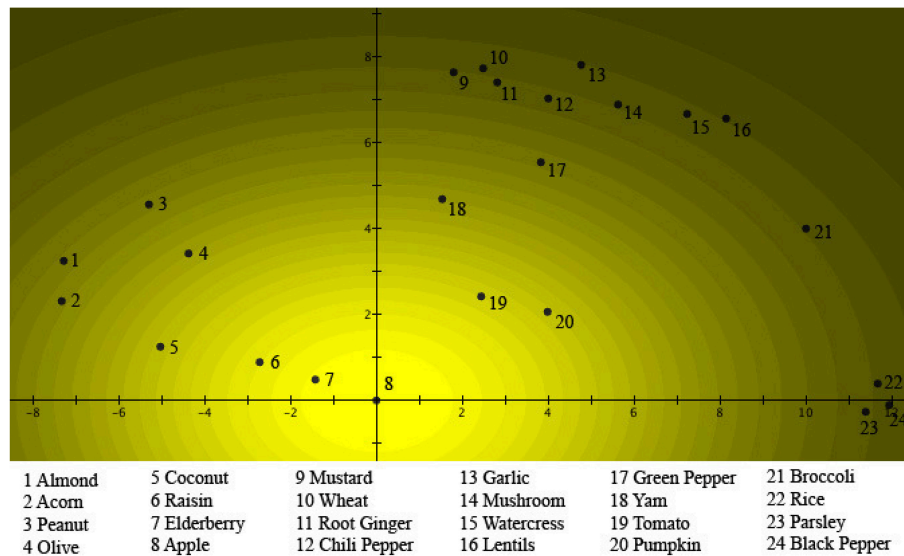
FIGURE 2 | The probabilities $\mu(A)_k$ of a person choosing the item k as a "good example" of _Fruits_ are fitted into a two-dimensional quantum wave function $\psi_A(x, y)$. The numbers are placed at the locations of the different items with respect to the Gaussian probability distribution $|\psi_A(x, y)|^2$. This can be seen as a light source shining through a hole centered on the origin, and regions where the different items are located. The brightness of the light source in a specific region corresponds to the probability that this item will be chosen as a "good example" of _Fruits_.
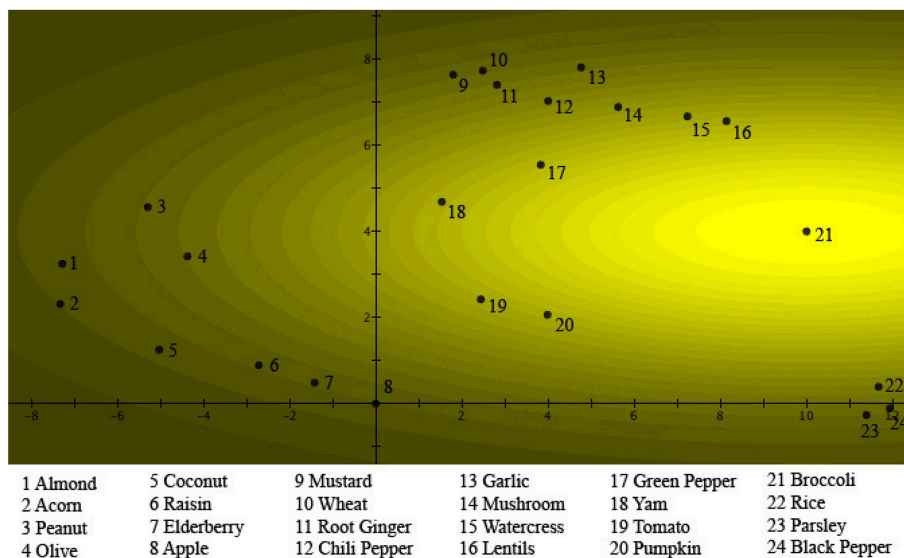
FIGURE 3 | The probabilities $\mu(B)_k$ of a person choosing the item k as a "good example" of _Vegetables_ are fitted into a two-dimensional quantum wave function $\psi_B(x, y)$. The numbers are placed at the locations of the different items with respect to the probability distribution $|\psi_B(x, y)|^2$. As in **Figure 2**, it can be seen as a light source shining through a hole centered on point 21, where _Broccoli_ is located. The brightness of the light source in a specific region corresponds to the probability that this item will be chosen as a "good example" of _Vegetables_.

probability, i.e., the probability to choose this item in the course of the study. What might not seem obvious is that in a plane it is always possible to find 24 locations for the 24 items such that a graded structure with center _Apple_ and a second graded structure with center _Broccoli_ can be defined, fitting exactly also the other items in their correct value of "graded light to dark," corresponding to the collapse probability weights in **Table 1**. Such a situation is what we show in **Figures 2**, **3**. It can be

proven mathematically that a solution always exists, although not a unique one, which means that **Figures 2**, **3** show one of these solutions.

We have chosen on purpose the graded structure form light to dark to be colored yellow, because we can interpret **Figures 2**, **3** such that an interesting analogy arises between our study of the 24 items and two concepts _Fruits_ and _Vegetables_, and the well-known double slit experiment with light in quantum

mechanics. It is this analogy that will also directly illustrate the "wave nature" of concepts. Suppose we consider a plane figuring in the experiment as a detection screen, and put counters for quantum light particles, i.e., photons, at the numbered spots on the plane. Then we send light through a first slit, which we call the *Fruits* slit, which is placed in front of the screen. The slit is placed such that the counters in the spots detect numbers of photons with fractions to the total number of photons send equal the collapse probability weights of the items represented by the respective spots with respect to the concept *Fruits*. The light received on the screen would then look like what is shown in **Figure 2**. Similarly, with counters placed in the same spots, we send light through a second slit, which we call the *Vegetable* slit. Now the counters detect numbers of photons with fractions to the total number of photons equal to the collapse probability weights of the same items with respect to the concept *Vegetables*. The light received on the screen would then look like what is shown in **Figure 3**. We can obtain the same figures directly for our psychological study, consisting of each participant choosing amongst the 24 items the one that he or she finds most characteristic of *Fruits* and *Vegetables* respectively. The relative frequencies of the first choice gives rise to the image in **Figure 2**, while the relative frequencies of the second choice gives rise to the image in **Figure 3**, if, for example, we would mark each chosen item by a fixed number of yellow light pixels on a computer screen.

Before we combine the two slits to give rise to interference, let us specify the mathematics of the quantum mechanical formalism that underlies the two Figures. The situation can be represented quantum mechanically by complex valued Schrödinger wave functions of two real variables $\psi_A(x, y)$, $\psi_B(x, y)$. For the light and the two slits, this situation is the "interaction of a photon with the two slits." For the human participants in the concepts study, this situation is the "interaction with the two concepts of the mind of a participant." We choose for $\psi_A(x, y)$ and $\psi_B(x, y)$ quantum wave packets, such that the radial part for both wave packets is a Gaussian in two dimensions. Considering **Figures 2, 3**, we choose the top of the first Gaussian in the origin where spot number 8 is located, and the top of the second Gaussian in the point (a, b), where spot number 21 is located. Hence

$$\psi_A(x, y) = \sqrt{D_A} e^{-\left(\frac{x^2}{4\sigma_{Ax}^2} + \frac{y^2}{4\sigma_{Ay}^2}\right)} e^{iS_A(x,y)}$$

$$\psi_B(x, y) = \sqrt{D_B} e^{-\left(\frac{(x-a)^2}{4\sigma_{Bx}^2} + \frac{(y-b)^2}{4\sigma_{By}^2}\right)} e^{iS_B(x,y)} \tag{51}$$

The phase parts of the wave packets $e^{iS_A(x,y)}$ and $e^{iS_B(x,y)}$ are determined by two phase fields $S_A(x, y)$ and $S_B(x, y)$ which will account for the interference and hence carry the wave nature. Of course, these phase parts vanish when we multiply each wave packet with its complex conjugate to find the connection with the collapse probabilities. Hence,

$$|\psi_A(x, y)|^2 = D_A e^{-\left(\frac{x^2}{2\sigma_{Ax}^2} + \frac{y^2}{2\sigma_{Ay}^2}\right)} \quad |\psi_B(x, y)|^2 = D_B e^{-\left(\frac{(x-a)^2}{2\sigma_{Bx}^2} + \frac{(y-b)^2}{2\sigma_{By}^2}\right)} \tag{52}$$

are the Gaussians to be seen in **Figures 2, 3**, respectively. Let us denote by Δ_k a small surface specifying the spot corresponding to the item number k in the plane of the two figures. We then calculate the collapse probabilities of this item k with respect to the concepts *Fruits* and *Vegetables* in a standard quantum mechanical way as follows

$$\mu_k(A) = \int_{\Delta_k} |\psi_A(x, y)|^2 dxdy = \int_{\Delta_k} D_A e^{-\left(\frac{x^2}{2\sigma_{Ax}^2} + \frac{y^2}{2\sigma_{Ay}^2}\right)} dxdy \tag{53}$$

$$\mu_k(B) = \int_{\Delta_k} |\psi_B(x, y)|^2 dxdy = \int_{\Delta_k} D_B e^{-\left(\frac{x^2}{2\sigma_{Bx}^2} + \frac{y^2}{2\sigma_{By}^2}\right)} dxdy \tag{54}$$

We can prove that the parameters of the Gaussians, D_A, σ_{Ax}, σ_{Ay}, D_B, σ_{Bx}, σ_{By} can be determined in such a way that the above equations come true, and for the images of **Figures 2, 3**, exactly as we have done—using an approximation for the integrals, which we explain later.

If we open both slits it will be the normalized superposition of the two wave packets that quantum mechanically describes the new situation

$$\psi_{A \text{or} B}(x, y) = \frac{1}{\sqrt{2}} (\psi_A(x, y) + \psi_B(x, y)) \tag{55}$$

We have

$$\mu_k(A \text{ or } B) = \int_{\Delta_k} \psi_{A \text{or} B}(x, y)^* \psi_{A \text{or} B}(x, y) dxdy$$
$$= \frac{1}{2} \left(\int_{\Delta_k} \psi_A(x, y)^* \psi_A(x, y) dxdy + \int_{\Delta_k} \psi_B(x, y)^* \psi_B(x, y) dxdy \right)$$
$$+ \int_{\Delta_k} \Re(\psi_A(x, y)^* \psi_B(x, y)) dxdy$$
$$= \frac{1}{2} (\mu_k(A) + \mu_k(B)) +$$
$$\int_{\Delta_k} \Re(\psi_A(x, y)^* \psi_B(x, y)) dxdy \tag{56}$$

Let us calculate $\int_{\Delta_k} \Re(\psi_A(x, y)^* \psi_B(x, y)) dxdy$. We have

$$\int_{\Delta_k} \Re(\psi_A(x, y)^* \psi_B(x, y)) dxdy$$
$$= \int_{\Delta_k} \left(\sqrt{D_A} e^{-\left(\frac{x^2}{4\sigma_{Ax}^2} + \frac{y^2}{4\sigma_{Ay}^2}\right)} \right) \left(\sqrt{D_B} e^{-\left(\frac{(x-a)^2}{4\sigma_{Bx}^2} + \frac{(y-b)^2}{4\sigma_{By}^2}\right)} \right)$$
$$\Re(e^{-iS_A(x,y)} e^{iS_B(x,y)}) dxdy$$
$$= \int_{\Delta_k} \left(\sqrt{D_A D_B} e^{-\left(\frac{x^2}{4\sigma_{Ax}^2} + \frac{(x-a)^2}{4\sigma_{Bx}^2} + \frac{y^2}{4\sigma_{Ay}^2} + \frac{(y-b)^2}{4\sigma_{By}^2}\right)} \right)$$
$$\Re(e^{i(S_B(x,y) - S_A(x,y))}) dxdy$$

$$= \int_{\Delta_k} \left(\sqrt{D_A D_B} e^{-(\frac{x^2}{4\sigma_{Ax}^2} + \frac{(x-a)^2}{4\sigma_{Bx}^2} + \frac{y^2}{4\sigma_{Ay}^2} + \frac{(y-b)^2}{4\sigma_{By}^2})} \right)$$
$$\cos(S_B(x,y) - S_A(x,y)) dx dy \qquad (57)$$

We can hence rewrite (Equation 56) in the following way

$$\int_{\Delta_k} f(x,y) \cos\theta(x,y) dx dy = f_k \qquad (58)$$

where

$$f(x,y) = \sqrt{D_A D_B} e^{-\left(\frac{x^2}{4\sigma_{Ax}^2} + \frac{(x-a)^2}{4\sigma_{Bx}^2} + \frac{y^2}{4\sigma_{Ay}^2} + \frac{(y-b)^2}{4\sigma_{By}^2}\right)} \qquad (59)$$

is a known Gaussian-like function, remember that we have determined D_A, D_B, σ_{Ax}, σ_{Ay}, σ_{Bx}, σ_{By} and a and b in choosing a solution to be seen in **Figures 3, 4**, and

$$f_k = \mu_k(A \text{ or } B) - \frac{1}{2}(\mu_k(A) + \mu_k(B)) \qquad (60)$$

are constants for each k determined by the data, and we have introduced

$$\theta(x,y) = S_B(x,y) - S_A(x,y) \qquad (61)$$

the field of phase differences of the two quantum wave packets. This field of phases differences will determine the interference pattern and it is the solution of the 24 nonlinear Equations in (58). This set of 24 equations cannot be solved exactly, but even a general numerical solution is not straightforwardly at reach

within actual optimization programs. We have introduces two steps of idealization to find a solution. First, we have looked for a solution where $\theta(x,y)$ is a large enough, polynomial in x and y, more specifically consisting of 24 independent sub-polynomials that are independent

$$\begin{aligned} \theta(x,y) = \ & F_1 + F_2 x + F_3 y + F_4 x^2 + F_5 xy + F_6 y^2 + F_7 x^3 \\ & + F_8 x^2 y + F_9 xy^2 + F_{10} y^3 + F_{11} x^4 + F_{12} x^3 y \\ & + F_{13} x^2 y^2 + F_{14} xy^3 + F_{15} y^4 + F_{16} x^5 + F_{17} x^4 y \\ & + F_{18} x^3 y^2 + F_{19} x^2 y^3 + F_{20} xy^4 + F_{21} y^5 + F_{22} x^6 \\ & + F_{23} x^5 y + F_{24} x^4 y^2 \end{aligned} \qquad (62)$$

Secondly, we suppose that $\Delta_k = \Delta$ is a sufficiently small square surface such that a good approximation of the integral in Equation (58)—and it is also the approximation we have used for the integrals (Equations 53 and 54)—is given by Δ times the value of the function under the integral in the center of Δ. This transforms the set of 24 nonlinear (Equation 58) into a set of 24 linear equations

$$\Delta f(x_k, y_k) \theta(x_k, y_k) = f_k \qquad (63)$$

We have solved them for the points (x_k, y_k) where the 24 items are located in **Figures 2, 3**, for $\Delta = 0.01$, which gives us $\theta(x,y)$, and hence also the expression for $|\psi_{A \text{ or } B}(x,y)|^2$ containing the expected interference term, and we have

$$\begin{aligned} |\psi_{A \text{ or } B}(x,y)|^2 = \ & \frac{1}{2}(|\psi_A(x,y)|^2 + |\psi_B(x,y)|^2) \\ & + |\psi_A(x,y)\psi_B(x,y)| \cos\theta(x,y) \end{aligned} \qquad (64)$$

1 Almond	5 Coconut	9 Mustard	13 Garlic	17 Green Pepper	21 Broccoli
2 Acorn	6 Raisin	10 Wheat	14 Mushroom	18 Yam	22 Rice
3 Peanut	7 Elderberry	11 Root Ginger	15 Watercress	19 Tomato	23 Parsley
4 Olive	8 Apple	12 Chili Pepper	16 Lentils	20 Pumpkin	24 Black Pepper
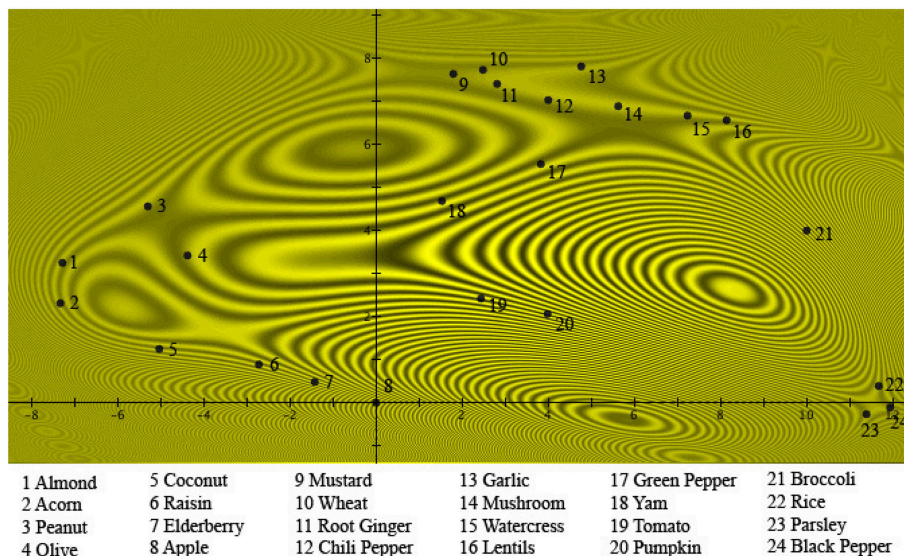
FIGURE 4 | The probabilities $\mu(A \text{ or } B)_k$ of a person choosing the item k as a "good example" of *Fruits or Vegetables* are fitted into the two-dimensional quantum wave function $\frac{1}{\sqrt{2}}(\psi_A(x,y) + \psi_B(x,y))$, which is the normalized superposition of the wave functions in Figures 2, 3. The numbers are placed at the locations of the different exemplars with respect to the probability distribution $|\psi_A(x,y) + \psi_B(x,y)|^2 = \frac{1}{2}(|\psi_A(x,y)|^2 + |\psi_B(x,y)|^2)$ $+ |\psi_A(x,y)\psi_B(x,y)| \cos\theta(x,y)$, where $\theta(x,y)$ is the quantum phase difference at (x,y). The values of $\theta(x,y)$ are given in **Table 1** for the locations of the different items. The interference pattern is clearly visible.

TABLE 2 | The parameters of the interference pattern solution illustrated in Figure 4.

			Parameters of the solution			
k	item	(x, y) coordinates of items	Sub-polynomial	Coefficients F_k	Gaussian	Parameters
1	Almond	(−7.2826, 3.24347)	1	87.6039	D_A	1.18412
2	Acorn	(−7.3316, 2.3116)	x	2792.02	σ_{Ax}	5.65390
3	Peanut	(−5.2957, 4.56032)	y	8425.01	σ_{Ay}	3.80360
4	Olive	(−4.3776, 3.41765)	x^2	19.36	D_B	1.28421
5	Coconut	(−5.0322, 1.24573)	xy	−2139.87	σ_{Bx}	8.20823
6	Raisin	(−2.7149, 0.896651)	y^2	−7322.26	σ_{By}	2.41578
7	Elderberry	(−1.420, 0.487598)	x^3	−39.2811		
8	Apple	(0, 0)	x^2y	−55.5263		
9	Mustard	(1.7978, 7.64549)	xy^2	586.674		
10	Wheat	(2.4786, 7.73915)	y^3	2205.81		
11	Root Ginger	(2.8164, 7.41004)	x^4	−2.22868		
12	Chili Pepper	(3.9933, 7.03549)	x^3y	4.19408		
13	Garlic	(4.7681, 7.81803)	x^2y^2	13.3579		
14	Mushroom	(5.6281, 6.89107)	xy^3	−72.233		
15	Watercress	(7.233, 6.67322)	y^4	−275.834		
16	Lentils	(8.1373, 6.56281)	x^5	0.426731		
17	Green Pepper	(3.8337, 5.55379)	x^4y	1.58764		
18	Yam	(1.5305, 4.69497)	x^3y^2	0.582536		
19	Tomato	(2.4348, 2.42612)	x^2y^3	−1.13167		
20	Pumpkin	(3.9873, 2.06652)	xy^4	3.44008		
21	Broccoli	(10, 4)	y^5	12.2584		
22	Rice	(11.6771, 0.392458)	x^6	−0.00943132		
23	Parsley	(11.3949, −0.268463)	x^5y	−0.0535881		
24	Black Pepper	(11.9389, −0.107151)	x^4y^2	−0.200688		

*The first column lists the different items, and the second column the coordinates of their locations in **Figures 2–4**. The third column contains the orthogonal set of sub-polynomials used as first approximation for the phase field θ(x, y), and the fourth column their values. The fifth and sixth columns contain the Gaussian parameters and their values of the solution.*

In **Figure 4** we have graphically represented this probability density $|\psi_{AorB}(x, y)|^2$. The interference pattern shown in **Figure 4** is very similar to well-known interference patterns of light passing through an elastic material under stress. In our case, it is the interference pattern corresponding to "Fruits or Vegetables" as a contextual, interfering prototype. The numerical values of the solutions represented in **Figures 2–4** are in **Table 2**.

We have thus completed our illustration of contextual interfering prototypes. It is, however, important to remember that this representation is at the subtle level of an illustration, while the real working representation of contextual interfering prototypes needs the complete quantum-mechanical formalism. It can be considered as a pre-representation, exactly as the wave-like representations by de Broglie and Schrödinger in the early days of quantum physics can be considered as useful pre-quantum representations that capture something of the wave aspects of microscopic particles.

5. DISCUSSION

In this paper we showed that a generalization of prototype theory can address the "Pet-Fish problem" and related combination issues. This was done by formalizing the effect of the cognitive context on the state of a concept using a SCoP formalism (Gabora and Aerts, 2002; Aerts and Gabora, 2005a,b; Gabora et al., 2008). We also developed a quantum-theoretic model in complex Hilbert space to show that, in this contextualized prototype theory, prototypes can interfere when concepts combine, as evidenced by data where typicality measurements are performed. This could then lead one to think that the general quantum approach to concepts only presupposes a (contextual) prototype theory. We now explain why this inference is not true.

Let us make more explicit the relationship between our quantum-conceptual approach and other concept theories, such as prototype theory, exemplar theory and theory theory. A deeper analysis shows that our approach is more than a contextual generalization of prototype theory. Roughly speaking, other theories make assumptions about the principles guiding the formation and intuitive representation of a concept in the human mind. Thus, prototype theory assumes that a concept is determined by a set of characteristic rather than defining features, the human mind has a privileged prototype for each concept, and typicality of a concrete item is determined by its similarity with the prototype (Rosch, 1973, 1978, 1983). Exemplar theory assumes instead that a concept is not determined by a

set of defining or characteristic features but, rather, by a set of salient instances of it stored in memory (Nosofsky, 1988, 1992). Theory theory assumes that concepts are determined by "mini-theories" or schemata, identifying the causal relationships among properties (Murphy and Medin, 1985; Rumelhart and Norman, 1988). These theories have all mainly been preoccupied with the question of "what predominantly determines a concept." We agree on the relevance of this question, though it is not the main issue focused on there. Transposed to our approach, these theories mainly investigate "what predominantly determines the state of a concept." Conversely, the main preoccupation of our approach has been to propose a theory with the following features:

(i) a well-defined ontology, i.e., a concept is in our approach an entity capable of different modes of being with respect to how it influences measurable semantic quantities such as typicality, membership weight and membership probability, and these modes are called "states";
(ii) the capacity to produce theoretical models fitting data on these measurable semantic quantities.

We seek to achieve (i) and (ii) independent of the question that is the focus of other theories of concepts. More concretely, and in accordance with the results of investigations into the question of "what predominantly determines a concept," as far as prototype theory, exemplar theory and theory theory are concerned, we believe that all approaches are partially valid. The state of a concept, i.e., its capability of influencing the values of measurable semantic quantities, such as typicality and membership weight, is influenced by the set of its characteristic features, but also by salient exemplars in memory, and in a considerable number of cases—where more causal aspects are at play—mini-theories might be appropriate to express this state. It is important that "a conceptual state is defined and gives rise, together with the context, to the values of the measurable semantic quantities." The fact that the specification of these values can be only probabilistic is a confirmation that potentiality and uncertainty occur even if the state is completely known, hence quantum structures are intrinsically needed.

It follows from the above that resorting and giving new life to prototype theory does not necessarily entail that contextual prototype theory is the only possible theory of concepts for what concerns the question of "what predominantly determines a concept." However, we choose to identify our general approach as a "generalized contextual interfering prototype theory," because the "ground state" of a concept is a fundamental notion of the theory, and this ground state is what corresponds to the prototype. There is not a similar affinity with exemplar theory and theory theory. However, the conceptual state and its interaction with the cognitive context can potentially capture the other conceptual aspects, exemplars and schemata, which are instead predominant in alternative concept theories. In this respect, an interesting analogy must be emphasized. The quantum-theoretic approach only aims at modeling concepts and their combinations in a unitary and coherent mathematical formalism. We do not pretend to give a universal definition of what a concept is and how it

forms. Using a known analogy in mathematics, we can say that the quantum-theoretic model is to a concept as a traditional Kolmogorov model is to a probability. A Kolmogorovian model specifies how a probability can be mathematically formalized independent of the definition of probability that is chosen (favorable over possible cases, large number limit of frequencies, subjective, etc.). Analogously, the quantum-theoretic framework for concepts enables mathematical modeling of conceptual entities independent of the definition that is adopted in a specific concept theory (prototype, exemplar, theory, etc.).

We conclude with an epistemological consideration. The quantum-theoretic framework presented here constitutes a step toward the elaboration of a general theory for the representation of any conceptual entity. Hence, it is not just a "cognitive model for typicality, membership weight or membership probability." Rather, we are investigating whether "quantum theory, in its Hilbert space formulation, is an appropriate theory to model human cognition." To understand what we mean by this let us consider an example taken from everyday life. As an example of a theory, we could introduce the theory of "how to make good clothes." A tailor needs to learn how to make good clothes for different types of people, men, women, children, old people, etc. Each cloth is a model in itself. Then, one can also consider intermediate situations where one has models of series of clothes. A specific body will not fit in any clothes: you need to adjust the parameters (length, size, etc.) to reach the desired fit. We think that a theory should be able to reproduce different experimental results by suitably adjusting the involved parameters, exactly as a theory of clothing. This is different from a set of models, even if the set can cope with a wide range of data.

There is a tendency, mainly in empirically-based disciplines, to be critical with respect to a theory that can cope with all possible situations it applies to. This is because the theory contains too many parameters, which may lead one to think that "any type of data can be modeled by allowing all these parameters to have different values." We agree that, in case we have to do with an "*ad-hoc* model," i.e., a model specially made for the circumstance of the situation it models, this suspicion is grounded. Adding parameters to such an *ad-hoc* model, or stretching the already contained parameters to other values, does not give rise to what we call a theory. On the other hand, a theory needs to be well defined, its rules, the allowed procedures, its theoretical, mathematical, and internal logical structure, "independent" of the structure of the models describing specific situations that can be coped with by the theory. Hence also the theory needs to contain a well defined description of "how to produce models for specific situations." Coming back to the theory of clothing, if a tailor knows the theory of clothing, obviously he or she can make a cloth for every human body, because the theory of clothing, although its structure is defined independently of a specific cloth, contains a prescription of how to apply it to any possible specific cloth. In this respect, we think that one should carefully distinguish between a model that is derived by a general theory, as the one presented in this paper, and a model specifically designed to test a number of experimental situations.

This brings us to the important question of the "predictive power" of existing quantum-theoretic models. Models derived from a theory will generally need more data from a bigger set of experiments to become predictive for the outcomes of other not yet performed experiments than this is the case for models that are more *ad-hoc*. The reason is that in principle such models—think of the analogy we present with the theory of clothing above—must be able to faithfully represent the data of all possible experiments that can be performed on the conceptual entity in the same state. A tailor knowing the theory of clothing can in principle make clothes for all human bodies but hence also predicts outcomes of not performed experiments, e.g., the measure of a specific part of the cloth, if enough data of a set of experiments are available to the tailor, e.g., data that determine the possible types of clothes still fitting these data and as a consequence also determine the measure of this part of the clothe. In general in quantum cognition, the scarcity of data is preventing models from having systematic and substantial predictive power. One can wonder, if predictive power is not yet predominantly available in the majority of existing quantum-theoretic models, why so much attention and value is actually attributed to them? Answering this question allows us to clarify an aspect of quantum cognition that is not obvious and even makes it special in a specific way, at least provisionally until more data is available. The success of quantum cognition is due to it "being able to convincingly model data that theoretically can be proven to be impossible to model with any model that relies on classical fuzzy set theory and/or classical Kolmogorovian probability theory." Hence, a different criterion than predictive power is provisionally used to identify the success of quantum cognition. Of course, as soon as more data are collected, the models will also be able to be tested for their predictive power. Recent work in quantum cognition is starting to reach the level of being predictive, for example study of order effects (Wang et al., 2014), and an elaboration and refinement of the model presented in this article (Aerts et al., 2015a,c). The latter model simultaneously investigates the "conjunction" and the "negation" of concepts, starting from data collected on such conceptual combinations. To explain the exact nature and also accurateness of the predictive power we gained in the model in Aerts et al. (2015a,c), consider the following mathematical expression

$$I_{ABA'B'} = 1 - \mu(A \text{ and } B) - \mu(A \text{ and } B') - \mu(A' \text{ and } B) - \mu(A' \text{ and } B') \quad (65)$$

where A and B are the concepts *Fruits* and *Vegetables*, respectively, while A' and B' are their negations. Thus, "A and B'" means *Fruits and not Vegetables*, while "A' and B" means *not Fruits and Vegetables* and "A' and B'" means *not Fruits and not Vegetables*. In Aerts et al. (2015a,c) we published the data for the outcomes of experiments that test the membership of the same 24 items which we considered in the present article, but this time not only for the conjunction of A and B, but also for the conjunctions "A and B'," "A' and B," and "A' and B'." Suppose that the data follow a classical probabilistic structure, then $I_{ABA'B'}$ has to be theoretically equal to zero for each considered item, and this purely follows from a general "law of probability calculus"

related to the so called "de Morgan laws" of classical probability. This means that, under the hypothesis of a classical probabilistic structure, if we measure the relative frequencies of "A and B," "A and B'" and "A' and B," and hence determine experimentally the values of $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$ and $\mu(A' \text{ and } B)$, a "prediction" for $\mu(A' \text{ and } B')$ can be made theoretically, namely,

$$\mu(A' \text{ and } B') = 1 - \mu(A \text{ and } B) - \mu(A \text{ and } B') - \mu(A' \text{ and } B) \quad (66)$$

for each considered item. Let is explain what are our findings in Aerts et al. (2015a,c) that make it possible for us to speak of some specific type of predictability for the more elaborated and refined model we developed for the combination of concepts and their negations. In Aerts et al. (2015a,c) we have collected data not only for the pair of concepts *Fruits* and *Vegetables* and the 24 items treated also in the present article, but for three more pairs of concepts, and for each of them again 24 items. Due to the already identified non classical nature of overextension of the conjunction we expected that $I_{ABA'B'}$ would not be equal to zero, and that indeed showed to be the case. However, we detected a high level of systematics of the value of $I_{ABA'B'}$ fluctuating around an average of -0.81. A statistical analysis showed the different values for individual items to be possible to be explained as fluctuations around this average (see Tables 1–4 in Aerts et al., 2015a). Next to the detailed statistical analysis to be found in Aerts et al. (2015a) we also put forward a theoretical explanation of this value. The elaborated and refined model for concept combinations developed in Aerts et al. (2015a) introduces within the model the combination of a pure quantum model and a classical model. It can be shown that for a pure quantum model the value of $I_{ABA'B'}$ would be -1. We also find that the quantum effects are dominant as compared to the classical effects in case concepts are combined, which explains why our refined model gives rise to a value of $I_{ABA'B'}$ in between the classical one, which is 0, and the pure quantum one, which is -1, but closer to the quantum one, hence -0.81. This finding can be turned into a predictive one in the following way. Suppose we measure $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$ and $\mu(A' \text{ and } B)$ for two arbitrary concepts and an item. Our model allows us to put forward the following prediction for $\mu(A' \text{ and } B')$

$$\mu(A' \text{ and } B') = 1.81 - \mu(A \text{ and } B) - \mu(A \text{ and } B') - \mu(A' \text{ and } B) \quad (67)$$

By comparing Equations (66) and (67), we get that the quantum-theoretic model in Aerts et al. (2015a) provides a "different prediction" from a classical probabilistic model satisfying the axioms of Kolmogorov, and experiments confirm the validity of the former over the latter. We add that the quantum model has different predictions from a classical model also for the values of other functions than $I_{ABA'B'}$, and these predictions are "parameter independent," in the sense that they do not depend on the values of free parameters that may accommodate the data.

The results above can be considered as a strong confirmation that quantum-theoretic models of concept

combinations provide predictions that deviate, in some situations, from the predictions of classical Kolmogorovian models, which is confirmed by experimental data.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Aerts, D. (1986). A possible explanation for the probabilities of quantum mechanics. *J. Math. Phys.* 27, 202–210. doi: 10.1063/1.527362

Aerts, D. (1999). Foundations of quantum physics: a general realistic and operational approach. *Int. J. Theor. Phys.* 38, 289–358. doi: 10.1023/A:1026605829007

Aerts, D. (2009a). Quantum structure in cognition. *J. Math. Psychol.* 53, 314–348. doi: 10.1016/j.jmp.2009.04.005

Aerts, D. (2009b). Quantum particles as conceptual entities: a possible explanatory framework for quantum theory. *Found. Sci.* 14, 361–411. doi: 10.1007/s10699-009-9166-y

Aerts, D., and Aerts, S. (1995). Applications of quantum statistics in psychological studies of decision processes. *Found. Sci.* 1, 85–97. doi: 10.1007/BF00208726

Aerts, D., Aerts, S., Broekaert, J., and Gabora, L. (2000). The violation of Bell inequalities in the macroworld. *Found. Phys.* 30, 1387–1414. doi: 10.1023/A:1026449716544

Aerts, D., Broekaert, J., Gabora, L., and Sozzo, S. (2013b). Quantum structure and human thought. *Behav. Brain Sci.* 36, 274–276. doi: 10.1017/S0140525X12002841

Aerts, D., and Gabora, L. (2005a). A theory of concepts and their combinations I: the structure of the sets of contexts and properties. *Kybernetes* 34, 167–191. doi: 10.1108/03684920510575799

Aerts, D., and Gabora, L. (2005b). A theory of concepts and their combinations II: a Hilbert space representation. *Kybernetes* 34, 192–221. doi: 10.1108/03684920510575807

Aerts, D., Gabora, L., and S. Sozzo, S. (2013a). Concepts and their dynamics: a quantum–theoretic modeling of human thought. *Top. Cogn. Sci.* 5, 737–772. doi: 10.1111/tops.12042

Aerts, D., and Sozzo, S. (2011). *Quantum Structure in Cognition. Why and How Concepts are Entangled. Quantum Interaction. Lecture Notes in Computer Science 7052*. Berlin: Springer.

Aerts, D., and Sozzo, S. (2014). Quantum entanglement in conceptual combinations. *Int. J. Theor. Phys.* 53, 3587–3603. doi: 10.1007/s10773-013-1946-z

Aerts, D., Sozzo, S., and Tapia, J. (2014). Identifying quantum structures in the Ellsberg paradox. *Int. J. Theor. Phys.* 53, 3666–3682. doi: 10.1007/s10773-014-2086-9

Aerts, D., Sozzo, S., and Veloz, T. (2015). Quantum nature of identity in human concepts: Bose-Einstein statistics for conceptual indistinguishability. *Int. J. Theor. Phys.* 54, 4430–4443. doi: 10.1007/s10773-015-2620-4

Aerts, D., Sozzo, S., and Veloz, T. (2015a). Quantum structure of negation and conjunction in human thought. *Front. Psychol.* 6:1447. doi: 10.3389/fpsyg.2015.01447

Aerts, D., Sozzo, S., and Veloz, T. (2015b). Quantum structure in cognition and the foundations of human reasoning. *Int. J. Theor. Phys.* 54, 4557–4569. doi: 10.1007/s10773-015-2717-9

Aerts, D., Sozzo, S., and Veloz, T. (2015c). New fundamental evidence of non-classical structure in the combination of natural concepts. *Philos. Trans. R. Soc. A* 374:20150095. doi: 10.1098/rsta.2015.0095

Alxatib, S., and Pelletier, J. (2011). "On the psychology of truth gaps," in *Vagueness in Communication*, eds R. Nouwen, R. van Rooij, U. Sauerland, and H.-C. Schmitz (Berlin; Heidelberg: Springer-Verlag), 13–36. doi: 10.1007/978-3-642-18446-8_2

Bonini, N., Osherson, D., Viale, R., and Williamson, T. (1999). On the psychology of vague predicates. *Mind Lang.* 14, 377–393. doi: 10.1111/1468-0017.00117

Busemeyer, J. R., and Bruza, P. D. (2012). *Quantum Models of Cognition and Decision*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511997716

Costello, J., and Keane, M. T. (2000). Efficient creativity: constraint-guided conceptual combination. *Cogn. Sci.* 24, 299–349. doi: 10.1207/s15516709cog2402_4

Fodor, J. (1994) Concepts: a potboiler. *Cognition* 50, 95–113. doi: 10.1016/0010-0277(94)90023-X

Gabora, L., and Aerts, D. (2002). Contextualizing concepts using a mathematical generalization of the quantum formalism. *J. Exp. Theor. Artif. Intell.* 14, 327–358. doi: 10.1080/09528130210162253

Gabora, L., Rosch, E., and Aerts, D. (2008). Toward an ecological theory of concepts. *Ecol. Psychol.* 20, 84–116. doi: 10.1080/10407410701766676

Gagne, C. L., and Spalding, T. L. (2009). Constituent integration during the processing of compound words: does it involve the use of relational structures? *J. Mem. Lang.* 60, 20–35. doi: 10.1016/j.jml.2008.07.003

Geeraerts, D. (1989). Prospects and problems of prototype theory. *Linguistics* 27, 587–612. doi: 10.1515/ling.1989.27.4.587

Hampton, J. A. (1988a). Overextension of conjunctive concepts: evidence for a unitary model for concept typicality and class inclusion. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 12–32. doi: 10.1037/0278-7393.14.1.12

Hampton, J. A. (1988b). Disjunction of natural concepts. *Mem. Cogn.* 16, 579–591. doi: 10.3758/BF03197059

Hampton, J. A. (1997). Conceptual combination: conjunction and negation of natural concepts. *Mem. Cogn.* 25, 888–909. doi: 10.3758/BF03211333

Haven, E., and Khrennikov, A. Y. (2013). *Quantum Social Science*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139003261

Kamp, H., and Partee, B. (1995). Prototype theory and compositionality. *Cognition* 57, 129–191. doi: 10.1016/0010-0277(94)00659-9

Khrennikov, A. Y., Basieva, I., Dzhafarov, E. N., and Busemeyer, J. R. (2014). Quantum models for psychological measurements: an unsolved problem. *PLoS ONE* 9:e110909. doi: 10.1371/journal.pone.0110909

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitrechnung, Ergebnisse Der Mathematik; translated as Foundations of Probability*. New York, NY: Chelsea Publishing Company.

Komatsu, L. K. (1992). Recent views on conceptual structure. *Psychol. Bull.* 112, 500–526. doi: 10.1037/0033-2909.112.3.500

Murphy, G. L., and Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychol. Rev.* 92, 289–316. doi: 10.1037/0033-295X.92.3.289

Nosofsky, R. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 700–708. doi: 10.1037/0278-7393.14.4.700

Nosofsky, R. (1992). "Exemplars, prototypes, and similarity rules," in *From Learning Theory to Connectionist Theory: Essays in Honor of William K. Estes*, eds A. Healy, S. Kosslyn, and R. Shiffrin (Hillsdale, NJ: Erlbaum), 149–167.

Osherson, D., and Smith, E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition* 9, 35–58. doi: 10.1016/0010-0277(81)90013-5

Osherson, D. N., and Smith, E. (1982). Gradeness and conceptual combination. *Cognition* 12, 299–318. doi: 10.1016/0010-0277(82)90037-3

Osherson, D. N., and Smith, E. (1997). On typicality and vagueness. *Cognition* 64, 189–206. doi: 10.1016/S0010-0277(97)00025-5

Pitowsky, I. (1989). *Quantum Probability, Quantum Logic. Lecture Notes in Physics*. Vol. 321. Berlin: Springer.

Pothos, E. M., and Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* 36, 255–274. doi: 10.1017/S0140525X12001525

Rips, L. J. (1995). The current status of research on concept combination. *Mind Lang.* 10, 72–104. doi: 10.1111/j.1468-0017.1995.tb00006.x

Rosch, E. (1973). Natural categories. *Cogn. Psychol.* 4, 328–350. doi: 10.1016/0010-0285(73)90017-0

Rosch, E. (1978). "Principles of categorization," in *Cognition and Categorization*, eds E. Rosch and B. Lloyd (Hillsdale, NJ: Lawrence Erlbaum), 133–179.

Rosch, E. (1983). "Prototype classification and logical classification: the two systems," in *New Trends in Conceptual Representation: Challenges to Piaget Theory?*, ed E. Scholnick (Hillsdale, NJ: Lawrence Erlbaum), 133–159.

Rumelhart, D. E., and Norman, D. A. (1988). "Representation in memory," in *Stevens Handbook of Experimental Psychology*, eds R. C. Atkinson, R. J. Hernsein, G. Lindzey, and R. L. Duncan (New York, NY: John Wiley & Sons), 511–587.

Sozzo, S. (2014). A quantum probability explanation in Fock space for borderline contradictions. *J. Math. Psychol.* 58, 1–12. doi: 10.1016/j.jmp.2013.11.001

Sozzo, S. (2015). Conjunction and negation of natural concepts: a quantum-theoretic modeling. *J. Math. Psychol.* 66, 83–102. doi: 10.1016/j.jmp.2015.01.005

Thagard, P., and Stewart, T. C. (2011). The AHA! experience: creativity through emergent binding in neural networks. *Cogn. Sci.* 35, 1–33. doi: 10.1111/j.1551-6709.2010.01142.x

Van Dantzig, S., Raffone, A., and Hommel, B. (2011). Acquiring contextualized concepts: a connectionist approach. *Cogn. Sci.* 35, 1162–1189. doi: 10.1111/j.1551-6709.2011.01178.x

Wang, Z., Solloway, T., Shiffrin, R. M., and Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9431–9436. doi: 10.1073/pnas.1407756111

Zadeh, L. (1982). A note on prototype theory and fuzzy sets. *Cognition* 12, 291–297. doi: 10.1016/0010-0277(82)90036-1

APPENDIX

A. Quantum Mathematics for Conceptual Modeling

We illustrate in this section how the mathematical formalism of quantum theory can be applied to model situations outside the microscopic quantum world, more specifically, in the representation of concepts and their combinations. We will limit technicalities to the essential.

When the quantum mechanical formalism is applied for modeling purposes, each considered entity—in our case a concept—is associated with a complex Hilbert space $\mathcal{H}$, that is, a vector space over the field $\mathbb{C}$ of complex numbers, equipped with an inner product $\langle \cdot | \cdot \rangle$ that maps two vectors $\langle A|$ and $|B\rangle$ onto a complex number $\langle A|B\rangle$. We denote vectors by using the bra-ket notation introduced by Paul Adrien Dirac, one of the pioneers of quantum theory. Vectors can be "kets," denoted by $|A\rangle$, $|B\rangle$, or "bras," denoted by $\langle A|$, $\langle B|$. The inner product between the ket vectors $|A\rangle$ and $|B\rangle$, or the bra-vectors $\langle A|$ and $\langle B|$, is realized by juxtaposing the bra vector $\langle A|$ and the ket vector $|B\rangle$, and $\langle A|B\rangle$ is also called a "bra-ket," and it satisfies the following properties:

(i) $\langle A|A \rangle \geq 0$;
(ii) $\langle A|B \rangle = \langle B|A \rangle^*$, where $\langle B|A \rangle^*$ is the complex conjugate of $\langle A|B \rangle$;
(iii) $\langle A|(z|B\rangle + t|C\rangle) = z\langle A|B\rangle + t\langle A|C\rangle$, for $z, t \in \mathbb{C}$, where the sum vector $z|B\rangle + t|C\rangle$ is called a "superposition" of vectors $|B\rangle$ and $|C\rangle$ in the quantum jargon.

From (ii) and (iii) follows that inner product $\langle \cdot | \cdot \rangle$ is linear in the ket and anti-linear in the bra, i.e., $(z\langle A| + t\langle B|)|C\rangle = z^*\langle A|C\rangle + t^*\langle B|C\rangle$.

The "absolute value" of a complex number is defined as the square root of the product of this complex number times its complex conjugate, that is, $|z| = \sqrt{z^*z}$. Moreover, a complex number z can either be decomposed into its cartesian form $z = x + iy$, or into its polar form $z = |z|e^{i\theta} = |z|(\cos\theta + i\sin\theta)$. As a consequence, we have $|\langle A|B\rangle| = \sqrt{\langle A|B\rangle\langle B|A\rangle}$. We define the "length" of a ket (bra) vector $|A\rangle$ ($\langle A|$) as $|||A\rangle|| = ||\langle A|||| = \sqrt{\langle A|A\rangle}$. A vector of unitary length is called a "unit vector." We say that the ket vectors $|A\rangle$ and $|B\rangle$ are "orthogonal" and write $|A\rangle \perp |B\rangle$ if $\langle A|B\rangle = 0$.

We have now introduced the necessary mathematics to state the first modeling rule of quantum theory, as follows.

A.1. First Quantum Modeling Rule
A state A of an entity—in our case a concept—modeled by quantum theory is represented by a ket vector $|A\rangle$ with length 1, that is $\langle A|A\rangle = 1$.

An orthogonal projection M is a linear operator on the Hilbert space, that is, a mapping $M : \mathcal{H} \rightarrow \mathcal{H}, |A\rangle \mapsto M|A\rangle$ which is Hermitian and idempotent. The latter means that, for every $|A\rangle, |B\rangle \in \mathcal{H}$ and $z, t \in \mathbb{C}$, we have:

(i) $M(z|A\rangle + t|B\rangle) = zM|A\rangle + tM|B\rangle$ (linearity);
(ii) $\langle A|M|B\rangle = \langle B|M|A\rangle^*$ (hermiticity);
(iii) $M \cdot M = M$ (idempotency).

The identity operator $\mathbb{1}$ maps each vector onto itself and is a trivial orthogonal projection. We say that two orthogonal projections M_k and M_l are orthogonal operators if each vector contained in $M_k(\mathcal{H})$ is orthogonal to each vector contained in $M_l(\mathcal{H})$, and we write $M_k \perp M_l$, in this case. The orthogonality of the projection operators M_k and M_l can also be expressed by $M_kM_l = 0$, where 0 is the null operator. A set of orthogonal projection operators $\{M_k \,|\, k = 1, \ldots, n\}$ is called a "spectral family" if all projectors are mutually orthogonal, that is, $M_k \perp M_l$ for $k \neq l$, and their sum is the identity, that is, $\sum_{k=1}^n M_k = \mathbb{1}$.

The above definitions give us the necessary mathematics to state the second modeling rule of quantum theory, as follows.

A.2. Second Quantum Modeling Rule
A measurable quantity Q of an entity—in our case a concept—modeled by quantum theory, and having a set of possible real values $\{q_1, \ldots, q_n\}$ is represented by a spectral family $\{M_k \,|\, k = 1, \ldots, n\}$ in the following way. If the entity—in our case a concept—is in a state represented by the vector $|A\rangle$, then the probability of obtaining the value q_k in a measurement of the measurable quantity Q is $\langle A|M_k|A\rangle = ||M_k|A\rangle||^2$. This formula is called the "Born rule" in the quantum jargon. Moreover, if the value q_k is actually obtained in the measurement, then the initial state is changed into a state represented by the vector

$$|A_k\rangle = \frac{M_k|A\rangle}{||M_k|A\rangle||} \qquad (A1)$$

This change of state is called "collapse" in the quantum jargon.

The tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is the Hilbert space generated by the set $\{|A_i\rangle \otimes |B_j\rangle\}$, where $|A_i\rangle$ and $|B_j\rangle$ are vectors of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, which means that a general vector of this tensor product is of the form $\sum_{ij}|A_i\rangle \otimes |B_j\rangle$. This gives us the necessary mathematics to introduce the third modeling rule.

A.3. Third Quantum Modeling Rule
A state C of a compound entity—in our case a combined concept—is represented by a unit vector $|C\rangle$ of the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of the two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ containing the vectors that represent the states of the component entities—concepts.

The above means that we have $|C\rangle = \sum_{ij} c_{ij}|A_i\rangle \otimes |B_j\rangle$, where $|A_i\rangle$ and $|B_j\rangle$ are unit vectors of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, and $\sum_{i,j} |c_{ij}|^2 = 1$. We say that the state C represented by $|C\rangle$ is a product state if it is of the form $|A\rangle \otimes |B\rangle$ for some $|A\rangle \in \mathcal{H}_A$ and $|B\rangle \in \mathcal{H}_B$. Otherwise, C is called an "entangled state."

The Fock space is a specific type of Hilbert space, originally introduced in quantum field theory. For most states of a quantum field the number of identical quantum entities is not conserved but is a variable quantity. The Fock space copes with this situation in allowing its vectors to be superpositions of vectors pertaining to different sectors for fixed numbers of identical quantum entities. MoreA explicitly, the k-th sector of a Aock space describes a fixed number of k identical quantum entities,

and it is of the form $\mathcal{H} \otimes \ldots \otimes \mathcal{H}$ of the tensor product of k identical Hilbert spaces $\mathcal{H}$. The Aock space A itself is the direct sum of all these sectors, hence

$$\mathcal{A} = \oplus_{k=1}^{j} \otimes_{l=1}^{k} \mathcal{H} \qquad (A2)$$

Aor our modeling we have only used Aock space for the "two" and "one quantum entity" case, hence $\mathcal{A} = \mathcal{H} \oplus (\mathcal{H} \otimes \mathcal{H})$. This is due to considering only combinations of two concepts. The sector $\mathcal{H}$ is called the "first sector," while the sector $\mathcal{H} \otimes \mathcal{H}$ is

called the "second sector." A unit vector $|F\rangle \in \mathcal{F}$ is then written as $|F\rangle = ne^{i\gamma}|C\rangle + me^{i\delta}(|A\rangle \otimes |B\rangle)$, where $|A\rangle$, $|B\rangle$ and $|C\rangle$ are unit vectors of $\mathcal{H}$, and such that $n^2 + m^2 = 1$. For combinations of j concepts, the general form of Fock space in Equation (A2) should be used.

The quantum modeling above can be generalized by allowing states to be represented by the so called "density operators" and measurements to be represented by the so called "positive operator valued measures." However, for the sake of brevity we will not dwell on this extension here.

Quantum structure of negation and conjunction in human thought

Diederik Aerts[1], Sandro Sozzo[2]* and Tomas Veloz[3]

[1] Center Leo Apostel for Interdisciplinary Studies (CLEA), Free University of Brussels (VUB), Brussels, Belgium, [2] School of Management, Institute for Quantum Social and Cognitive Science, University of Leicester, Leicester, UK, [3] Department of Mathematics, University of British Columbia, Kelowna, BC, Canada

We analyze in this paper the data collected in a set of experiments investigating how people combine natural concepts. We study the mutual influence of conceptual conjunction and negation by measuring the membership weights of a list of exemplars with respect to two concepts, e.g., *Fruits* and *Vegetables*, and their conjunction *Fruits And Vegetables*, but also their conjunction when one or both concepts are negated, namely, *Fruits And Not Vegetables*, *Not Fruits And Vegetables*, and *Not Fruits And Not Vegetables*. Our findings sharpen and advance existing analysis on conceptual combinations, revealing systematic deviations from classical (fuzzy set) logic and probability theory. And, more important, our results give further considerable evidence to the validity of our quantum-theoretic framework for the combination of two concepts. Indeed, the representation of conceptual negation naturally arises from the general assumptions of our two-sector Fock space model, and this representation faithfully agrees with the collected data. In addition, we find a new significant and a priori unexpected deviation from classicality, which can exactly be explained by assuming that human reasoning is the superposition of an "emergent reasoning" and a "logical reasoning," and that these two processes are represented in a Fock space algebraic structure.

Keywords: cognition, concept theory, quantum structures, fock space, conceptual emergence, concept formation

1. Introduction

Substantial evidence of presence of quantum structures in processes connected with human behavior and cognition has been put forward in the last decade. More specifically, such quantum structures were identified in situations of decision making and in the structure of language (see e.g., Aerts, 2009; Khrennikov, 2010; Busemeyer and Bruza, 2012; Aerts et al., 2013b; Haven and Khrennikov, 2013; Pothos and Busemeyer, 2013; Wang et al., 2014). The success of this quantum modeling is interpreted as due to "descriptive effectiveness of the mathematical apparatus of quantum theory as formal instrument to model cognitive dynamics and structures in situations where classical set-based approaches are problematical." In particular, the mathematics of quantum theory in Hilbert space has proved very successful in modeling combinations of two concepts (Aerts, 2009; Aerts and Sozzo, 2001, 2014; Aerts and Gabora, 2005a,b; Aerts et al., 2013a,b; Sozzo, 2014, 2015).

The "combination problem," that is, the question of how the representation of the combination of two or more natural concepts can be connected to the representation of the component concepts, has been studied experimentally and within classical concept theories in great

detail in the last 30 years. The main experimental challenges to traditional modeling approaches to concepts combinations are sketched in the following[1].

(i) The "Guppy effect" in concept conjunction, also known as the "Pet-Fish problem" (Osherson and Smith, 1981, 1982). If one measures the typicality of specific exemplars with respect to the concepts *Pet* and *Fish* and their conjunction *Pet-Fish*, then one systematically finds exemplars such as *Guppy* that are very typical examples of *Pet-Fish*, while neither being typical examples of *Pet* nor of *Fish*.

(ii) The deviation from classical (fuzzy) set-theoretic membership weights of exemplars with respect to pairs of concepts and their conjunction or disjunction (Hampton, 1988a,b). If one measures the membership weight of an exemplar with respect to a pair of concepts and their conjunction (disjunction), then one experimentally finds that there is an abundance of cases where the membership weight of the exemplar with respect to the conjunction (disjunction) is greater (less) than the membership weight of the exemplar with respect to at least one of the component concepts.

(iii) The existence of "borderline contradictions" in sentences expressing vague properties (Bonini et al., 1999; Alxatib and Pelletier, 2011). Roughly speaking, a borderline contradiction is a sentence of the form $\mathscr{P}(x) \wedge \neg\mathscr{P}(x)$, for a vague predicate $\mathscr{P}$ and a borderline case x, e.g., the sentence "John is tall and John is not tall." If one investigates how people estimate the truth value of such a sentence, a significant number of them will find it as true, in particular for borderline cases.

What one typically finds in the above situations is a failure of set-theoretic approaches (classical set, fuzzy set, Kolmogorovian probability) to supply satisfactory theoretic models for the experimentally observed patterns. Indeed, all traditional approaches to concept theory [mainly, "prototype theory" (Rosch, 1973, 1978, 1983), "exemplar theory" (Nosofsky, 1988, 1992), and "theory theory" (Murphy and Medin, 1985; Rumelhart and Norman, 1988)] and concept representation [mainly, "extensional" membership-based (Zadeh, 1982; Rips, 1995) and "intensional" attribute-based (Hampton, 1988b; Minsky, 1975; Hampton, 1997)] have structural difficulties to cope with the experimental data exactly where the "graded," or "vague" nature of these data abundantly violates (fuzzy) set-theoretic structures (Osherson and Smith, 1982; Zadeh, 1982), indicating that this violation of set-theoretic structures is

the core of the problem. This situation is experienced as one of the major problems in the domain of traditional concept theories and an obstacle for progress (Komatsu, 1992; Fodor, 1994; Kamp and Partee, 1995; Rips, 1995; Hampton, 1997; Osherson and Smith, 1997).

Important results in concept research and modeling have been obtained in the last decade within the approach of quantum cognition in which our research group has substantially contributed. We cannot report in detail the results attained in our approach, for obvious reasons of space limits. We limit ourselves to summarize the fundamentals and attach relevant bibliographic sources in the following.

(a) The structural aspects of the approach rest on the results of older research on the foundations of quantum theory (Aerts, 1999), the origins of quantum probability (Aerts, 1986; Pitowsky, 1989) and the identification of typically quantum aspects outside the microscopic domain of quantum physics (Aerts and Aerts, 1995; Aerts et al., 2000). A first major step was taken in considering a concept as an "entity in a specific state changing under the influence of a context," rather than as a "container of instantiations," like in most the traditional approaches to concepts. This led to the development of a "State Context Property" (SCoP) formalism for the description of any conceptual entity in terms of its states, contexts and properties. In a cognitive process, such as a typicality estimation, the cognitive context changes the state of the conceptual entity, exactly as in a micro physics process the measurement context changes the state of the quantum particle that is measured. In this perspective, the state p_{Guppy} of the concept *Pet* (*Fish*) scores a low typicality in absence of any context, while it scores a high typicality when the concept *Pet* (*Fish*) is under the context *Fish* (*Pet*). This insight was made concrete by means of an explicit quantum representation of the Guppy effect situation (Aerts and Gabora, 2005a,b).

(b) Continuing in this direction the mathematical formalism of quantum theory was employed to model the overextension and underextension of membership weights measured in Hampton (1988a,b). More specifically, the overextension for conjunctions of concepts measured in Hampton (1988a) was described as an effect of quantum interference and superposition (Aerts, 2009; Aerts et al., 2013b). The existence of superposed quantum states allows for the description of quantum interference, i.e., the deviation from the classically expected pattern, in the two-slit experiment with quantum particles. Analogously, the representation of the conjunction of two concepts by means of a superposed quantum state allows for the modeling of overextension as an expression of quantum interference, i.e., deviation from the classically expected behavior. Quantum interference and superposition also play a primary role in the description of both overextension and underextension for disjunctions of two concepts (Hampton, 1988b). Successively, a two-sector Fock space structure enabled a complete representation of data on conjunctions and disjunctions of two concepts (Aerts, 2009; Aerts et al., 2013b).

---

[1]One typically gains insight into how people combine concepts by gathering data on "typicalities" or "membership weights." To obtain data on "typicalities," participants are given a concept, and a list of instances or exemplars, and asked to pick which exemplar they consider most typical of the concept. A membership weight is instead obtained by asking people to estimate the membership of specific exemplars with respect to a concept. This estimation can, e.g., be quantified by using 7-point (Likert) scale and then converted into a relative frequency and then into a probability called the "membership weight." We have worked on both typicality measurements, as in the analysis of the Guppy effect, and membership weights measurements, as in the analysis of Hampton's experiments and in the present paper.

(c) This quantum-theoretic framework was successfully applied to describe more complex situations, such as borderline vagueness (Sozzo, 2014) and the effects of negation on conceptual conjunction (Sozzo, 2015). In addition, specific conceptual combinations experimentally revealed the presence of further genuine quantum effects, namely, entanglement (Aerts et al., 2013a,b; Aerts and Sozzo, 2001, 2014) and quantum-type indistinguishability (Aerts et al., 2015c). Finally, other phenomena related to concept combination, such as "Ellsberg and Machina decision making paradoxes" (Ellsberg, 1961; Machina, 2009) were successfully modeled in the same quantum-conceptual framework (Aerts et al., 2012, 2014).

There has been very little research on how people interpret and combine negated concepts. In a seminal study, Hampton (1997) considered in a set of experiments both conjunctions of the form *Games Which Are Also Sports* and conjunctions of the form *Games Which Are Not Sports*. His work confirmed overextension in both types of conjunctions, also showing a violation of Boolean classical logical rules for the negation, which has recently been confirmed by ourselves (Sozzo, 2015). In the present paper we extend the collection of data in Sozzo (2015) with the aim of further exploring the use of negation in conceptual combinations and, more generally, the underlying logical structures being at work in human thought in the course of cognitive processes (Aerts et al., 2015a). Let us first put forward a specific comment with respect to the "negation of a concept." From the perspective of prototype theory, for quite some concepts the negation of a concept can be considered as a "singular concept," since it does not have a well defined prototype. In fact, while it is plain to determine the non-membership of, e.g., *Fruit*, this does not seem to lead to the determination involving a similarity with some prototype of *Not Fruit*. Some authors maintain, for this reason, that single negated concepts have little meaning and that conceptual negations can be evaluated only in conjunctions of the form *Fruits Which Are Not Vegetables* (Hampton, 1997). We agree that there is an asymmetry between the way people estimate the membership of an exemplar, e.g., *Apple*, with respect to a positive concept, e.g., *Fruits*, and the way people estimate the membership of the same exemplar with respect to its negative counterpart, e.g., *Not Fruits*. Notwithstanding this, we believe it is meaningful to explicitly introduce the concept *Not Fruits* in our research. First of all, because we do not confine our concept modeling to prototype theory, on the contrary, our approach is more general, the basic structure of prototype theory can be recovered if we limit the concepts to be in their ground states (Aerts and Gabora, 2005a,b). Secondly, we will see that the quantum modeling elaborated in the present paper sheds light exactly on this problem, namely, the "negated concept" only appears as a full concept in "one part of the representation," while is treated as "non-membership with respect to the positive concept" in the other part. Hence, quantum-conceptual framework copes with this problem in a natural way.

Let us proceed by steps, summarizing the major findings in this paper, as follows.

In Section 2 we illustrate design and procedure of the four cognitive experiments we performed. In the first experiment, we tested the membership weights of four sets of exemplars with respect to four pairs (A, B) of concepts and their conjunction "A and B." In the second experiment, we tested the membership weights of the same four sets of exemplars with respect to the same four pairs (A, B) of concepts, but negating the second concept, hence actually considering A, B' and the conjunction "A and B'." In the third experiment, we tested the membership weights obtained considering A', B and the conjunction "A' and B." Finally, in the fourth experiment, we considered the membership weights obtained by negating both concepts, hence actually considering A', B' and the conjunction "A' and B'."

We investigate the representability of the collected data, reported in Appendix A3, in a "single classical Kolmogorovian probability space" (Kolmogorov, 1933). Basic notions and results on probability measures and classical modeling are briefly reviewed in Appendix A1. We prove theorems providing necessary and sufficient conditions for the modeling of the conceptual conjunctions "A and B," "A and B'," "A' and B," and "A' and B'" in such a single classical Kolmogorovian framework. Then, we observe that the data significantly violate our theorems. More specifically, our analysis of classicality for the presence of conjunction and negation together leads to five classicality conditions that should be simultaneously satisfied by the data to fit into one classical probability framework together. When we analyze the deviations of our data with respect to these five conditions we also find a very strong, stable and systematic pattern of violation, i.e., the deviation has the same numerical values even over different pairs of concepts. That the violation is numerically the same independently of the considered pair of concepts indicates that we have identified a non-classical mechanism in human thought which is linked to the depth of concept formation itself, independent of the specific meaning for a specific pair of concepts and a specific set of considered exemplars. This was for ourselves a first surprising and unexpected finding, and we have recently devoted an article to investigate it in depth (Aerts et al., 2015b).

A second major and equally unexpected finding was that the numerical size of the "deviation of classicality pattern" can exactly be predicted in our quantum-theoretic model in two-sector Fock space. And, more, it can be explained by assuming that human reasoning is the superposition of two simultaneous processes, a "logical reasoning" and a "conceptual," or "emergent," "reasoning." Logical reasoning combines cognitive entities (concepts, combinations of concepts, propositions, etc.) by applying the rules of logic, though generally in a probabilistic way. Emergent reasoning instead enables formation of combined cognitive entities as newly emerging entities (new concepts, new propositions, etc.), carrying new meaning, linked to the meaning of the component cognitive entities, but with a connection not defined by the algebra of logic. Emergent reasoning can be modeled in first sector of Fock space and, at variance with widespread beliefs, is dominant in our approach. Logical reasoning can be modeled in second sector of Fock space, hence one expects that classical logical rules

hold in this sector, like we explicitly prove here for conceptual conjunctions and negations (see also Aerts et al., 2015b).

Our quantum-theoretic model in two-sector Fock space for conceptual negations and conjunctions is elaborated in Section 3. It naturally extends the model in Aerts (2009) and follows the general lines traced in Sozzo (2014, 2015). It is however important to notice that the simultaneous modeling of conjunction and negation requires the introduction of two new conceptual steps which were not needed in the modeling of conjunction pairs: (i) the introduction of entangled states in second sector of Fock space, which enables formalizing the situation where probabilities in second sector can be formed by a "product procedure," even if they are not independent—this is an aspect of the Fock space model we had not understood in our earlier modeling, hence we could consider it a further new surprising finding of the investigation presented in this paper; (ii) the handling of "negation" in second sector by "logical inversion," similarly like we handled conjunction in second sector by "product," more concretely, an experiment with "negation" with respect to a concept is treated by "negating logically" an experiment on the concept itself. This is also the way in which our Fock space model naturally copes with the general non-prototypicality of a negated concept, as already mentioned above.

We see in Section 4 that a large amount of data can be faithfully represented in our two-sector Fock space, and construct an explicit representation for some relevant cases that are classically problematical. A complete representation of the data is provided in the Supplementary Material attached to this paper. As we can see the findings presented in this paper provide strong and independent confirmations to our quantum-theoretic framework, and we devote Section 5 to comment on our results and extensively discuss novelties and corroboration of our approach. Technical appendices A4 and A5 complete the paper.

2. Description of Experiments and Classicality Analysis

James Hampton identified in his cognitive tests systematic deviations from classical (fuzzy) set predictions for membership weights of exemplars with respect to conjunctions and disjunctions of two concepts, and named these deviations "overextensions" and "underextensions" (Hampton, 1988a,b). Cases of "double overextension" were also observed. More explicitly, if the membership weight of an exemplar x with respect to the conjunction "A and B" of two concepts A and B is higher than the membership weight of x with respect to one concept (both concepts), we say that the membership weight of x is "overextended" ("double overextended") with respect to the conjunction (by abuse of language, we say that x is overextended (double overextended) with respect to the conjunction, in this case). If the membership weight of an exemplar x with respect to the disjunction "A or B" of two concepts A and B is less than the membership weight of x with respect to one concept (both concepts), we say that the membership weight of x is "underextended" ("double underextended") with respect to the disjunction (by abuse of language, we say that x is underextended

(double underextended) with respect to the disjunction, in this case).

Similar effects were identified by Hampton in his experiments on conjunction and negation of two concepts (Hampton, 1997). The analysis in Aerts (2009) revealed further deviations from classicality in Hampton's experiments, due to the impossibility to generally represent the collected data in a classical probability framework satisfying the axioms of Kolmogorov. In Sozzo (2015) we moved along this direction and performed an experiment in which we tested both conjunctions of the form "A and B" and conjunctions of the form "A and B'," for specific pairs (A, B) of natural concepts. We showed that very similar deviations from classicality are observed in our experiment too.

In the present paper we aim to generalize the results in Sozzo (2015), providing an extensive analysis of conceptual conjunction and negation and investigating their reciprocal influences. To this end we complete the experiment in Sozzo (2015) by performing a more general cognitive test, as described in the following sections.

2.1. Participants and Design

The participants to our experimental study—40 persons, chosen among our colleagues and friends—were asked to fill in a questionnaire in which they had to estimate the membership of four different sets of exemplars with respect to four different pairs (A, B) of natural concepts, and their conjunctions "A and B," "A and B'," "A' and B," and "A' and B'," where A' and B' denote the negations of the concepts A and B, respectively. We devised a "within-subjects design" for our experiments, hence all participants were exposed to every treatment or condition. The participants were presented with a preliminary text where we made explicit, by means of suitable examples, what one usually means by "membership of an exemplar with respect to a specific conceptual category." Further, we chose participants with different backgrounds, not only academics, to avoid issues connected with "selection biases."

We considered four pairs of natural concepts, namely (*Home Furnishing, Furniture*), (*Spices, Herbs*), (*Pets, Farmyard Animals*), and (*Fruits, Vegetables*). For each pair, we considered 24 exemplars and measured their membership with respect to these pairs of concepts and the conjunctions of these pairs mentioned above.

Conceptual membership was estimated by using a "7-point scale." The participants were asked to choose a number from the set $+3, +2, +1, 0, -1, -2, -3$, where the positive numbers $+1, +2,$ and $+3$ meant that they considered "the exemplar to be a member of the concept"—$+3$ indicated a strong membership, $+1$ a relatively weak membership. The negative numbers $-1, -2,$ and -3 meant that the participant considered "the exemplar to not be a member of the concept"—-3 indicated a strong non-membership, -1 a relatively weak non-membership.

Although we explicitly measured the "amount of membership" on a 7-point scale, for the scopes of this paper, we only need the data of a sub-experiment, namely the one tested for "membership" or "non-membership"—our plan is to use the "amount of membership data" for a following study leading to a graphical representation of the data, as we did with Hampton's data for the disjunction in earlier work (Aerts et al., 2013a,b).

A second reason for this specific form of the experiments is that we wanted to stay as close as possible to the disjunction experiments by Hampton (1988b), since we plan to investigate later the connections of our conjunction data with Hampton's disjunction data, for example to investigate the way in which the "de Morgan laws" take form in our modeling of the data. This is why we performed the full experiment measuring "amount of membership" and also testing simultaneously for "membership or non-membership," while it is only the latter sub-experiment that we use in the investigation presented in this article. The data of this sub-experiment give rise to relative frequencies for testing membership or not membership, which means that we can interpret them as probabilities in the limit of large numbers. More concretely, $\mu(A \text{ and } B)$ is the large number limit of the relative frequency for x to be a member of "A and B" in the performed experiment. We get to this by converting the values collected on the 7-point scale by associating a value $+1$ to each positive value on the 7-point scale, -1 to each negative number, and 0.5 to each 0 on the same 7-point scale.

2.2. Procedure and Materials
This experimental study was carried out in accordance with the recommendations of the "University of Leicester Code of Practice and Research Code of Conduct, Research Ethics Committee of the School of Management" with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. For each pair (A, B) of natural concepts, the 40 participants were involved in four subsequent experiments, e_{AB}, $e_{AB'}$, $e_{A'B}$, and $e_{A'B'}$, corresponding to the conjunctions "A and B," "A and B'," "A' and B," and "A' and B'," respectively. More specifically, the four sequential experiments can be illustrated as follows.

For the conceptual pair (*Home Furnishing*, *Furniture*), we firstly asked the 40 participants to estimate the membership of the first set of 24 exemplars with respect to the concepts *Home Furnishing*, *Furniture*, and their conjunction *Home Furnishing And Furniture*. Then, we asked the same 40 participants to estimate the membership of the same set of 24 exemplars with respect to the concept *Home Furnishing*, the negation *Not Furniture* of the concept *Furniture*, and their conjunction *Home Furnishing And Not Furniture*. Subsequently, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negation *Not Home Furnishing* of the concept *Home Furnishing*, the concept *Furniture*, and their conjunction *Not Home Furnishing And Furniture*. Finally, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negations *Not Home Furnishing*, *Not Furniture*, and their conjunction *Not Home Furnishing And Not Furniture*. The corresponding membership weights are reported in **Table A1**.

For the conceptual pair (*Spices*, *Herbs*), we firstly asked the 40 participants to estimate the membership of the second set of 24 exemplars with respect to the concepts *Spices*, *Herbs*, and their conjunction *Spices And Herbs*. Then, we asked the same 40 participants to estimate the membership of the same set of 24 exemplars with respect to the concept *Spices*, the negation *Not Herbs* of the concept *Herbs*, and their

conjunction *Spices And Not Herbs*. Subsequently, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negation *Not Spices* of the concept *Spices*, the concept *Herbs*, and their conjunction *Not Spices And Herbs*. Finally, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negations *Not Spices*, *Not Herbs*, and their conjunction *Not Spices And Not Herbs*. The corresponding membership weights are reported in **Table A2**.

For the conceptual pair (*Pets*, *Farmyard Animals*), we firstly asked the 40 participants to estimate the membership of the third set of 24 exemplars with respect to the concepts *Pets*, *Farmyard Animals*, and their conjunction *Pets And Farmyard Animals*. Then, we asked the same 40 participants to estimate the membership of the same set of 24 exemplars with respect to the concept *Pets*, the negation *Not Farmyard Animals* of the concept *Farmyard Animals*, and their conjunction *Pets And Not Farmyard Animals*. Subsequently, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negation *Not Pets* of the concept *Pets*, the concept *Farmyard Animals*, and their conjunction *Not Pets And Farmyard Animals*. Finally, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negations *Not Pets*, *Not Farmyard Animals*, and their conjunction *Not Pets And Not Farmyard Animals*. The corresponding membership weights are reported in **Table A3**.

For the conceptual pair (*Fruits*, *Vegetables*), we firstly asked the 40 participants to estimate the membership of the third set of 24 exemplars with respect to the concepts *Fruits*, *Vegetables*, and their conjunction *Fruits And Vegetables*. Then, we asked the same 40 participants to estimate the membership of the same set of 24 exemplars with respect to the concept *Fruits*, the negation *Not Vegetables* of the concept *Vegetables*, and their conjunction *Fruits And Not Vegetables*. Subsequently, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negation *Not Fruits* of the concept *Fruits*, the concept *Vegetables*, and their conjunction *Not Fruits And Vegetables*. Finally, we asked the 40 participants to estimate the membership of the 24 exemplars with respect to the negations *Not Fruits*, *Not Vegetables*, and their conjunction *Not Fruits And Not Vegetables*. The corresponding membership weights are reported in **Table A4**.

2.3. Methodology
A first inspection of tables **Tables A1–A4** already reveals that some exemplars present overextension with respect to all conjunctions "A and B," "A and B'," "A' and B," "A' and B'." This is the case, e.g., for the exemplar *Lamp* with respect to the concepts *Home Furnishing* and *Furniture* (**Table A1**), the exemplar *Salt* with respect to *Spices* and *Herbs* (**Table A2**), the exemplar *Goldfish* with respect to *Pets* and *Farmyard Animals* (**Table A3**), and the exemplar *Mustard* with respect to *Fruits* and *Vegetables* (**Table A4**). Hence, manifest deviations from classicality occurred in our experiments. When we say "deviations from classicality," we actually mean that the collected data behave in such a way that they cannot generally be modeled by using the usual connectives of classical (fuzzy set) logic for conceptual conjunctions, neither

the rules of classical probability for their membership weights. In order to systematically identify such deviations from classicality we need however a characterization of the representability of these data in a classical probability space. To this end we derive in the following step by step conditions that will give us an overall picture of the classicality of conceptual conjunctions and negations. Finally, we arrive to a set of five conditions, formulated in Theorem 3 as a set of necessary and sufficient conditions of classicality for a pair of concepts, its negations and conjunctions to be representable within a classical Kolmogorovian probability model. Symbols and notions are introduced in Appendix A1. Let us mention that to our knowledge the "necessary and sufficient conditions for probabilities $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ to be represented within in classical Kolmogorovian probability model, have not yet been systematically derived, and hence are not known in the literature. However, the "necessary and sufficient conditions for probabilities $\mu(A)$, $\mu(B)$ and $\mu(A \text{ and } B)$ to be represented within in classical Kolmogorovian probability model have been systematically studied (Pitowsky, 1989), their direct derivation can for example be found in Aerts (2009), theorem 1 of Section 1.3. We will start our investigation of the classicality condition by making use of the conditions that could be derived for $\mu(A)$, $\mu(B)$, and $\mu(A \text{ and } B)$ and applying them additionally to $\mu(A')$, $\mu(B')$, and $\mu(A' \text{ and } B')$, and to add some intermediate conditions connecting $\mu(A)$, $\mu(B)$, and $\mu(A \text{ and } B)$ and $\mu(A')$, $\mu(B')$, and $\mu(A' \text{ and } B')$, to also imply classicality for the mixed situations such as $\mu(A)$, $\mu(B')$, and $\mu(A \text{ and } B')$. We can prove the following theorems (see also Appendix A4).

Theorem 1. *The membership weights $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ of an exemplar x with respect to the concepts A, B, the negations "not A," "not B," the conjunctions "A and B," "A and B'," "A' and B," and "A' and B'" are classical conjunction data i.e., can be represented in a classical Kolmogorovian probability model, if and only if they satisfy the following conditions.*

$$0 \leq \mu(A \text{ and } B) \leq \mu(A) \leq 1 \tag{1}$$
$$0 \leq \mu(A \text{ and } B) \leq \mu(B) \leq 1 \tag{2}$$
$$0 \leq \mu(A' \text{ and } B') \leq \mu(A') \leq 1 \tag{3}$$
$$0 \leq \mu(A' \text{ and } B') \leq \mu(B') \leq 1 \tag{4}$$
$$\mu(A) - \mu(A \text{ and } B) = \mu(B') - \mu(A' \text{ and } B') \tag{5}$$
$$= \mu(A \text{ and } B')$$
$$\mu(B) - \mu(A \text{ and } B) = \mu(A') - \mu(A' \text{ and } B') \tag{6}$$
$$= \mu(A' \text{ and } B)$$
$$1 - \mu(A) - \mu(B) + \mu(A \text{ and } B) = \mu(A' \text{ and } B') \tag{7}$$
$$1 - \mu(A') - \mu(B') + \mu(A' \text{ and } B') = \mu(A \text{ and } B) \tag{8}$$

Theorem 2. *The membership weights $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ of an exemplar x with respect to the concepts A, B, A', and B' and the*

conjunctions "A and B," "A and B'," "A' and B," and "A' and B'" are classical conjunction data if and only if they satisfy the following conditions.

$$0 \leq \mu(A \text{ and } B) \leq \mu(A) \leq 1 \tag{9}$$
$$0 \leq \mu(A \text{ and } B) \leq \mu(B) \leq 1 \tag{10}$$
$$\mu(A) - \mu(A \text{ and } B) = \mu(B') - \mu(A' \text{ and } B') \tag{11}$$
$$= \mu(A \text{ and } B')$$
$$\mu(B) - \mu(A \text{ and } B) = \mu(A') - \mu(A' \text{ and } B') \tag{12}$$
$$= \mu(A' \text{ and } B)$$
$$0 \leq 1 - \mu(A) - \mu(B) + \mu(A \text{ and } B) = \mu(A' \text{ and } B') \tag{13}$$

The classicality requirements in Theorems 1 and 2 are not symmetric with respect to the exchange of A with A' and B with B'. Thus, we can look for equivalent and more symmetric sets of requirements. These include validity of the "marginal law" of classical probability. We see this in Theorem 3, whose proof preliminarily requires the following lemma.

Lemma 1. The four equalities defined in Equations (5) and (6) are equivalent with the following four equalities expressing the marginal law for all elements to be satisfied.

$$\mu(A) = \mu(A \text{ and } B) + \mu(A \text{ and } B') \tag{14}$$
$$\mu(B) = \mu(A \text{ and } B) + \mu(A' \text{ and } B) \tag{15}$$
$$\mu(A') = \mu(A' \text{ and } B') + \mu(A' \text{ and } B) \tag{16}$$
$$\mu(B') = \mu(A' \text{ and } B') + \mu(A \text{ and } B') \tag{17}$$

Theorem 3. *The membership weights $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ of an exemplar x with respect to the concepts A, B, A', and B' and the conjunctions "A and B," "A and B'," "A' and B," and "A' and B'" are classical conjunction data if and only if they satisfy the following conditions.*

$$0 \leq \mu(A \text{ and } B) \leq \mu(A) \leq 1 \tag{18}$$
$$0 \leq \mu(A \text{ and } B) \leq \mu(B) \leq 1 \tag{19}$$
$$\mu(A) = \mu(A \text{ and } B) + \mu(A \text{ and } B') \tag{20}$$
$$\mu(B) = \mu(A \text{ and } B) + \mu(A' \text{ and } B) \tag{21}$$
$$\mu(A') = \mu(A' \text{ and } B') + \mu(A' \text{ and } B) \tag{22}$$
$$\mu(B') = \mu(A' \text{ and } B') + \mu(A \text{ and } B) \tag{23}$$
$$0 \leq 1 - \mu(A \text{ and } B) - \mu(A \text{ and } B') - \mu(A' \text{ and } B) \tag{24}$$
$$= \mu(A' \text{ and } B')$$

The conditions above can be further simplified by observing that the membership weights we collected in our experiments are large number limits of relative frequencies, thus all measured quantities are already contained in the interval [0, 1]. Therefore, we have

$$\mu(A), \mu(B), \mu(A'), \mu(B'), \mu(A \text{ and } B), \mu(A \text{ and } B'), \tag{25}$$
$$\mu(A' \text{ and } B), \mu(A' \text{ and } B') \in [0, 1]$$

Now, when Equation (25) is satisfied, we have that from Equations (21) and (22) follows that

$$\mu(A \text{ and } B) \leq \mu(A)$$
$$\mu(A \text{ and } B) \leq \mu(B)$$

This entails that Equations (18) and (19) are satisfied, when Equations (21) and (22) are. Hence, we can amazingly enough formulate Theorem 3 a new, with only five conditions to be satisfied—four conditions expressing the marginal law.

Theorem 3′. *If the membership weights $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ of an exemplar x with respect to the concepts A, B, A′, and B′ and the conjunctions "A and B," "A and B′," "A′ and B," and "A′ and B′" are all contained in the interval $[0, 1]$, they are classical conjunction data if and only if they satisfy the following conditions.*

$$\mu(A) = \mu(A \text{ and } B) + \mu(A \text{ and } B') \quad (26)$$
$$\mu(B) = \mu(A \text{ and } B) + \mu(A' \text{ and } B) \quad (27)$$
$$\mu(A') = \mu(A' \text{ and } B') + \mu(A' \text{ and } B) \quad (28)$$
$$\mu(B') = \mu(A' \text{ and } B') + \mu(A \text{ and } B') \quad (29)$$
$$\mu(A \text{ and } B) + \mu(A \text{ and } B') + \mu(A' \text{ and } B) \quad (30)$$
$$+ \mu(A' \text{ and } B') = 1$$

Equations (26–30) express classicality conditions in their most symmetric form. A more traditional way to quantify deviations from classical conjunction in real data is resorting to the following parameters.

$$\Delta_{AB} = \mu(A \text{ and } B) - \min\{\mu(A), \mu(B)\} \quad (31)$$
$$\Delta_{AB'} = \mu(A \text{ and } B') - \min\{\mu(A), \mu(B')\} \quad (32)$$
$$\Delta_{A'B} = \mu(A' \text{ and } B) - \min\{\mu(A'), \mu(B)\} \quad (33)$$
$$\Delta_{A'B'} = \mu(A' \text{ and } B') - \min\{\mu(A'), \mu(B')\} \quad (34)$$

In fact, the quantities Δ_{AB}, $\Delta_{AB'}$, $\Delta_{A'B}$, and $\Delta_{A'B'}$ typically measure overextension with respect to the conjunctions "A and B," "A and B′," "A′ and B," and "A′ and B′," respectively (Hampton, 1988a). However, overextension-type deviations are generally not the only way in which membership for conjunction of concepts can deviate from classicality. Let us now introduce the following quantities:

$$k_{AB} = 1 - \mu(A) - \mu(B) + \mu(A \text{ and } B) \quad (35)$$
$$k_{AB'} = 1 - \mu(A) - \mu(B') + \mu(A \text{ and } B') \quad (36)$$
$$k_{A'B} = 1 - \mu(A') - \mu(B) + \mu(A' \text{ and } B) \quad (37)$$
$$k_{A'B'} = 1 - \mu(A') - \mu(B') + \mu(A' \text{ and } B') \quad (38)$$

The quantities k_{AB}, $k_{AB'}$, $k_{A'B}$, and $k_{A'B'}$ have been named "Kolmogorovian conjunction factors" and studied in detail in Aerts (2009). The Kolmogorovian factors measure a deviation that can be understood as of "opposite type" than the deviation measured by the overextension. Namely, the condition for k_{AB} is

violated when both $\mu(A)$ and $\mu(B)$ are "too large" compared with $\mu(A \text{ and } B)$. Finally, we introduce a new type of quantities that measure the deviations of classicality as expressed by Equations (27–30), hence essentially deviations from the marginal law of classical probability:[2]

$$I_{ABA'B'} = 1 - \mu(A \text{ and } B) - \mu(A \text{ and } B') - \mu(A' \text{ and } B) \quad (39)$$
$$- \mu(A' \text{ and } B')$$
$$I_A = \mu(A) - \mu(A \text{ and } B) - \mu(A \text{ and } B') \quad (40)$$
$$I_B = \mu(B) - \mu(A \text{ and } B) - \mu(A' \text{ and } B) \quad (41)$$
$$I_{A'} = \mu(A') - \mu(A' \text{ and } B') - \mu(A' \text{ and } B) \quad (42)$$
$$I_{B'} = \mu(B') - \mu(A' \text{ and } B') - \mu(A \text{ and } B') \quad (43)$$

Finally, Theorem 3′ can be reformulated by means of the introduced parameters as follows.

Theorem 3″. *If the membership weights $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ of an exemplar x with respect to concepts A, B, A′, and B′ and the conjunctions "A and B," "A and B′," "A′ and B," and "A′ and B′," are all contained in the interval $[0, 1]$, they are classical conjunction data if and only if*

$$I_{ABA'B'} = I_A = I_B = I_{A'} = I_{B'} = 0 \quad (44)$$

2.4. Results

Let us now come back to our experiments. Theorems 1–3 are manifestly violated in several cases, and we report in Appendix A3 the relevant conditions that should hold in a classical setting. Since the conditions $k_{AB} > 0$, $k_{AB'} > 0$, $k_{A'B} > 0$, and $k_{A'B'} > 0$ are always satisfied, they are not explicitly inserted in **Tables A1–A4**. On the contrary, Δ_{XY}, I_X, I_Y, $X = A, A'$, $Y = B, B'$, and $I_{ABA'B'}$ are systematically violated. This means that deviations from a classical probability model in our experimental data are due to both overextension in the conjunctions and violations of classicality in the negations. We consider some relevant cases in the following.

The exemplar *Apple* scores $\mu(A) = 1$ with respect to the concept *Fruits*, $\mu(B) = 0.23$ with respect to the concept *Vegetables*, and $\mu(A \text{ and } B) = 0.6$ with respect to the conjunction *Fruits And Vegetables*, hence it has $\Delta_{AB} = 0.38$ (**Table A4**). The exemplar *Prize Bull* scores $\mu(A) = 0.13$ with respect to *Pets*, $\mu(B) = 0.76$ with respect to the concept *Farmyard Animals*, and $\mu(A \text{ and } B) = 0.43$ with respect to the conjunction *Pets And Farmyard Animals*, hence it has $\Delta_{AB} = 0.29$ (**Table A3**). The membership weight of *Chili Pepper* with respect to *Spices* is 0.98, with respect to *Herbs* is 0.53, while its membership weight with respect to the conjunction *Spices And Herbs* is 0.8, hence $\Delta_{AB} = 0.27$, thus giving rise to overextension (**Table A2**). Even stronger deviations are observed in the combination *Fruits And Vegetables*. For example, the exemplar *Broccoli* scores 0.09 with respect to *Fruits*, 1 with respect to *Vegetables*, and 0.59 with

---

[2] Remark that, if we set $I_{AA'} = 1 - \mu(A) - \mu(A')$ and $I_{BB'} = 1 - \mu(B) - \mu(B')$, we have $I_{AA'} = I_{ABA'B'} - I_A - I_{A'}$ and $I_{BB'} = I_{ABA'B'} - I_B - I_{B'}$, which means that the parameters $I_{AA'}$ and $I_{BB'}$ used in Sozzo (2015) can be derived from the parameters $I_{ABA'B'}$, I_A, I_B, $I_{A'}$, If the membership weights and $I_{B'}$.

respect to *Fruits And Vegetables* ($\Delta_{AB} = 0.49$). A similar pattern is observed for *Parsley*, which scores 0.02 with respect to *Fruits*, 0.78 with respect to *Vegetables* and 0.45 with respect to *Fruits And Vegetables* ($\Delta_{AB} = 0.43$, **Table A4**).

Overextension is present when one concept is negated. More explicitly:

(i) in the conjunction "*A* and *B'*." Indeed, the membership weights of *Shelves* with respect to *Home Furnishing*, *Not Furniture*, and *Home Furnishing And Not Furniture* is 0.85, 0.13, and 0.39, respectively, for a $\Delta_{AB'} = 0.26$ (**Table A1**). Then, *Pepper* scores 0.99 with respect to *Spices*, 0.58 with respect to *Not Herbs*, and 0.9 with respect to *Spices and Not Herbs*, for a $\Delta_{AB'} = 0.32$ (**Table A2**). Finally, *Doberman Guard Dog* gives 0.88 and 0.27 with respect to *Pets* and *Not Farmyard Animals*, respectively, while it scores 0.55 with respect to *Pets And Not Farmyard Animals*, hence it scores $\Delta_{AB'} = 0.28$ (**Table A3**).

(ii) in the conjunction "*A'* and *B*." Indeed, the membership weights of *Desk* with respect to *Not Home Furnishing*, *Furniture* and *Not Home Furnishing And Furniture* is 0.31, 0.95, and 0.75, respectively, for a $\Delta_{A'B} = 0.44$ (**Table A1**). The exemplar *Oregano* scores 0.21 with respect to *Not Spices*, 0.86 with respect to *Herbs*, and 0.5 with respect to *Not Spices and Herbs*, for a $\Delta_{A'B} = 0.29$ (**Table A2**). Finally, again *Doberman Guard Dog* gives 0.14 and 0.76 with respect to *Not Pets* and *Farmyard Animals*, respectively, while it scores 0.45 with respect to *Not Pets And Farmyard Animals*, hence it scores $\Delta_{A'B} = 0.45$ (**Table A3**).

When two concepts are negated—"*A'* and *B''*"—we have, for example, $\mu(A') = 0.12$, $\mu(B') = 0.81$ and $\mu(A' \text{ and } B') = 0.43$ for *Goldfish*, with respect to *Not Pets* and *Not Farmyard Animals*, hence $\Delta_{A'B'} = 0.31$, in this case (**Table A3**). More, the exemplar *Garlic* scores $\mu(A') = 0.88$ with respect to *Not Fruits* and $\mu(B') = 0.24$ with respect to *Not Vegetables*, and $\mu(A' \text{ and } B') = 0.45$ with respect to *Not Fruits And Not Vegetables*, for a $\Delta_{A'B'} = 0.21$ (**Table A4**).

Double overextension is also present in various cases. For example, the membership weight of *Olive* with respect to *Fruits And Vegetables* is 0.65, which is greater than both 0.53 and 0.63, i.e., the membership weights of *Olive* with respect to *Fruits* and *Vegetables*, respectively (**Table A4**). Furthermore, *Prize Bull* scores 0.13 with respect to *Pets* and 0.26 with respect to *Not Farmyard Animals*, but its membership weight with respect to *Pets And Not Farmayard Animals* is 0.28 (**Table A3**). Also, *Door Bell* gives 0.32 with respect to *Not Home Furnishing* and 0.33 with respect to *Furniture*, while it gives 0.34 with respect to *Not Home Furnishing And Furniture*.

Significant deviations from classicality are also due to conceptual negation, in the form of violation of the marginal law of classical probability theory. By again referring to **Tables A1–A4**, we have that the exemplar *Field Mouse* has $I_{ABA'B'}$ in Equation (40) equal to -0.46 (**Table A3**), while the exemplar *Doberman Guard Dog* has $I_{ABA'B'} = -1.03$ (**Table A3**). Both exemplars thus violate Equation (30). Analogously, *Chili Pepper* has I_A in Equation (40) equal to -0.73 (**Table A2**), hence it violates Equation (26), while *Pumpkin* has $I_{B'}$ in Equation (43) equal to -0.13 (**Table A4**), hence it violates Equation (29).

We performed a statistical analysis of the data, estimating the probability that the experimentally identified deviations from classicality would be due to chance. We specifically considered the classicality conditions Equations (26–30) with the aim to prove that the deviations I_X, $X = A, A'$, I_Y, $Y = B, B'$ and $I_{ABA'B'}$ in Equations (40–43) were statistically significant. We firstly performed a "two-tail *t*-test for paired two samples for means" to test deviations from the marginal law of classical probability, that is, we tested violations of Equations (26–29) by comparing $\mu(X)$ with respect to $\sum_{Y=B,B'} \mu(X, Y)$, $X = A, A'$, and $\mu(Y)$ with respect to $\sum_{X=A,A'} \mu(X \text{ and } Y)$, $Y = B, B'$. Then, we performed a "two-tail *t*-test for one sample for means" to test $\sum_{X=A,A'} \sum_{Y=B,B'} \mu(X \text{and } Y)$ with respect to 1. The corresponding *p*-values for df $= 37$ are reported in **Tables A5A–E**. Due to the high number of multiple comparisons—24 null hypotheses were tested for each pair (X, Y) of concepts—we applied a "Bonferroni correction procedure" to avoid the so-called "family-wise error rate" (FWER). Hence, we compared the obtained *p*-values with the reference value $0.05/24 \approx 0.002$. We found *p*-values systematically much lower than this reference value, for all exemplars and pairs of concepts, which makes it possible to conclude that the experimentally tested deviations from classicality are not due to chance.

In addition, our data analysis reveals a new, fundamental and a priori unexpected deviation from classicality. The numerical values of I_A, I_B, $I_{A'}$, $I_{B'}$, and $I_{ABA'B'}$ in Equations (40–43) are reported in Aerts et al. (2015b). They are such that the corresponding pattern of violation exhibits specific features:

(i) it cannot be explained by means of traditional classical probabilistic approaches, since we should have $I_A = I_B = I_{A'} = I_{B'} = I_{ABA'B'} = 0$, in that case (see Theorem 3″).

(ii) it is "highly stable," in the sense that the functions I_A, I_B, $I_{A'}$, $I_{B'}$, and $I_{ABA'B'}$ are very likely between -1 and 0;

(iii) is is "systematic," in the sense that the values of I_A, I_B, $I_{A'}$, $I_{B'}$, and $I_{ABA'B'}$ are approximately the same across all exemplars;

(iv) it is "regular," in the sense that the functions I_A, I_B, $I_{A'}$, $I_{B'}$, and $I_{ABA'B'}$ do not depend on the pair of concepts that are considered.

Observations (i–iv) were for us a clue that I_A, I_B, $I_{A'}$, $I_{B'}$, and $I_{ABA'B'}$ are constant functions across all exemplars and pairs of concepts. This is indeed the case, as we have proved in Aerts et al. (2015b) by means of a "linear regression statistical analysis." This pattern is so unexpectedly stable, systematic and regular, being independent of exemplars, concepts and conceptual connectives, that it constitutes for us a fundamental new finding. We believe that this deviation from classicality occurs at a deeper level than the known deviations due to overextension and underextension, and that it expresses a fundamental mechanism of concept formation.

These results could already be considered as crucial for claiming that the violation of classicality occurs at a deep structural conceptual level, but this is not the end of the story. We will see in Section 5 that the stability of this violation can exactly be explained in a quantum-theoretic framework in two-sector Fock space elaborated by ourselves. Hence, we devote Sections 3 and 4 to expose this modeling framework (the essentials of the

formalism we apply are reviewed in Appendix A2, and we refer to it for symbols and notation).

3. Quantum Modeling Conceptual Conjunctions and Negations

In Aerts (2009) we proved that a big amount of the experimental data collected in Hampton (1988a,b) on conjunctions and disjunctions of two concepts can be modeled by using the mathematical formalism of quantum theory. A two-sector Fock space then provided an optimal algebraic setting for this modeling. In Sozzo (2014) we proved that this quantum-theoretic framework was suitable to model the data collected in Alxatib and Pelletier (2011) on conjunctions of the form "A and A'," and in Sozzo (2015) we were able to prove that also the experimental data collected on conjunctions of the form "A and B" and "A and B'," for specific pairs (A, B) of concepts can be represented by using the same quantum mathematics. However, a complete modeling of data on both conjunctions and negations requires performing new experiments, where the conceptual conjunctions "A and B" and "A and B'" are tested together with the conceptual conjunctions "A' and B" and "A' and B'." The complete collection of these experiments has been discussed in Section 2. As anticipated in Section 1, we undertake this modeling task here. It is natural to observe that the modeling in Aerts (2009) needs a suitable generalization, since conceptual negation should be taken into account as well. But, we will see later in this section that such a generalization is completely compatible with the original model, because it rests on the assumption that (probabilistic versions of) logical rules hold only in second sector of Fock space. By introducing this quite natural assumption, we were able to model conceptual conjunctions and disjunctions in Fock space. We show now that conceptual conjunctions and negations can be modeled in Fock space by introducing the same assumption.

To model conceptual negations we also need a new theoretical step which was not necessary in our previous formulations, namely, the introduction of "entangled states" in second sector of Fock space to formalize situations where the membership weights are not independent. This introduction, together with the application of quantum logical rules in second sector of Fock space, are compatible with previous formulations, but they make our generalization in this paper highly non-obvious. We will extensively discuss the novelties of the present modeling in the next sections. Let us first proceed with our mathematical construction.

Let us denote by $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ the membership weights of a given exemplar x with respect to the concepts A, B, the negations A', B' and the conjunctions "A and B," "A and B'," "A' and B," and "A' and B'," respectively.

The decision measurement testing whether a specific exemplar x is a member or not a member of a concept A is represented by the spectral decomposition of the identity consisting of the two orthogonal projection operators M (generally depending on x, we omit such dependence, for the sake of brevity) and $\mathbb{1} - M$ defined in a complex Hilbert space $\mathcal{H}$. The concepts A and B are

represented by orthogonal unit vectors $|A\rangle$ and $|B\rangle$, respectively, of $\mathcal{H}$. Hence we have

$$\langle A|A\rangle = \langle B|B\rangle = 1 \quad \langle A|B\rangle = 0 \tag{45}$$

By using standard rules for quantum probabilities (see Appendix A2), we have the following

$$\mu(A) = \langle A|M|A\rangle \quad \mu(B) = \langle B|M|B\rangle \tag{46}$$

where $\mu(A)$ and $\mu(B)$ are the measured membership weights of x with respect to the concepts A and B, respectively, in the performed experiment.

The conceptual negations A' and B' are represented by another pair of orthogonal unit vectors $|A'\rangle$ and $|B'\rangle$, respectively, such that the set $\{|A\rangle, |B\rangle, |A'\rangle, |B'\rangle\}$, is an orthonormal set. Hence we have

$$\mu(A') = \langle A'|M|A'\rangle \quad \mu(B') = \langle B'|M|B'\rangle \tag{47}$$

where $\mu(A')$ and $\mu(B')$ are the measured membership weights of x with respect to the negations A' and B' of the concepts A and B, respectively, in the performed experiment.

3.1. The First Sector Analysis

Let us first analyze the situation where we look for a modeling solution in the Hilbert space $\mathcal{H}$ —which for our complete quantum model in Fock space will be the first sector of this Fock space, as we will show in detail later. In the Hilbert space $\mathcal{H}$, the concepts "A and B," "A and B'," "A' and B," and "A' and B'" are respectively represented by the superposition vectors[3] $\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$, $\frac{1}{\sqrt{2}}(|A\rangle + |B'\rangle)$, $\frac{1}{\sqrt{2}}(|A'\rangle + |B\rangle)$, and $\frac{1}{\sqrt{2}}(|A'\rangle + |B'\rangle)$.

Let us analyze in detail the aspects of this situation with the aim of resulting in a view on the possible solutions. Geometric considerations induce to observe that, if we look for a solution in the complex Hilbert space $\mathbb{C}^8$, we will find the most general type of solution. Indeed, since we consider four orthonormal vectors $|A\rangle$, $|A'\rangle$, $|B\rangle$, and $|B'\rangle$, our Hilbert space will contain a four dimensional subspace generated by these vectors. Further we have two orthogonal projection operators M and $\mathbb{1} - M$, that work on this four dimensional subspace. The image of a projection operator has dimension not bigger than the definition domain of it, which means that the image of M of the four dimensional subspace is at maximum equal to four, and this is also the case for the image of $\mathbb{1} - M$ of the four dimensional subspace. Since M and $\mathbb{1} - M$ are orthogonal, this can give rise to a eight dimensional subspace, but not more. It means that we can incorporate all what we need in an eight dimensional complex Hilbert space. This is the reason that we look for our representation starting with $\mathbb{C}^8$, knowing that the choice of a Hilbert space with more than eight dimensions would not add degrees of freedom that can give rise to additional solutions to those that can be found in $\mathbb{C}^8$. Hence, we explicitly use this Hilbert space in what follows reminding, however, that our results hold in any higher dimensional Hilbert

---

[3]We introduce in this model a superposition vector with equal weights on the two vectors. The general case of a weighted superposition can be considered in future investigation, and it is an interesting line of research in itself, as the interpretation of the weights is not trivial.

space. Let $\{|1\rangle = (1, \ldots, 0), |2\rangle = (0, 1, \ldots, 0), \ldots, |8\rangle = (0, 0, \ldots, 1)\}$ denote the canonical base of $\mathbb{C}^8$. We construct a representation in $\mathbb{C}^8$ where M projects on the subspace $\mathbb{C}^4$ generated by the last four vectors of this canonical base, and $\mathbb{1} - M$ on the subspace $\mathbb{C}^4$ generator by the first four vectors of it. If we set

$$|A\rangle = e^{i\phi_A}(a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8) \tag{48}$$

$$|A'\rangle = e^{i\phi_{A'}}(a'_1, a'_2, a'_3, a'_4, a'_5, a'_6, a'_7, a'_8) \tag{49}$$

$$|B\rangle = e^{i\phi_B}(b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8) \tag{50}$$

$$|B'\rangle = e^{i\phi_{B'}}(b'_1, b'_2, b'_3, b'_4, b'_5, b'_6, b'_7, b'_8) \tag{51}$$

then Equations (46) and (47) become

$$\mu(A) = \langle A|M|A\rangle = a_5^2 + a_6^2 + a_7^2 + a_8^2 \tag{52}$$

$$1 - \mu(A) = \langle A|\mathbb{1} - M|A\rangle = a_1^2 + a_2^2 + a_3^2 + a_4^2 \tag{53}$$

$$\mu(A') = \langle A'|M|A'\rangle = a'^2_5 + a'^2_6 + a'^2_7 + a'^2_8 \tag{54}$$

$$1 - \mu(A') = \langle A'|\mathbb{1} - M|A'\rangle = a'^2_1 + a'^2_2 + a'^2_3 + a'^2_4 \tag{55}$$

$$\mu(B) = \langle B|M|B\rangle = b_5^2 + b_6^2 + b_7^2 + b_8^2 \tag{56}$$

$$1 - \mu(B) = \langle B|\mathbb{1} - M|B\rangle = b_1^2 + b_2^2 + b_3^2 + b_4^2 \tag{57}$$

$$\mu(B') = \langle B'|M|B'\rangle = b'^2_5 + b'^2_6 + b'^2_7 + b'^2_8 \tag{58}$$

$$1 - \mu(B') = \langle B'|\mathbb{1} - M|B'\rangle = b'^2_1 + b'^2_2 + b'^2_3 + b'^2_4 \tag{59}$$

and the orthogonality conditions become

$$0 = \langle A|A'\rangle = a_1 a'_1 + a_2 a'_2 + a_3 a'_3 + a_4 a'_4 + a_5 a'_5 \\ + a_6 a'_6 + a_7 a'_7 + a_8 a'_8 \tag{60}$$

$$0 = \langle B|B'\rangle = b_1 b'_1 + b_2 b'_2 + b_3 b'_3 + b_4 b'_4 + b_5 b'_5 \\ + b_6 b'_6 + b_7 b'_7 + b_8 b'_8 \tag{61}$$

$$0 = \langle A|B\rangle = a_1 b_1 + a_2 b_2 + a_3 b_3 + a_4 b_4 + a_5 b_5 \\ + a_6 b_6 + a_7 b_7 + a_8 b_8 \tag{62}$$

$$0 = \langle A|B'\rangle = a_1 b'_1 + a_2 b'_2 + a_3 b'_3 + a_4 b'_4 + a_5 b'_5 \\ + a_6 b'_6 + a_7 b'_7 + a_8 b'_8 \tag{63}$$

$$0 = \langle A'|B\rangle = a'_1 b_1 + a'_2 b_2 + a'_3 b_3 + a'_4 b_4 + a'_5 b_5 \\ + a'_6 b_6 + a'_7 b_7 + a'_8 b_8 \tag{64}$$

$$0 = \langle A'|B'\rangle = a'_1 b'_1 + a'_2 b'_2 + a'_3 b'_3 + a'_4 b'_4 + a'_5 b'_5 \\ + a'_6 b'_6 + a'_7 b'_7 + a'_8 b'_8 \tag{65}$$

A solution of Equations (52–65) gives us a configuration of the four orthonormal vectors $|A\rangle$, $|A'\rangle$, $|B\rangle$, and $|B'\rangle$ in $\mathbb{C}^8$, such that self-adjoint operator formed by the spectral decomposition of the two orthogonal projections M and $\mathbb{1} - M$ give rise to the values $\mu(A)$, $1 - \mu(A)$, $\mu(A')$, $1 - \mu(A')$, $\mu(B)$, $1 - \mu(B)$, and $\mu(B')$, $1 - \mu(B')$, corresponding to the measured data.

By using standard rules for quantum probabilities we have that the membership weights for the conjunctions corresponding to the measured data should satisfy the following equations:

$$\mu(A \text{ and } B) = \frac{1}{\sqrt{2}}(\langle A| + \langle B|)M\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$$

$$= \frac{1}{2}(\mu(A) + \mu(B)) + \Re\langle A|M|B\rangle \tag{66}$$

$$\mu(A \text{ and } B') = \frac{1}{\sqrt{2}}(\langle A| + \langle B'|)M\frac{1}{\sqrt{2}}(|A\rangle + |B'\rangle)$$

$$= \frac{1}{2}(\mu(A) + \mu(B')) + \Re\langle A|M|B'\rangle \tag{67}$$

$$\mu(A' \text{ and } B) = \frac{1}{\sqrt{2}}(\langle A'| + \langle B|)M\frac{1}{\sqrt{2}}(|A'\rangle + |B\rangle)$$

$$= \frac{1}{2}(\mu(A') + \mu(B)) + \Re\langle A'|M|B\rangle \tag{68}$$

$$\mu(A' \text{ and } B') = \frac{1}{\sqrt{2}}(\langle A'| + \langle B'|)M\frac{1}{\sqrt{2}}(|A'\rangle + |B'\rangle)$$

$$= \frac{1}{2}(\mu(A') + \mu(B')) + \Re\langle A'|M|B'\rangle \tag{69}$$

Hence, in $\mathbb{C}^8$ these equations become

$$\mu(A \text{ and } B) = \frac{1}{2}(\mu(A) + \mu(B)) + \Re\langle A|M|B\rangle$$

$$= \frac{1}{2}(\mu(A) + \mu(B)) + (a_5 b_5 + a_6 b_6 + a_7 b_7 + a_8 b_8) \\ \cos(\phi_B - \phi_A) \tag{70}$$

$$\mu(A \text{ and } B') = \frac{1}{2}(\mu(A) + \mu(B')) + \Re\langle A|M|B'\rangle$$

$$= \frac{1}{2}(\mu(A) + \mu(B')) + (a_5 b'_5 + a_6 b'_6 + a_7 b'_7 + a_8 b'_8) \\ \cos(\phi_{B'} - \phi_A) \tag{71}$$

$$\mu(A' \text{ and } B) = \frac{1}{2}(\mu(A') + \mu(B)) + \Re\langle A'|M|B\rangle$$

$$= \frac{1}{2}(\mu(A') + \mu(B)) + (a'_5 b_5 + a'_6 b_6 + a'_7 b_7 + a'_8 b_8) \\ \cos(\phi_B - \phi_{A'}) \tag{72}$$

$$\mu(A' \text{ and } B') = \frac{1}{2}(\mu(A') + \mu(B')) + \Re\langle A'|M|B'\rangle$$

$$= \frac{1}{2}(\mu(A') + \mu(B')) + (a'_5 b'_5 + a'_6 b'_6 + a'_7 b'_7 + a'_8 b'_8) \\ \cos(\phi_{B'} - \phi_{A'}) \tag{73}$$

The conditions that should be satisfied by experimental data in order to represent them in $\mathbb{C}^8$ are reported in Appendix A5. By analogy with what we found in Aerts (2009) and Sozzo (2014, 2015), we however expect that our experimental data in Appendix A3 cannot be generally modeled in the complex Hilbert space $\mathbb{C}^8$, or first sector of Fock space, but a second sector $\mathbb{C}^8 \otimes \mathbb{C}^8$ of Fock space is also needed. Consider for example a simple case that applies for classical logics $\mu(A) = 1$, $\mu(B) = 0$ and $\mu(A \text{ and } B) = 0$. This case is consistent with the minimum conjunction rule (Zadeh, 1982) but not with our first sector of Fock space (Hilbert space) model. We will see that this type of cases is compatible with second sector of Fock space, and show that a framework that encompasses both "logical" and "emergent" reasoning about membership in these situations requires the general properties of a Fock space. To make our description complete however, we first have to introduce a new, conceptually relevant, ingredient.

3.2. Introducing Entanglement in Conceptual Combinations

In Aerts (2009) and Sozzo (2014, 2015) we successfully modeled conjunctions of the form "A and B" in a Fock space constructed as the direct sum of an individual Hilbert space $\mathcal{H}$, or "first sector of Fock space," and a tensor product Hilbert space $\mathcal{H} \otimes \mathcal{H}$, or "second sector of Fock space." The concepts A and B were respectively represented by the unit vectors $|A\rangle$ and $|B\rangle$ of $\mathcal{H}$, while the conjunction "A and B" was represented by the unit vector $\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$ in first sector, and by the tensor product vector $|A\rangle \otimes |B\rangle$ in second sector. The decision measurement of a person who estimates whether a given exemplar x is a member of "A and B" was represented by the orthogonal projection operator M in first sector, and by the tensor product projection operator $M \otimes M$ in second sector. The conjunction "A and B" was represented by a unit vector of the form $\psi(A, B) = me^{i\theta}|A\rangle \otimes |B\rangle + ne^{i\rho}\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$ in the Fock space $\mathcal{H} \oplus (\mathcal{H} \otimes \mathcal{H})$, while the decision measurement was represented by the orthogonal projection operator $M \oplus (M \otimes M)$ in the same Fock space. By using quantum probabilistic rules, one could then write the membership weight of x with respect to "A and B" as $\mu(A \text{ and } B) = \langle\psi(A, B)|(M \otimes M) \oplus M|\psi(A, B)\rangle = m^2\mu(A)\mu(B) + n^2(\frac{1}{2}(\mu(A) + \mu(B)) + \Re\langle A|M|B\rangle)$. This treatment needs now to be generalized to the decision measurement of the concepts A, B, the negations A', B' and the conjunctions "A and B," "A and B'," "A' and B," and "A' and B'." The first sector situation has already been analysed in Section 3.1 where we have also constructed an explicit representation in the complex Hilbert space $\mathbb{C}^8$. Here we analyse the second sector situation, but we allow for the possibility of representing concepts by entangled states too.

Is it possible to introduce some "type of entanglement" in second sector $\mathbb{C}^8 \otimes \mathbb{C}^8$? This question is interesting, since it is reasonable to believe that the outcomes of experiments for A are not independent of the outcomes of experiments for B. For example, in case a specific exemplar x is strongly a member of *Fruits*, this will influence the strength of membership of *Vegetables*, and viceversa, because the meanings of *Fruits* and *Vegetables* are not independent. And this apparently occurs for all human concepts. Suppose we combined, for example, *Fruits* with *Not Fruits*, then one would expect to exist, for any exemplar, a substantial amount of anti-correlation between it being a member of *Fruits* and it being a member of *Not Fruits*. How can we express the general situation, where anti-correlation, as well as correlation, are possible to be increased or decreased by parameters? In the foregoing modeling (Aerts, 2009; Aerts et al., 2013b) we chose the simplest representation for the situation in second sector, namely the product state $|A\rangle \otimes |B\rangle$, which leads to a situation of complete independence between A and B, for whatever exemplar tested. Let us investigate what would be a situation for a general entangled state, and how a two-sector Fock space already incorporates this possibility.

Suppose that the concept "A and B" is not represented by the product state vector $|A\rangle \otimes |B\rangle$ in second sector of Fock space $\mathbb{C}^8 \otimes \mathbb{C}^8$, but by a general entangled state vector $|C\rangle$ of $\mathbb{C}^8 \otimes \mathbb{C}^8$. We remind that $\mathbb{C}^8$ is the concrete Hilbert space we have constructed in Section 3.1. In the canonical base $\{|i\rangle\}_{i=1,\ldots,8}$ of $\mathbb{C}^8$, we have

$$|C\rangle = \sum_{i,j=1}^{8} c_{ij}e^{i\gamma_{ij}}|i\rangle \otimes |j\rangle \qquad (74)$$

and

$$1 = \langle C|C\rangle = \sum_{i,j=1}^{8} c_{ij}^2 \qquad (75)$$

We now express the effect, as described in second sector, of the experiments where participants were asked to decide for membership (or non-membership) of a specific exemplar with respect to the concepts A and B, as follows. Membership with respect to A, as a yes-no measurement, is represented by the orthogonal projection operators $M \otimes \mathbb{1}$, $(\mathbb{1} - M) \otimes \mathbb{1}$, as spectral family of the corresponding self-adjoint operator. Hence in second sector, tests on concept A, are described in the first component of the tensor product Hilbert space $\mathbb{C}^8 \otimes \mathbb{C}^8$, which forms second sector. In an analogous way, tests on concept B, are described in the second component of the tensor product, by the orthogonal projection operators $\mathbb{1} \otimes M$, $\mathbb{1} \otimes (\mathbb{1} - M)$, as spectral family of the corresponding self-adjoint operator. Remark that we do not introduce in any way the concepts A' and B' for the second sector description, and also not conjunction of A and B with them, at least the aspect of these conjunctions that represent new emergent concepts. The concept A' and B' are indeed "emergent entities" because the negation on a concept to give rise to a new concept, namely the negation concept.

Also the experimental data collected on A', B' and combinations of them with A and B do not appear in second sector. All emergence is indeed modeled in first sector. This means that also the conjunction of A and B as a new emergent concept does not appear in second sector, it only appears in first sector modeled there by the superposition. All non-emergent equivalents of these are described by the tensor product of the two self-adjoint operators corresponding to the yes-no experiments with respect to membership performed on concepts A and consequently on concept B, exactly as in our real life experiment that gave rise to our data on A and B. This means that the orthogonal projectors of the spectral family of this tensor product self-adjoint operator describe all cases of non-emergence. This family consists of $\{M \otimes M, M \otimes (\mathbb{1} - M), (\mathbb{1} - M) \otimes M, (\mathbb{1} - M) \otimes (\mathbb{1} - M)\}$. Let us express this on a general entangled state vector $|C\rangle$. We have

$$\mu(A) = \langle C|M \otimes \mathbb{1}|C\rangle = \sum_{i=5}^{8}\sum_{j=1}^{8} c_{ij}^2 \qquad (76)$$

$$\mu(B) = \langle C|\mathbb{1} \otimes M|C\rangle = \sum_{i=1}^{8}\sum_{j=5}^{8} c_{ij}^2 \qquad (77)$$

$$1 - \mu(A) = \langle C|(\mathbb{1} - M) \otimes \mathbb{1}|C\rangle = \sum_{i=1}^{4}\sum_{j=1}^{8} c_{ij}^2 \qquad (78)$$

$$1 - \mu(B) = \langle C|\mathbb{1} \otimes (\mathbb{1} - M)|C\rangle = \sum_{i=1}^{8}\sum_{j=1}^{4} c_{ij}^2 \qquad (79)$$

And further we have

$$\langle C|M \otimes M|C\rangle = \sum_{i=5}^{8}\sum_{j=5}^{8} c_{ij}^2 \qquad (80)$$

$$\langle C|M \otimes (\mathbb{1} - M)|C\rangle = \sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2 \qquad (81)$$

$$\langle C|(\mathbb{1} - M) \otimes M|C\rangle = \sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2 \qquad (82)$$

$$\langle C|(\mathbb{1} - M) \otimes (\mathbb{1} - M)|C\rangle = \sum_{i=1}^{4}\sum_{j=1}^{4} c_{ij}^2 \qquad (83)$$

The values of $\langle C|M \otimes M|C\rangle$, $\langle C|M \otimes (\mathbb{1} - M)|C\rangle$, $\langle C|(\mathbb{1} - M) \otimes M|C\rangle$, and $\langle C|(\mathbb{1} - M) \otimes (\mathbb{1} - M)|C\rangle$ will respectively represent the amounts that within our Fock space model second sector contributes to the values of $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$. We can prove that the second sector theoretical values, allowing the state to be a general entangled state in our $\mathbb{C}^8 \otimes \mathbb{C}^8$ Hilbert space model, reach exactly the values to be found for the case the classicality conditions Equations (26–30) in Theorem $3'$ are satisfied. More explicitly, the following theorem holds (see Appendix A4 for its proof).

Theorem 4. *If the experimentally collected membership weights $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ can be represented in second sector of Fock space for a given choice of the entangled state vector $|C\rangle$ and the decision measurement projection operator M, then the membership weights satisfy Equations (26–30), hence they are classical data. Viceversa, if $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ satisfy Equations (26–30), hence they are classical data, then an entangled state vector $|C\rangle$ and a decision measurement projection operator M can always be found such that $\mu(A)$, $\mu(B)$, $\mu(A')$, $\mu(B')$, $\mu(A \text{ and } B)$, $\mu(A \text{ and } B')$, $\mu(A' \text{ and } B)$, and $\mu(A' \text{ and } B')$ can be represented in second sector of Fock space.*

Theorem 4 implies that the tensor product Hilbert space model (second sector of Fock space) has exactly the same generality as the most general classical conditions for conjunction and negation. More specifically, given any data that satisfy the five classicality conditions of Theorem $3'$, we can construct an entangled state such that in second sector "exactly" these classicality conditions are satisfied. Moreover, it clarifies that entangled states in our general Fock space modeling of data on conceptual conjunction and negation play a fundamental unexpected role in the combination of human concepts. The fact that classical logical rules are satisfied, in a probabilistic form, in second sector of Fock space provides an important confirmation to our two-sector quantum framework, as we will see in Section 5.

3.3. A complete modeling in Fock space

In Section 3.1 we have considered the situation of first sector of Fock space, representing the starting concepts A and B by the state vectors $|A\rangle$ and $|B\rangle$, respectively, of a Hilbert space $\mathcal{H}$. Then, we have introduced the state vectors $|A'\rangle$ and $|B'\rangle$, which represent the conceptual negations "not A" and "not B," respectively. Since first sector of Fock space describes "emergence," the state vectors $|A'\rangle$ and $|B'\rangle$ can be interpreted as representing the newly emergent concepts "not A" and "not B," respectively, compatibly with the core of the approach we developed in our quantum modeling of combinations of concepts. We have also seen in Section 3.1 that

also the newly emergent concept "A and B," or one of the other conjunction combinations, "A and B'," "A' and B," and "A' and B'," are directly represented in this first sector of Fock space by state vectors, more specifically by the superposition state vectors of the corresponding state vectors, namely $\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$, $\frac{1}{\sqrt{2}}(|A\rangle + |B'\rangle)$, $\frac{1}{\sqrt{2}}(|A'\rangle + |B\rangle)$ and $\frac{1}{\sqrt{2}}(|A'\rangle + |B'\rangle)$), respectively.

Following Aerts (2009) and Sozzo (2014, 2015), we should also take into account the logical aspects of conceptual conjunctions and negations in second sector of Fock space, mathematically formed by the tensor product of the Hilbert space $\mathcal{H}$ of first sector. In previous papers we had represented the state of the concept A and B in second sector by the product vector $|A\rangle \otimes |B\rangle$ of this tensor product $\mathcal{H} \otimes \mathcal{H}$. However, this leads inevitably to the probability for the conjunction $\mu(A \text{ and } B)$ to be equal to the product $\mu(A)\mu(B)$, as we have seen in Section 3.2. In classical probability theory this means that the probabilities are probabilistically independent. Now, quite obviously, since the concepts A and B are related by their meaning, these probabilities are not probabilistically independent. Suppose that B is the negation A', like in the borderline effect (Sozzo, 2014). Then, we obviously would have an anti-correlation between $\mu(A)$ and $\mu(B)$. But, even in this not simple case, any meaning connection between A and B would give rise to probabilities that are not independent. On the other hand, we have seen in Section 3.2 that we can model any type of classical probabilistic dependence by introducing the proper entangled state for the concept representation of A and of B in second sector. This means that we should not in principle use $|A\rangle \otimes |B\rangle$ to represent the concepts in second sector, but a properly chosen entangled state.

Let us denote, following our analysis in Section 3.2, such a general entangled state in $\mathbb{C}^8$ by means of

$$|C\rangle = \sum_{i,j=1}^{8} c_{ij}e^{\gamma_{ij}}|i\rangle \otimes |j\rangle \qquad (84)$$

where $|i\rangle$ and $|j\rangle$ are the canonical base vectors of $\mathbb{C}^8$.

The state vector representing the concept "A and B" in its totality, hence its first sector part, describing emergent human thought, i.e., the formation of the new concept "A and B," and its second sector part, describing quantum logical human thought, i.e., the conjunctive connective structure "A and B," is then the following

$$\psi(A, B) = m_{AB}e^{i\theta}|C\rangle + \frac{n_{AB}e^{i\rho}}{\sqrt{2}}(|A\rangle + |B\rangle) \qquad (85)$$

with $m_{AB}^2 + n_{AB}^2 = 1$. It is indeed the superposition of two vectors, one vector given by $\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$ in first sector of Fock space, accounting for the emergent part of human thought with respect to the conjunction, and a second vector given by $|C\rangle$ in second sector of Fock space, accounting for the quantum logical part of human thought with respect to the conjunction. By using Equations (84) and (85), we then get the following general expression for the membership weight of the conjunction

$$\mu(A \text{ and } B) = \langle \psi(A, B)|(M \otimes M \oplus M)|\psi(A, B)\rangle$$
$$= m_{AB}^2(\langle C|)M \otimes M|(|C\rangle)$$

$$+ \frac{n_{AB}^2}{2}(\langle A| + \langle B|)M(|A\rangle + |B\rangle)$$
$$= m_{AB}^2 \sum_{i,j=5}^{8} c_{ij}^2 + \frac{n_{AB}^2}{2}(\langle A|M|A\rangle + \langle B|M|B\rangle$$
$$+ \langle A|M|B\rangle + \langle B|M|A\rangle)$$
$$= m_{AB}^2 \sum_{i,j=5}^{8} c_{ij}^2 + n_{AB}^2(\frac{1}{2}(\mu(A) + \mu(B)) + \Re\langle A|M|B\rangle)$$
$$= m_{AB}^2 \sum_{i,j=5}^{8} c_{ij}^2 + n_{AB}^2(\frac{1}{2}(\mu(A) + \mu(B))$$
$$+ (a_5 b_5 + a_6 b_6 + a_7 b_7 + a_8 b_8)\cos(\phi_B - \phi_A)) \quad (86)$$

where we have used Equation (70) in the last line of Equation (86).

What is the procedure corresponding to emergent and quantum logical parts of human thought when we also take into account negations, i.e., when we consider the conjunctions "A and B'," "A' and B," and "A' and B'"? As for first sector of Fock space, we have already made it explicit in Section 3.1. Indeed, the new emergent concepts "A'" and "B'" are described by the state vectors $|A'\rangle$ and $|B'\rangle$, and the respective conjunctions, i.e., their emergent aspects as a new concept, each time by means of the corresponding superposition state vector. This is the way the emergence of negation and conjunction are jointly modeled in first sector of Fock space—new state vectors model the new emergent concepts due to negation, and the emergent conjunctions are modeled by the respective superpositions.

In second sector of Fock space, we however have a specific situation to solve. Namely, exactly as we did for the conjunction, we need to identify what is the quantum logical structure related with negation, independent of its provoking the emergence of a new concept, i.e., the negation of the original concept. In second sector of Fock space we indeed only express the quantum logical reasoning in human thought and not the emergent reasoning. For the conjunction "A and B" we did this by means of the entangled state $|C\rangle$. Let us reflect about the negation, for example, with respect to the concept B. To make things clear let us introduce the following two expressions. We are in the experimental situation where the membership of an exemplar x, or the non-membership of this exemplar, is to be decided about, by a person participating in the experiment. The concept B can be involved, and the concept B' can be involved.

Expression 1. The considered exemplar x is a member of the concept B'.

Expression 2. The considered exemplar x is "not" a member of the concept B.

Our theoretic proposal is that:

(1) the first expression describes what happens in a human mind when emergent thought is dominant with respect to a concept B, its negation B' and an exemplar x. Indeed, the focus is on "membership" of this exemplar x with respect to the new emergent concept "B'";
(2) the second expression describes what happens in a human mind when quantum logical thought is dominant with respect to a concept B, its negation B' and an exemplar x. Indeed, the focus is on "non-membership" of this exemplar x with respect to the old existing concept B.

Expressions (1) and (2) are two structurally speaking subtle deeply different possibilities of reasoning related to a concept and its negation.

Our third theoretic proposal is that:

(3) human thought, when confronted with this situation, follows a dynamics described by a quantum superposition of the two modes (1) and (2).

We will see in the following that the mathematical structure of Fock space enables modeling this in an impecable way.

Indeed, expression (1) will be modeled in first sector of our Fock space, and it is mathematically realised by making M work on $|B'\rangle$. Expression (2) will instead be modeled in second sector of Fock space, and it is mathematically realised by making $\mathbb{1} \otimes (\mathbb{1} - M)$ work on $|C\rangle$. In the complete Fock space, direct sum of its first and second sectors, mathematically a superposition of the whole dynamics can be realised, by considering the superposition state which we already specified in Equation (85), and consider different structures of the projection operator on the whole of Fock space. More specifically, $(M \otimes M) \oplus M$ for "A and B," $(M \otimes (\mathbb{1} - M)) \oplus M$ for "A and B'," $((\mathbb{1} - M) \otimes M) \oplus M$ for "A' and B," and $((\mathbb{1} - M) \otimes (\mathbb{1} - M)) \oplus M$ for "A' and B'." However, for each of the combinations the vector representing the combination in second sector of Fock space will be $|C\rangle$. So, no vector appears in second sector of Fock space, since the negation is expressed quantum logically here, hence by M becoming $\mathbb{1} - M$. While in first sector of Fock space, the negation is expressed emergently, hence by $|A\rangle$ becoming $|A'\rangle$ and $|B\rangle$ becoming $|B'\rangle$, and M remaining M, since the focus in this first sector of Fock space, with emergent reasoning of human thought, is always on "membership," while in second sector, with quantum logical reasoning, the focus of negation is on "non-membership," described by $\mathbb{1} - M$.

The above conceptual analysis makes it possible for us to write the complete Fock space formulas for the other combinations. More specifically, if we represent the concept "A and B'" by the unit vector

$$\psi(A, B') = m_{AB'} e^{i\theta}|C\rangle + \frac{n_{AB'} e^{i\rho}}{\sqrt{2}}(|A\rangle + |B'\rangle) \quad (87)$$

with $m_{AB'}^2 + n_{AB'}^2 = 1$, then, by using Equations (84) and (87), we get

$$\mu(A \text{ and } B') = \langle \psi(A, B')|M \otimes (\mathbb{1} - M) \oplus M|\psi(A, B')\rangle$$
$$= m_{AB'}^2(\langle C|)M \otimes (\mathbb{1} - M)|(|C\rangle)$$
$$+ \frac{n_{AB'}^2}{2}(\langle A| + \langle B'|)M(|A\rangle + |B'\rangle)$$
$$= m_{AB'}^2 \sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2 + \frac{n_{AB'}^2}{2}(\langle A|M|A\rangle + \langle B'|M|B'\rangle$$
$$+ \langle A|M|B'\rangle + \langle B'|M|A\rangle)$$
$$= m_{AB'}^2 \sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2 + n_{AB'}^2\left(\frac{1}{2}(\mu(A) + \mu(B'))\right.$$
$$\left. + \Re\langle A|M|B'\rangle\right)$$

$$= m_{AB'}^2 \sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2 + n_{AB'}^2(\frac{1}{2}(\mu(A)+\mu(B')))$$
$$+ (a_5 b_5' + a_6 b_6' + a_7 b_7' + a_8 b_8')\cos(\phi_{B'}-\phi_A)) \tag{88}$$

where we have used Equation (71) in the last line of Equation (88). Analogously, if we represent the concept "A' and B" by the unit vector

$$\psi(A',B) = m_{A'B}e^{i\theta}|C\rangle + \frac{n_{A'B}e^{i\rho}}{\sqrt{2}}(|A'\rangle + |B\rangle) \tag{89}$$

with $m_{A'B}^2 + n_{A'B}^2 = 1$, then, by using Equations (84) and (89), we get

$$\mu(A' \text{ and } B) = \langle\psi(A',B)|(\mathbb{1}-M)\otimes M \oplus M|\psi(A',B)\rangle$$
$$= m_{A'B}^2(\langle C|)(\mathbb{1}-M)\otimes M|(|C\rangle)$$
$$+ \frac{n_{A'B}^2}{2}(\langle A'|+\langle B|)M(|A'\rangle+|B\rangle)$$
$$= m_{A'B}^2 \sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2 + \frac{n_{A'B}^2}{2}(\langle A'|M|A'\rangle + \langle B|M|B\rangle$$
$$+ \langle A'|M|B\rangle + \langle B|M|A'\rangle)$$
$$= m_{A'B}^2 \sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2 + n_{A'B}^2(\frac{1}{2}(\mu(A')+\mu(B))$$
$$+ \Re\langle A'|M|B\rangle)$$
$$= m_{A'B}^2 \sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2 + n_{A'B}^2(\frac{1}{2}(\mu(A')+\mu(B))$$
$$+ (a_5' b_5 + a_6' b_6 + a_7' b_7 + a_8' b_8)\cos(\phi_B-\phi_{A'})) \tag{90}$$

where we have used Equation (72) in the last line of Equation (90). Finally, if we represent the concept "A' and B'" by the unit vector

$$\psi(A',B') = m_{A'B'}e^{i\theta}|C\rangle + \frac{n_{A'B'}e^{i\rho}}{\sqrt{2}}(|A'\rangle + |B'\rangle) \tag{91}$$

with $m_{A'B'}^2 + n_{A'B'}^2 = 1$, then, by using Equations (84) and (91), we get

$$\mu(A' \text{ and } B') = \langle\psi(A',B')|(\mathbb{1}-M)\otimes(\mathbb{1}-M) \oplus M|\psi(A',B')\rangle$$
$$= m_{A'B'}^2(\langle C|)(\mathbb{1}-M)\otimes(\mathbb{1}-M)|(|C\rangle)$$
$$+ \frac{n_{A'B'}^2}{2}(\langle A'|+\langle B'|)M(|A'\rangle+|B'\rangle)$$
$$= m_{A'B'}^2 \sum_{i,j=1}^{4} c_{ij}^2 + \frac{n_{A'B'}^2}{2}(\langle A'|M|A'\rangle$$
$$+ \langle B'|M|B'\rangle + \langle A'|M|B'\rangle + \langle B'|M|A'\rangle)$$
$$= m_{A'B'}^2 \sum_{i,j=1}^{4} c_{ij}^2 + n_{A'B'}^2(\frac{1}{2}(\mu(A')+\mu(B'))$$
$$+ \Re\langle A'|M|B'\rangle)$$

$$= m_{A'B'}^2 \sum_{i,j=1}^{4} c_{ij}^2 + n_{A'B'}^2(\frac{1}{2}(\mu(A')+\mu(B'))$$
$$+ (a_5' b_5' + a_6' b_6' + a_7' b_7' + a_8' b_8')\cos(\phi_{B'}-\phi_{A'})) \tag{92}$$

where we have used Equation (73) in the last line of Equation (92).

Equations (86), (88), (90), and (92) contain the probabilistic expressions for simultaneously representing experimental data on conjunctions and negations of two concepts in a quantum-theoretic framework. These equations express the membership weights of the conjunctions "A and B," "A and B'," "A' and B," and "A' and B'" in terms of the memership weights of A, B, A', and B', for suitable values of the following modeling parameters:[4]

(i)   the angles $\phi_B - \phi_A$, $\phi_{B'} - \phi_A$, $\phi_B - \phi_{A'}$ and $\phi_{B'} - \phi_{A'}$,
(ii) the pairs of convex coefficients (m_{AB}, n_{AB}), $(m_{AB'}, n_{AB'})$, $(m_{A'B}, n_{A'B})$ and $(m_{A'B'}, n_{A'B'})$,
(iii) the normalized coefficients $c_{11}^2, \ldots, c_{88}^2$.

As we can see our two-sector Fock space framework is able to cope with conceptual negation in a very natural way. In fact, the latter negation is modeled by using the general assumption that emergent aspects of a concept are represented in first sector of Fock space, while logical aspects of a concept are represented in second sector. This will be made explicit in Section 5. It is however important to stress that, for a given experiment e_{XY}, with $X = A, A'$, $Y = B, B'$ described in Section 2, there is no guarantee that sets of these parameters can be found such that Equations (86–92) are simultaneously satisfied. For this reason, we provide in Appendix A5 the conditions that should be satisfied by the experimental data $\mu(A), \mu(B), \ldots, \mu(A' \text{ and } B), \mu(A' \text{ and } B)$ such that these sets exist.

The conclusion we draw from the analysis above is that finding solutions for a given set of experimental data in our quantum-theoretic modeling it is highly non-obvious, which makes the results in the next section even more significant.

4. Representation of experimental data in Fock space

Most of these data in **Tables A1–A4**, are compatible with the intervals in Equations (A26), (A29), (A32), and (A35). Hence, almost all our data can be successfully modeled by using the quantum probabilistic equations in Equations (86), (88), (90), and (92). Let us consider some interesting cases, distinguishing them by: (i) situations with double overextension, (ii) situations with complete overextension, (iii) situations requiring both sectors of Fock space and/or entanglement, (iv) partially classical situations. Complete modeling is presented in the Supplementary Material attached to this article.

(i)   Let us start with exemplars that are double overextended.
      Olive, with respect to (*Fruits*, *Vegetables*) (double overextension with respect to *Fruits And Vegetables*). *Olive* scored $\mu(A) = 0.53$ with respect to *Fruits*, $\mu(B) = 0.63$

---

[4]We remind that $n_{XY}^2 = 1 - m_{XY}^2$, $X = A, A'$, $Y = B, B'$. In addition, only the sums $\sum_{i,j=5}^{8} c_{ij}^2$, $\sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2$, $\sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2$, and $\sum_{i,j=5}^{8} c_{ij}^2$ appear in Equations (86), (88), (90), and (92), respectively.

with respect to *Vegetables*, $\mu(A') = 0.47$ with respect to *Not Fruits*, $\mu(B') = 0.44$ with respect to *Not Vegetables*, $\mu(A \text{ and } B) = 0.65$ with respect to *Fruits And Vegetables*, $\mu(A \text{ and } B') = 0.34$ with respect to *Fruits And Not Vegetables*, $\mu(A' \text{ and } B) = 0.51$ with respect to *Not Fruits And Vegetables*, and $\mu(A' \text{ and } B') = 0.36$ with respect to *Not Fruits And Not Vegetables*. If one first looks for a representation of *Olive* in the Hilbert space $\mathbb{C}^8$, then the concepts *Fruits* and *Vegetables* are represented by the unit vectors $|A\rangle = e^{i\phi_A}$ $(-0.02, -0.47, 0.5, -0.02, -0.07, -0.31, -0.18, -0.63)$ and $|B\rangle = e^{i\phi_B}$ $(0.04, 0.02, -0.6, 0.03, -0.26, 0.35, -0.39, -0.53)$, respectively, and their negations *Not Fruits* and *Not Vegetables* by the unit vectors $|A'\rangle = e^{i\phi_{A'}}$ $(0.06, -0.47, -0.55, 0.03, -0.02, -0.64, -0.06, 0.25)$, and $|B'\rangle = e^{i\phi_{B'}}$ $(-0.03, 0.75, -0.01, -0.01, -0.08, -0.6, -0.18, -0.19)$, respectively.

The interference angles $\phi_{AB} = \phi_B - \phi_A = 57.31°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 95.32°$, $\phi_{A'B} = \phi_B - \phi_{A'} = 103.43°$ and $\phi_{A'B'} = \phi_{B'} - \phi_{A'} = 85.56°$ complete the Hilbert space representation in $\mathbb{C}^8$. A complete modeling in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$ satisfying Equations (86), (88), (90), and (92) is given by an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.44^2$, $\sum_{i=5}^{8} \sum_{i=1}^{4} c_{ij}^2 = 0.58^2$, $\sum_{i=1}^{4} \sum_{j=5}^{8} c_{ij}^2 = 0.66^2$ and $\sum_{i,j=1}^{4} c_{ij}^2 = 0.18^2$, and convex weights $m_{AB} = 0.42$, $n_{AB} = 0.91$, $m_{AB'} = 0.1$, $n_{AB'} = 0$, $m_{A'B} = 0.78$, $n_{A'B} = 0.63$, $m_{A'B'} = 0.52$, and $n_{A'B'} = 0.86$.

Prize Bull, with respect to (*Pets*, *Farmyard Animals*) (double overextension with respect to *Pets And Not Farmyard Animals*). *Prize Bull* scored $\mu(A) = 0.13$ with respect to *Pets*, $\mu(B) = 0.76$ with respect to *Farmyard Animals*, $\mu(A') = 0.88$ with respect to *Not Pets*, $\mu(B') = 0.26$ with respect to *Not Farmyard Animals*, $\mu(A \text{ and } B) = 0.43$ with respect to *Pets And Farmyard Animals*, $\mu(A \text{ and } B') = 0.28$ with respect to *Pets And Not Farmyard Animals*, $\mu(A' \text{ and } B) = 0.83$ with respect to *Not Pets And Farmyard Animals*, and $\mu(A' \text{ and } B') = 0.34$ with respect to *Not Pets And Not Farmyard Animals*. If one first looks for a representation of *Prize Bull* in the Hilbert space $\mathbb{C}^8$, then the concepts *Pets* and *Farmyard Animals*, and their negations *Not Pets* and *Not Farmyard Animals* are respectively represented by the unit vectors $|A\rangle = e^{i\phi_A}$ $(0.07, -0.39, -0.84, 0.03, -0.06, -0.35, 0.04, -0.01)$ and $|B\rangle = e^{i\phi_B}$ $(0.03, 0.21, -0.44, 0.01, 0.01, 0.81, -0.2, -0.25)$, and $|A'\rangle = e^{i\phi_{A'}}$ $(0.01, 0.29, -0.19, 0, 0.11, 0.06, -0.2, 0.91)$ and $|B'\rangle = e^{i\phi_{B'}}$ $(0.01, 0.84, -0.19, 0, -0.17, -0.41, -0.01, -0.26)$.

The interference angles $\phi_{AB} = \phi_B - \phi_A = 105.71°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 40.23°$, $\phi_{A'B}E = \phi_B - \phi_{A'} = 111.25°$ and $\phi_{A'B'} = \phi_{B'} - \phi_{A'} = 52.51°$ complete the Hilbert space representation in $\mathbb{C}^8$. A complete modeling in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$ satisfying Equations (86), (88), (90), and (92) is given by an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.24^2$, $\sum_{i=5}^{8} \sum_{i=1}^{4} c_{ij}^2 = 0.27^2$, $\sum_{i=1}^{4} \sum_{j=5}^{8} c_{ij}^2 = 0.84^2$, and $\sum_{i,j=1}^{4} c_{ij}^2 = 0.41^2$, and convex weights $m_{AB} = 0.46$, $n_{AB} = 0.89$, $m_{AB'} = 0.41$, $n_{AB'} = 0.91$, $m_{A'B} = 0.54$, $n_{A'B} = 0.84$, $m_{A'B'} = 0.52$, and $n_{A'B'} = 0.85$.

Door Bell, with respect to (*Home Furnishing*, *Furniture*) (double overextension with respect to *Not Home Furnishing And Furniture*). *Door Bell* scored $\mu(A) = 0.75$ with respect to *Home Furnishing*, $\mu(B) = 0.33$ with respect to *Furniture*, $\mu(A') = 0.32$ with respect to *Not Home Furnishing*, $\mu(B') = 0.79$ with respect to *Not Furniture*, $\mu(A \text{ and } B) = 0.5$ with respect to *Home Furnnishing And Furniture*, $\mu(A \text{ and } B') = 0.64$ with respect to *Home Furnishing And Not Furniture*, $\mu(A' \text{ and } B) = 0.34$ with respect to *Not Home Furnishing And Furniture*, and $\mu(A' \text{ and } B') = 0.51$ with respect to *Not Home Furnishing And Not Furniture*. If one first looks for a representation of *Door Bell* in the Hilbert space $\mathbb{C}^8$, then the concepts *Home Furnishing* and *Furniture*, and their negations *Not Home Furnishing* and *Not Furniture* are respectively represented by the unit vectors $|A\rangle = e^{i\phi_A}$ $(0, 0.33, 0.37, -0.05, 0.04, -0.29, 0, 0.81)$ and $|B\rangle = e^{i\phi_B}$ $(-0.14, 0.77, 0.17, -0.16, 0.24, -0.19, 0.07, -0.48)$, and $|A'\rangle = e^{i\phi_{A'}}$ $(0.21, -0.43, 0.66, 0.13, 0.22, -0.39, 0.22, -0.27)$ and $|B'\rangle = e^{i\phi_{B'}}$ $(-0.08, -0.03, -0.45, -0.02, -0.17, -0.52, 0.7, 0.04)$.

The interference angles $\phi_{AB} = \phi_B - \phi_A = 102.81°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 117.67°$, $\phi_{A'B} = \phi_B - \phi_{A'} = 67.37°$ and $\phi_{A'B'} = \phi_{B'} - \phi_{A'} = 77.65°$ complete the Hilbert space representation in $\mathbb{C}^8$. A complete modeling in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$ satisfying Equations (86), (88), (90), and (92) is given by an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.35^2$, $\sum_{i=5}^{8} \sum_{i=1}^{4} c_{ij}^2 = 0.79^2$, $\sum_{i=1}^{4} \sum_{j=5}^{8} c_{ij}^2 = 0.46^2$ and $\sum_{i,j=1}^{4} c_{ij}^2 = 0.2^2$, and convex weights $m_{AB} = 0.48$, $n_{AB} = 0.88$, $m_{AB'} = 0.91$, $n_{AB'} = 0.41$, $m_{A'B} = 0.65$, $n_{A'B} = 0.76$, $m_{A'B'} = 0.43$, and $n_{A'B'} = 0.9$.

(ii) Let us now come to the exemplars that present complete overextension, that is, exemplars that are overextended in all experiments.

Goldfish, with respect to (*Pets*, *Farmyard Animals*) (big overextension in all experiments, but also double overextension with respect to *Not Pets And Farmyard Animals*). *Goldfish* scored $\mu(A) = 0.93$ with respect to *Pets*, $\mu(B) = 0.17$ with respect to *Farmyard Animals*, $\mu(A') = 0.12$ with respect to *Not Pets*, $\mu(B') = 0.81$ with respect to *Not Farmyard Animals*, $\mu(A \text{ and } B) = 0.43$ with respect to *Pets And Farmyard Animals*, $\mu(A \text{ and } B') = 0.91$ with respect to *Pets And Not Farmyard Animals*, $\mu(A' \text{ and } B) = 0.18$ with respect to *Not Pets And Farmyard Animals*, and $\mu(A' \text{ and } B') = 0.43$ with respect to *Not Pets And Not Farmyard Animals*. If one first looks for a representation of *Goldfish* in the Hilbert space $\mathbb{C}^8$, then the concepts *Pets* and *Farmyard Animals*, and their negations *Not Pets* and *Not Farmyard Animals* are respectively represented by the unit vectors $|A\rangle = e^{i\phi_A}$ $(-0.05, 0.16, -0.21, -0.01, -0.71, 0.22, 0.33, 0.51)$ and $|B\rangle = e^{i\phi_B}$ $(-0.24, 0.26, -0.84, -0.07, 0.38, -0.11, -0.01, 0.12)$, and $|A'\rangle = e^{i\phi_{A'}}$ $(0.18, 0.85, 0.35, 0.09, 0.2, -0.12, -0.03, 0.25)$ and $|B'\rangle = e^{i\phi_{B'}}$ $(0.01, -0.41, 0.14, -0.01, 0.27, -0.32, -0.13, 0.79)$.

The interference angles $\phi_{AB} = \phi_B - \phi_A = 78.9°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 43.15°$, $\phi_{A'B} = \phi_B - \phi_{A'} = 54.74°$ and $\phi_{A'B'} = \phi_{B'} - \phi_{A'} = 77.94°$ complete the Hilbert space representation in $\mathbb{C}^8$. A complete modeling in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$ satisfying Equations (86), (88), (90), and (92) is given by an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.35^2$, $\sum_{i=5}^{8} \sum_{i=1}^{4} c_{ij}^2 = 0.9^2$, $\sum_{i=1}^{4} \sum_{j=5}^{8} c_{ij}^2 = 0.22^2$ and $\sum_{i,j=1}^{4} c_{ij}^2 = 0.17^2$, and convex weights $m_{AB} = 0.45$, $n_{AB} = $

0.89, $m_{AB'} = 0.45$, $n_{AB'} = 0.9$, $m_{A'B} = 0.48$, $n_{A'B} = 0.88$, $m_{A'B'} = 0.45$, and $n_{A'B'} = 0.89$.

Parsley, with respect to (*Spices, Herbs*) (overextension in all experiments). *Parsley* scored $\mu(A) = 0.54$ with respect to *Spices*, $\mu(B) = 0.9$ with respect to *Herbs*, $\mu(A') = 0.54$ with respect to *Not Spices*, $\mu(B') = 0.09$ with respect to *Not Herbs*, $\mu(A \text{ and } B) = 0.68$ with respect to *Spices And Herbs*, $\mu(A \text{ and } B') = 0.26$ with respect to *Spices And Not Herbs*, $\mu(A' \text{ and } B) = 0.73$ with respect to *Not Spices And Herbs*, and $\mu(A' \text{ and } B') = 0.18$ with respect to *Not Spices And Not Herbs*. If one first looks for a representation of *Parsley* in the Hilbert space $\mathbb{C}^8$, then the concepts *Spices* and *Herbs*, and their negations *Not Spices* and *Not Herbs* are respectively represented by the unit vectors $|A\rangle = e^{i\phi_A}$ (0, 0.25, −0.63, −0.02, −0.02, 0.5, −0.06, 0.54) and $|B\rangle = e^{i\phi_B}$ (0, 0.02, −0.32, −0.01, 0.09, −0.84, −0.23, 0.37), and $|A'\rangle = e^{i\phi_{A'}}$ (0, 0.17, 0.66, 0.02, −0.17, 0.01, 0.14, 0.7) $|B'\rangle = e^{i\phi_{B'}}$ (0, −0.95, −0.06, −0.01, −0.04, 0.11, 0.02, 0.27).

The interference angles $\phi_{AB} = \phi_B - \phi_A = 97.66°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 84.49°$, $\phi_{A'B} = \phi_B - \phi_{A'} = 68.25°$ and $\phi_{A'B'} = \phi_{B'} - \phi_{A'} = 113.49°$ complete the Hilbert space representation in $\mathbb{C}^8$. A complete modeling in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$ satisfying Equations (86), (88), (90), and (92) is given by an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.66^2$, $\sum_{i=5}^{8}\sum_{i=1}^{4} c_{ij}^2 = 0.32^2$, $\sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2 = 0.68^2$ and $\sum_{i,j=1}^{4} c_{ij}^2 = 0$, and convex weights $m_{AB} = 0.48$, $n_{AB} = 0.88$, $m_{AB'} = 0.55$, $n_{AB'} = 0.84$, $m_{A'B} = 0.46$, $n_{A'B} = 0.89$, $m_{A'B'} = 0.5$, and $n_{A'B'} = 0.87$.

(iii) Let us then illustrate some relevant exemplars that either cannot be modeled in a pure Hilbert space framework, or cannot be represented by product states in second sector of Fock space.

Raisin, with respect to (*Fruits, Vegetables*). *Raisin* scored $\mu(A) = 0.88$ with respect to *Fruits*, $\mu(B) = 0.27$ with respect to *Vegetables*, $\mu(A') = 0.13$ with respect to *Not Fruits*, $\mu(B') = 0.76$ with respect to *Not Vegetables*, $\mu(A \text{ and } B) = 0.53$ with respect to *Fruits And Vegetables*, $\mu(A \text{ and } B') = 0.75$ with respect to *Fruits And Not Vegetables*, $\mu(A' \text{ and } B) = 0.25$ with respect to *Not Fruits And Vegetables*, and $\mu(A' \text{ and } B') = 0.34$ with respect to *Not Fruits And Not Vegetables*. If one first looks for a representation of *Raisin* in the Hilbert space $\mathbb{C}^8$, then the concepts *Fruits* and *Vegetables*, and their negations *Not Fruits* and *Not Vegetables* are respectively represented by the unit vectors $|A\rangle = e^{i\phi_A}$ (0.05, −0.01, 0.34, 0.01, −0.1, −0.51, 0.23, −0.75) and $|B\rangle = e^{i\phi_B}$ (−0.41, −0.15, −0.73, −0.1, −0.38, −0.17, −0.19, −0.25), and $|A'\rangle = e^{i\phi_{A'}}$ (0.56, −0.73, −0.09, 0.1, −0.08, −0.28, −0.15, 0.16) and $|B'\rangle = e^{i\phi_{B'}}$ (0.07, 0.46, −0.14, 0.04, 0.13, −0.76, −0.11, 0.4).

However, a complete representation satisfying Equations (86), (88), (90), and (92) can only be worked out in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$. This occurs for interference angles $\phi_{AB} = \phi_B - \phi_A = 80.79°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 160°$, $\phi_{A'B} = \phi_B - \phi_{A'} = 18.15°$ and $\phi_{A'B'} = \phi_{B'} - \phi_{A'} = 92.88°$, and for an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.41^2$, $\sum_{i=5}^{8}\sum_{i=1}^{4} c_{ij}^2 = 0.85^2$, $\sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2 = 0.32^2$ and

$\sum_{i,j=1}^{4} c_{ij}^2 = 0.13^2$, and convex weights $m_{AB} = 0.45$, $n_{AB} = 0.89$, $m_{AB'} = 0.65$, $n_{AB'} = 0.76$, $m_{A'B} = 0.26$, $n_{A'B} = 0.97$, $m_{A'B'} = 0.48$, and $n_{A'B'} = 0.88$.

Fox, with respect to (*Pets, Farmyard Animals*). *Fox* scored $\mu(A) = 0.13$ with respect to *Pets*, $\mu(B) = 0.3$ with respect to *Farmyard Animals*, $\mu(A') = 0.86$ with respect to *Not Pets*, $\mu(B') = 0.68$ with respect to *Not Farmyard Animals*, $\mu(A \text{ and } B) = 0.18$ with respect to *Pets And Farmyard Animals*, $\mu(A \text{ and } B') = 0.29$ with respect to *Pets And Not Farmyard Animals*, $\mu(A' \text{ and } B) = 0.46$ with respect to *Not Pets And Farmyard Animals*, and $\mu(A' \text{ and } B') = 0.59$ with respect to *Not Pets And Not Farmyard Animals*. If one first looks for a representation of *Fox* in the Hilbert space $\mathbb{C}^8$, then the concepts *Pets* and *Farmyard Animals*, and their negations *Not Pets* and *Not Farmyard Animals* are respectively represented by the unit vectors $|A\rangle = e^{i\phi_A}$ (−0.07, −0.84, −0.39, −0.03, −0.02, −0.31, 0.02, 0.19) and $|B\rangle = e^{i\phi_B}$ (−0.01, 0.17, −0.82, 0.01, −0.01, 0.28, −0.01, −0.47), and $|A'\rangle = e^{i\phi_{A'}}$ (−0.05, 0.19, −0.31, −0.02, 0.12, 0.39, −0.02, 0.83) and $|B'\rangle = e^{i\phi_{B'}}$ (−0.14, 0.47, −0.26, −0.08, −0.08, −0.8, 0.04, 0.17).

However, a complete representation satisfying Equations (86), (88), (90), and (92) can only be worked out in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$. This occurs for interference angles $\phi_{AB} = \phi_B - \phi_A = 96.58°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 95.05°$, $\phi_{A'B} = \phi_B - \phi_{A'} = 85.68°$ and $\phi_{A'B'} = \phi_{B'} - \phi_{A'} = -20°$, and for an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.05^2$, $\sum_{i=5}^{8}\sum_{i=1}^{4} c_{ij}^2 = 0.36^2$, $\sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2 = 0.55^2$ and $\sum_{i,j=1}^{4} c_{ij}^2 = 0.76^2$, and convex weights $m_{AB} = 0.51$, $n_{AB} = 0.86$, $m_{AB'} = 0.61$, $n_{AB'} = 0.79$, $m_{A'B} = 0.61$, $n_{A'B} = 0.79$, $m_{A'B'} = 0.66$, and $n_{A'B'} = 0.75$.

(iv) Let us finally describe the quantum-theoretic representation of an exemplar that does not present overextension in any conjunction, but still does not admit a representation in a classical Kolmogorovian probability framework.

Window Seat, with respect to (*Home Furnishing, Furniture*). *Window Seat* scored $\mu(A) = 0.5$ with respect to *Home Furnishing*, $\mu(B) = 0.48$ with respect to *Furniture*, $\mu(A') = 0.47$ with respect to *Not Home Furnishing*, $\mu(B') = 0.55$ with respect to *Not Furniture*, $\mu(A \text{ and } B) = 0.45$ with respect to *Home Furnnishing And Furniture*, $\mu(A \text{ and } B') = 0.49$ with respect to *Home Furnishing And Not Furniture*, $\mu(A' \text{ and } B) = 0.39$ with respect to *Not Home Furnishing And Furniture*, and $\mu(A' \text{ and } B') = 0.41$ with respect to *Not Home Furnishing And Not Furniture*. If one first looks for a representation of *Window Seat* in the Hilbert space $\mathbb{C}^8$, then the concepts *Home Furnishing* and *Furniture*, and their negations *Not Home Furnishing* and *Not Furniture* are respectively represented by the unit vectors $|A\rangle = e^{i\phi_A}$ (−0.01, 0.69, 0.14, −0.01, −0.13, −0.66, −0.2, 0.11) and $|B\rangle = e^{i\phi_B}$ (−0.08, −0.39, −0.6, 0, −0.03, −0.4, −0.17, 0.54), and $|A'\rangle = e^{i\phi_{A'}}$ (0.13, −0.19, 0.69, 0.01, 0.09, 0.05, −0.05, 0.67) and $|B'\rangle = e^{i\phi_{B'}}$ (−0.09, 0.57, −0.34, −0.02, 0.17, 0.54, 0.11, 0.47).

The interference angles $\phi_{AB} = \phi_B - \phi_A = 76.57°$, $\phi_{AB'} = \phi_{B'} - \phi_A = 103.86°$, $\phi_{A'B} = \phi_B - \phi_{A'} = 84.42°$ and $\phi_{A'B'} =$

$\phi_{B'} - \phi_{A'} = 85.94°$ complete the Hilbert space representation in $\mathbb{C}^8$. A complete modeling in the Fock space $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$ satisfying Equations (86), (88), (90), and (92) is given by an entangled state characterized by $\sum_{i,j=5}^{8} c_{ij}^2 = 0.31^2$, $\sum_{i=5}^{8} \sum_{j=1}^{4} c_{ij}^2 = 0.64^2$, $\sum_{i=1}^{4} \sum_{j=5}^{8} c_{ij}^2 = 0.62^2$ and $\sum_{i,j=1}^{4} c_{ij}^2 = 0.34^2$, and convex weights $m_{AB} = 0.51$, $n_{AB} = 0.86$, $m_{AB'} = 0.77$, $n_{AB'} = 0.63$, $m_{A'B} = 1$, $n_{A'B} = 0$, $m_{A'B'} = 0.54$, and $n_{A'B'} = 0.84$.

The theoretic analysis on the representatibility of the data in **Tables A1–A4** is thus concluded. We stress that the majority of these data can be faithfully modeled by using the mathematical formalism of quantum theory in Fock space. We finally observe that a big amount of the collected data can be modeled by using only the first sector of Fock space, while almost all the weights of n^2-type in first sector prevail over the weights of m^2-type in second sector. The reasons of this will be clear after the discussion in Section 5.

5. Discussion

Our experimental data on conjunctions and negations of natural concepts confirm that classical probability does not generally work when people combine concepts, as we have seen in the previous sections. And, more, we have proved here that the deviations from classicality cannot be reduced to overextension and underextension, but they also include a very strong and fundamental pattern of violation. On the other side, our quantum-theoretic framework in Fock space has received remarkable corroboration. Thus, we think it worth to summarize and stress the novelties that have emerged in this article with respect to our approach.

We have recently put forward an explanatory hypothesis with respect to the deviations from classical logical reasoning that have been observed in human cognition (Aerts et al., 2015a). According to our explanatory hypothesis, human reasoning is a specifically structured superposition of two processes, a "logical reasoning" and a "conceptual reasoning" (also called "emergent reasoning"). The former "logical reasoning" combines cognitive entities, such as concepts, combinations of concepts, or propositions, by applying the rules of logic, though generally in a probabilistic way. The latter "emergent reasoning" enables formation of combined cognitive entities as newly emerging entities, in the case of concepts, new concepts, in the case of propositions, new propositions, carrying new meaning, linked to the meaning of the constituent cognitive entities, but with a linkage not defined by the algebra of logic. The two mechanisms act simultaneously and in superposition in human thought during a reasoning process, the first one is guided by an algebra of "logic," the second one follows a mechanism of "emergence." In this perspective, human reasoning can be mathematically formalized in a two-sector Fock space. More specifically, first sector of Fock space models "conceptual emergence," while second sector of Fock space models a conceptual combination from the combining concepts by requiring the rules of logic for the logical connective used for the combining to be satisfied in a probabilistic setting. The relative prevalence of emergence or logic in a specific cognitive process is measured by the "degree of participation" of second and first sectors, respectively.

The abundance of evidence of deviations from classical logical reasoning in concrete human decisions (paradoxes, fallacies, effects, contradictions), together with our results, led us to draw the conclusion that emergence constitutes the dominant dynamics of human reasoning, while logic is only a secondary form of dynamics.

Now, if one reflects on how we represented conceptual negation in Section 3, one realizes at once that its modeling directly and naturally follows from the general hypothesis stated above. Indeed, suppose that a person is asked to estimate whether a given exemplar x is a member of the concepts A, B', "A and B' (a completely equivalent explanation can be given for the conjunctions "A' and B" and "A' and B'"). Then, our quantum mathematics can be interpreted by assuming that a "logical thought" acts, where the person considers two copies of x and estimates whether the first copy belongs to A and the second copy of x "does not" belong to B, thus applying logical rules, though in a probabilistic way. But also a "conceptual thought" acts, where the person estimates whether the exemplar x belongs to the newly emergent concept "A and B'." The place whether these superposed processes can be suitably structured is the two-sector Fock space. First sector of Fock space hosts the latter process, second sector hosts the former, hence one expects that classical logical rules are valid in this sector, though they are generally violated whenever both sectors are considered. The weights $m_{AB'}^2$ and $n_{AB'}^2$ indicate whether the overall process is mainly guided by logic or emergence.

The second confirmation of our quantum-conceptual framework comes from the significantly stable deviations from classicality in Equations (26–30). We have seen in Section 2.4 that these deviations occur at a different, deeper, level than overextension and underextension. We think we have identified a general mechanism determining how concepts are formed in the human mind. And this would already be convincing even without mentioning a Fock space modeling. But, this very stable pattern can exactly be explained in our two-sector Fock space framework by assuming that emergence plays a primary role in the human reasoning process, but also aspects of logic are systematically present. Indeed, suppose that, for every exemplar x and every $X = A, A', Y = B, B', n_{XY}^2 = 1$ and $m_{XY}^2 = 0$, that is, the decision process only occurs in first sector of Fock space. This assumption corresponds, from our quantum modeling perspective (see Equations 86–92), to a situation where only emergence is present. We have then argued in Aerts et al. (2015b) that, for every $X = A, A', Y = B, B', I_X = I_Y = -0.5$ and $I_{ABA'B'} = -1$, in this case. Then, an immediate comparison with the experimental values of I_X, I_Y, and $I_{ABA'B'}$ in Section 2.4 reveals that a component of second sector of Fock space is also present, which is generally smaller than the component of first sector but systematic across all exemplars. The consequence is immediate: both emergence and logic play a role in the decision process—emergence is dominant, but also logic is systematically present. We believe that this finding is really a fundamental one, and it deserves further investigation in the future.

The third strong confirmation of this two-layered structure of human thought and its representation in two-sector Fock space comes from the peculiarities of conceptual negation. Indeed, being pushed to cope with conceptual conjunction and negation simultaneously, we have found a new insight which we had not noticed before, namely, the emergent non-classical properties of the

conjunction *Fruits And Vegetables* are naturally accounted for in first sector of Fock space, while the fact that *Not Vegetables* does not have a well defined prototype, but it is rather the negation of *Vegetables*, is accounted for in second sector of Fock space, where logic occurs. In both cases, the Fock space model has naturally suggested us the right directions to follow.

The fourth corroboration derives from the fact that our Fock space indicates how and why introducing entanglement. In our previous attempts to model conceptual combinations, we had not recognized that by representing the combined concept by a tensor product vector $|A\rangle \otimes |B\rangle$, we implicitly assumed that membership weights probabilities are factorised in second sector, that is, the membership weights $\mu(A)$ and $\mu(B)$ correspond to independent events in this sector. In Section 3.2 we have showed that, if one introduces entangled states to represent combined concepts in second sector, one is able to fully reproduce all classicality conditions Equations (26–30) in this sector. And, more, one can formalize the fact that, for certain exemplars, the probabilities associated with memberships of, say *Fruits* and *Vegetables*, are not independent. Therefore, Fock space has suggested how to capture this relevant aspect in depth.

The discussion above shows, in our opinion, that the merits of our two-sector Fock space framework go beyond faithful representation of one or more sets of experimental data. It captures some fundamental aspects of the mechanisms through which concepts are formed, combine and interact in human cognition.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01447

References

Aerts, D., and Sozzo, S. (2011). Quantum structure in cognition. Why and how concepts are entangled. *Quantum Interact. Lect. Notes Comput. Sci.* 7052, 116–127. doi: 10.1007/978-3-642-24971-6/12

Aerts, D., and Aerts, S. (1995). Applications of quantum statistics in psychological studies of decision processes. *Found. Sci.* 1, 85–97. doi: 10.1007/BF00208726

Aerts, D., and Gabora, L. (2005a). A theory of concepts and their combinations I: the structure of the sets of contexts and properties. *Kybernetes* 34, 167–191. doi: 10.1108/03684920510575799

Aerts, D., and Gabora, L. (2005b). A theory of concepts and their combinations II: a Hilbert space representation. *Kybernetes* 34, 192–221. doi: 10.1108/03684920510575807

Aerts, D., and Sozzo, S. (2014). Quantum entanglement in conceptual combinations. *Int. J. Theor. Phys.* 53, 3587–3603. doi: 10.1007/s10773-013-1946-z

Aerts, D., Aerts, S., Broekaert, J., and Gabora, L. (2000). The violation of Bell inequalities in the macroworld. *Found. Phys.* 30, 1387–1414. doi: 10.1023/A:1026449716544

Aerts, D., Sozzo, S., and Tapia, J. (2012). A quantum model for the ellsberg and machina paradoxes. *Quantum Interact. Lect. Notes Comput. Sci.* 7620, 48–59. doi: 10.1007/978-3-642-35659-9/5

Aerts, D., Broekaert, J., Gabora, L., and Sozzo, S. (2013). Quantum structure and human thought. *Behav. Brain Sci.* 36, 274–276. doi: 10.1017/S0140525X12002841

Aerts, D., Gabora, L., and S. Sozzo, S. (2013). Concepts and their dynamics: a quantum–theoretic modeling of human thought. *Top. Cogn. Sci.* 5, 737–772. doi: 10.1111/tops.12042

Aerts, D., Sozzo, S., and Tapia, J. (2014). Identifying quantum structures in the Ellsberg paradox. *Int. J. Theor. Phys.* 53, 3666–3682. doi: 10.1007/s10773-014-2086-9

Aerts, D., Sozzo, S., and Veloz, T. (2015a). Quantum structure in cognition and the foundations of human reasoning. *Int. J. Theor. Phys.* doi: 10.1007/s10773-015-2717-9. [Epub ahead of print].

Aerts, D., Sozzo, S., and Veloz, T. (2015b). A new fundamental evidence of non-classical structure in the combination of natural concepts. *Philos. Trans. R. Soc. A.* arXiv: 1505.04981.

Aerts, D., Sozzo, S., and Veloz, T. (2015c). Quantum nature of identity in human concepts: bose-einstein statistics for conceptual indistinguisibility. *Int. J. Theor. Phys.* doi: 10.1007/s10773-015-2620-4. [Epub ahead of print].

Aerts, D. (1986). A possible explanation for the probabilities of quantum mechanics. *J. Math. Phys.* 27, 202–210. doi: 10.1063/1.527362

Aerts, D. (1999). Foundations of quantum physics: a general realistic and operational approach. *Int. J. Theor. Phys.* 38, 289–358. doi: 10.1023/A:1026605829007

Aerts, D. (2009). Quantum structure in cognition. *J. Math. Psychol.* 53, 314–348. doi: 10.1016/j.jmp.2009.04.005

Alxatib, S., and Pelletier, J. (2011). "On the psychology of truth gaps," in *Vagueness in Communication*, eds R. Nouwen, R. van Rooij, U. Sauerland, and H.-C. Schmitz (Berlin; Heidelberg: Springer-Verlag), 13–36.

Bonini, N., Osherson, D., Viale, R., and Williamson, T. (1999). On the psychology of vague predicates. *Mind Lang.* 14, 377–393. doi: 10.1111/1468-0017.00117

Busemeyer, J. R., and Bruza, P. D. (2012). *Quantum Models of Cognition and Decision.* Cambridge: Cambridge University Press.

Dirac, P. A. M. (1958). *Quantum mechanics, 4th Edn.* London: Oxford University Press.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Q. J. Econ.* 75, 643–669. doi: 10.2307/1884324

Fodor, J. (1994). Concepts: a potboiler. *Cognition* 50, 95–113. doi: 10.1016/0010-0277(94)90023-X

Hampton, J. A. (1988a). Overextension of conjunctive concepts: evidence for a unitary model for concept typicality and class inclusion. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 12–32. doi: 10.1037/0278-7393.14.1.12

Hampton, J. A. (1988b). Disjunction of natural concepts. *Mem. Cogn.* 16, 579–591. doi: 10.3758/BF03197059

Hampton, J. A. (1997). Conceptual combination: conjunction and negation of natural concepts. *Mem. Cogn.* 25, 888–909. doi: 10.3758/BF03211333

Haven, E., and Khrennikov, A. Y. (2013). *Quantum Social Science.* Cambridge: Cambridge University Press.

Kamp, H., and Partee, B. (1995). Prototype theory and compositionality. *Cognition* 57, 129–191. doi: 10.1016/0010-0277(94)00659-9

Khrennikov, A. Y. (2010). *Ubiquitous Quantum Structure.* Berlin: Springer.

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitrechnung,* Ergebnisse Der Mathematik; translated as *Foundations of Probability.* New York, NY: Chelsea Publishing Company, 1950.

Komatsu, L. K. (1992). Recent views on conceptual structure. *Psychol. Bull.* 112, 500–526. doi: 10.1037/0033-2909.112.3.500

Machina, M. J. (2009). Risk, ambiguity, and the dark dependence axioms. *Am. Econ. Rev.* 99, 385–392. doi: 10.1257/aer.99.1.385

Minsky, M. (1975). "A framework for representing knowledge," in *The Psychology of Computer Vision,* ed P. H. Winston (New York, NY: McGraw-Hill), 211–277.

Murphy, G. L., and Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychol. Rev.* 92, 289–316. doi: 10.1037/0033-295X.92.3.289

Nosofsky, R. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 700–708. doi: 10.1037/0278-7393.14.4.700

Nosofsky, R. (1992). "Exemplars, prototypes, and similarity rules," in *From Learning Theory to Connectionist Theory: Essays in Honor of William K. Estes,* eds A. Healy, S. Kosslyn, and R. Shiffrin (Hillsdale, NJ: Erlbaum), 149–167.

Osherson, D., and Smith, E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition* 9, 35–58. doi: 10.1016/0010-0277(81)90013-5

Osherson, D. N., and Smith, E. (1982). Gradedness and conceptual combination. *Cognition* 12, 299–318. doi: 10.1016/0010-0277(82)90037-3

Osherson, D. N., and Smith, E. (1997). On typicality and vagueness. *Cognition* 64, 189–206. doi: 10.1016/S0010-0277(97)00025-5

Pitowsky, I. (1989). *Quantum Probability, Quantum Logic*, Vol. 321. Lecture Notes in Physics. Berlin: Springer.

Pothos, E. M., and Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* 36, 255–274. doi: 10.1017/S0140525X12001525

Rips, L. J. (1995). The current status of research on concept combination. *Mind Lang.* 10, 72–104. doi: 10.1111/j.1468-0017.1995.tb00006.x

Rosch, E. (1973). Natural categories. *Cogn. Psychol.* 4, 328–350. doi: 10.1016/0010-0285(73)90017-0

Rosch, E. (1978). "Principles of categorization," in *Cognition and Categorization*, eds E. Rosch and B. Lloyd (Hillsdale, NJ: Lawrence Erlbaum), 133–179.

Rosch, E. (1983). "Prototype classification and logical classification: the two systems," in *New Trends in Conceptual Representation: Challenges to Piaget's Theory?*, ed E. K. Scholnick (Hillsdale, NJ: Lawrence Erlbaum), 133–159.

Rumelhart, D. E., and Norman, D. A. (1988). "Representation in memory," in *Stevens Handbook of Experimental Psychology*, eds R. C. Atkinson, R. J. Hernsein, G. Lindzey, and R. L. Duncan (New York, NY: John Wiley & Sons), 511–587.

Sozzo, S. (2014). A quantum probability explanation in Fock space for borderline contradictions. *J. Math. Psychol.* 58, 1–12. doi: 10.1016/j.jmp.2013.11.001

Sozzo, S. (2015). Conjunction and negation of natural concepts: a quantum-theoretic modeling. *J. Math. Psychol.* 66, 83–102. doi: 10.1016/j.jmp.2015.01.005

Wang, Z., Solloway, T., Shiffrin, R. M., and Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9431–9436. doi: 10.1073/pnas.1407756111

Zadeh, L. (1982). A note on prototype theory and fuzzy sets. *Cognition* 12, 291–297. doi: 10.1016/0010-0277(82)90036-1

Appendix

A1. Fundamentals of Modeling in a Classical Framework

We introduce in this section the elementary measure-theoretic notions that are needed to express the classicality of experimental data coming from the membership weights of two concepts A and B with respect to the conceptual negation "not B" and the conjunctions "A and B," "A and not B," "not A and B," and "not A and not B." As we have anticipated in Section 2, by "classicality of a collection of experimental date" we actually mean the possibility to represent them in a "classical," or "Kolmogorovian," probability model. We avoid in our presentation superfluous technicalities, but aim to be synthetic and rigorous at the same time.

Let us start by the definition of a σ-algebra over a set.

First definition. A σ-algebra over a set Ω is a non-empty collection $\sigma(\Omega)$ of subsets of Ω that is closed under complementation and countable unions of its members. It is a Boolean algebra, completed to include countably infinite operations.

Measure structures are the most general classical structures devised by mathematicians and physicists to structure weights. A Kolmogorovian probability measure is such a measure applied to statistical data. It is called "Kolmogorovian," because Andrey Kolmogorov was the first to axiomatize probability theory in this way (Kolmogorov, 1933).

Second definition. A *measure P* is a function defined on a σ-algebra $\sigma(\Omega)$ over a set Ω and taking values in the extended interval $[0, \infty]$ such that the following three conditions are satisfied:

(i)   the empty set has measure zero;
(ii) if E_1, E_2, E_3, ... is a countable sequence of pairwise disjoint sets in $\sigma(\Omega)$, the measure of the union of all the E_i is equal to the sum of the measures of each E_i (*countable additivity*, or σ-*additivity*);
(iii) the triple $(\Omega, \sigma(\Omega), P)$ satisfying (i) and (ii) is then called a measure space, and the members of $\sigma(\Omega)$ are called measurable sets.

A *Kolmogorovian probability measure* is a measure with total measure one. A Kolmogorovian probability space $(\Omega, \sigma(\Omega), P)$ is a measure space $(\Omega, \sigma(\Omega), P)$ such that P is a Kolmogorovian probability. The three conditions expressed in a mathematical way are:

$$P(\emptyset) = 0 \quad P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i) \quad P(\Omega) = 1 \qquad (A1)$$

Let us now come to the possibility to represent a set of experimental data on two concepts and their conjunction in a classical Kolmogorovian probability model.

Third definition. We say that the membership weights $\mu(A)$, $\mu(B)$ and $\mu(A$ and $B)$ of the exemplar x with respect to the pair of concepts A and B and their conjunction "A and B," respectively, can be represented in a *classical Kolmogorovian probability model* if there exists a Kolmogorovian probability space $(\Omega, \sigma(\Omega), P)$ and events $E_A, E_B \in \sigma(\Omega)$ of the events algebra $\sigma(\Omega)$ such that

$$P(E_A) = \mu(A) \quad P(E_B) = \mu(B) \quad \text{and} \quad P(E_A \cap E_B) = \mu(A \text{ and } B) \tag{A2}$$

Let us finally come to the representability a set of experimental data on a concept and its negation in a classical Kolmogorovian probability model.

Fourth definition. We say that the membership weights $\mu(B)$ and $\mu(\text{not } B)$ of the exemplar x with respect to the concept B and its negation "not B," respectively, can be represented in a classical Kolmogorovian probability model if there exists a Kolmogorovian probability space $(\Omega, \sigma(\Omega), P)$ and an event $E_B \in \sigma(\Omega)$ of the events algebra $\sigma(\Omega)$ such that

$$P(E_B) = \mu(B) \quad P(\Omega \setminus E_B) = \mu(\text{not } B) \tag{A3}$$

A2. Quantum Mathematics for Conceptual Modeling

We illustrate in this section how the mathematical formalism of quantum theory can be applied to model situations outside the microscopic quantum world, more specifically, in the representation of concepts and their combinations. As in Appendix A1, we will limit technicalities to the essential.

When the quantum mechanical formalism is applied for modeling purposes, each considered entity—in our case a concept—is associated with a complex Hilbert space $\mathcal{H}$, that is, a vector space over the field $\mathbb{C}$ of complex numbers, equipped with an inner product $\langle \cdot | \cdot \rangle$ that maps two vectors $\langle A |$ and $| B \rangle$ onto a complex number $\langle A | B \rangle$. We denote vectors by using the bra-ket notation introduced by Paul Adrien Dirac, one of the pioneers of quantum theory (Dirac, 1958). Vectors can be "kets," denoted by $|A\rangle$, $|B\rangle$, or "bras," denoted by $\langle A |$, $\langle B |$. The inner product between the ket vectors $|A\rangle$ and $|B\rangle$, or the bra-vectors $\langle A |$ and $\langle B |$, is realized by juxtaposing the bra vector $\langle A |$ and the ket vector $|B\rangle$, and $\langle A | B \rangle$ is also called a "bra-ket," and it satisfies the following properties:

(i) $\langle A | A \rangle \geq 0$;
(ii) $\langle A | B \rangle = \langle B | A \rangle^*$, where $\langle B | A \rangle^*$ is the complex conjugate of $\langle A | B \rangle$;
(iii) $\langle A | (z | B \rangle + t | C \rangle) = z \langle A | B \rangle + t \langle A | C \rangle$, for $z, t \in \mathbb{C}$, where the sum vector $z | B \rangle + t | C \rangle$ is called a "superposition" of vectors $|B\rangle$ and $|C\rangle$ in the quantum jargon.

From (ii) and (iii) follows that inner product $\langle \cdot | \cdot \rangle$ is linear in the ket and anti-linear in the bra, i.e., $(z \langle A | + t \langle B |) | C \rangle = z^* \langle A | C \rangle + t^* \langle B | C \rangle$.

We recall that the "absolute value" of a complex number is defined as the square root of the product of this complex number times its complex conjugate, that is, $|z| = \sqrt{z^* z}$. Moreover, a complex number z can either be decomposed into its cartesian form $z = x + iy$, or into its polar form $z = |z| e^{i\theta} = |z|(\cos \theta + i \sin \theta)$. As a consequence, we have $|\langle A | B \rangle| = \sqrt{\langle A | B \rangle \langle B | A \rangle}$. We define the "length" of a ket (bra) vector $|A\rangle$ ($\langle A |$) as $|| |A\rangle || = || \langle A | || = \sqrt{\langle A | A \rangle}$. A vector of unitary length is called a "unit vector'. We say that the ket vectors $|A\rangle$ and $|B\rangle$ are "orthogonal" and write $|A\rangle \perp |B\rangle$ if $\langle A | B \rangle = 0$.

We have now introduced the necessary mathematics to state the first modeling rule of quantum theory, as follows.

First quantum modeling rule: A state A of an entity—in our case a concept—modeled by quantum theory is represented by a ket vector $|A\rangle$ with length 1, that is $\langle A|A\rangle = 1$.

An orthogonal projection M is a linear operator on the Hilbert space, that is, a mapping $M : \mathcal{H} \rightarrow \mathcal{H}, |A\rangle \mapsto M|A\rangle$ which is Hermitian and idempotent. The latter means that, for every $|A\rangle, |B\rangle \in \mathcal{H}$ and $z, t \in \mathbb{C}$, we have:

(i) $M(z|A\rangle + t|B\rangle) = zM|A\rangle + tM|B\rangle$ (linearity);
(ii) $\langle A|M|B\rangle = \langle B|M|A\rangle^*$ (hermiticity);
(iii) $M \cdot M = M$ (idempotency).

The identity operator $\mathbb{1}$ maps each vector onto itself and is a trivial orthogonal projection. We say that two orthogonal projections M_k and M_l are orthogonal operators if each vector contained in $M_k(\mathcal{H})$ is orthogonal to each vector contained in $M_l(\mathcal{H})$, and we write $M_k \perp M_l$, in this case. The orthogonality of the projection operators M_k and M_l can also be expressed by $M_k M_l = 0$, where 0 is the null operator. A set of orthogonal projection operators $\{M_k \,|\, k = 1, \ldots, n\}$ is called a "spectral family" if all projectors are mutually orthogonal, that is, $M_k \perp M_l$ for $k \neq l$, and their sum is the identity, that is, $\sum_{k=1}^{n} M_k = \mathbb{1}$.

The above definitions give us the necessary mathematics to state the second modeling rule of quantum theory, as follows.

Second quantum modeling rule: A measurable quantity Q of an entity—in our case a concept— modeled by quantum theory, and having a set of possible real values $\{q_1, \ldots, q_n\}$ is represented by a spectral family $\{M_k \,|\, k = 1, \ldots, n\}$ in the following way. If the entity—in our case a concept—is in a state represented by the vector $|A\rangle$, then the probability of obtaining the value q_k in a measurement of the measurable quantity Q is $\langle A|M_k|A\rangle = ||M_k|A\rangle||^2$. This formula is called the "Born rule" in the quantum jargon. Moreover, if the value q_k is actually obtained in the measurement, then the initial state is changed into a state represented by the vector

$$|A_k\rangle = \frac{M_k|A\rangle}{||M_k|A\rangle||} \tag{A4}$$

This change of state is called "collapse" in the quantum jargon.

The tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is the Hilbert space generated by the set $\{|A_i\rangle \otimes |B_j\rangle\}$, where $|A_i\rangle$ and $|B_j\rangle$ are vectors of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, which means that a general vector of this tensor product is of the form $\sum_{ij} |A_i\rangle \otimes |B_j\rangle$. This gives us the necessary mathematics to introduce the third modeling rule.

Third quantum modeling rule: A state C of a compound entity—in our case a combined concept—is represented by a unit vector $|C\rangle$ of the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of the two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ containing the vectors that represent the states of the component entities—concepts.

The above means that we have $|C\rangle = \sum_{ij} c_{ij}|A_i\rangle \otimes |B_j\rangle$, where $|A_i\rangle$ and $|B_j\rangle$ are unit vectors of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, and $\sum_{i,j} |c_{ij}|^2 = 1$. We say that the state C represented by $|C\rangle$ is a product state if it is of the form $|A\rangle \otimes |B\rangle$ for some $|A\rangle \in \mathcal{H}_A$ and $|B\rangle \in \mathcal{H}_B$. Otherwise, C is called an "entangled state".

The Fock space is a specific type of Hilbert space, originally introduced in quantum field theory. For most states of a quantum field the number of identical quantum entities is not conserved but is a variable quantity. The Fock space copes with this situation in allowing its vectors to be superpositions of vectors pertaining to different sectors for fixed numbers of identical quantum entities. More explicitly, the k-th sector of a Fock space describes a fixed number of k identical quantum entities, and it is of the form $\mathcal{H} \otimes \ldots \otimes \mathcal{H}$ of the tensor product of k identical Hilbert spaces $\mathcal{H}$. The Fock space F itself is the direct sum of all these sectors, hence

$$\mathcal{F} = \oplus_{k=1}^{j} \otimes_{l=1}^{k} \mathcal{H} \tag{A5}$$

For our modeling we have only used Fock space for the "two" and "one quantum entity" case, hence $\mathcal{F} = \mathcal{H} \oplus (\mathcal{H} \otimes \mathcal{H})$. This is due to considering only combinations of two concepts. The sector $\mathcal{H}$ is called the "first sector," while the sector $\mathcal{H} \otimes \mathcal{H}$ is called the "second sector". A unit vector $|F\rangle \in \mathcal{F}$ is then written as $|F\rangle = ne^{i\gamma}|C\rangle + me^{i\delta}(|A\rangle \otimes |B\rangle)$, where $|A\rangle, |B\rangle$ and $|C\rangle$ are unit vectors of $\mathcal{H}$, and such that $n^2 + m^2 = 1$. For combinations of j concepts, the general form of Fock space in (A5) should be used.

A3. Data Modeling Tables and Statistical Analysis

Tables A1–A5A–E.

A4. Proofs of Theorems 1–4

Proof of Theorem 1. If $\mu(A), \mu(B), \mu(A'), \mu(B')$ and $\mu(A \text{ and } B), \mu(A \text{ and } B'), \mu(A' \text{ and } B), \mu(A' \text{ and } B')$ are classical conjunction and negation data, then there exists a Kolmogorovian probability space $(\Omega, \sigma(\Omega), P)$ and events $E_A, E_B \in \sigma(\Omega)$ such that $P(E_A) = \mu(A), P(E_B) = \mu(B), P(\Omega \setminus E_A) = \mu(A'), P(\Omega \setminus E_B) = \mu(B'), P(E_A \cap E_B) = \mu(A \text{ and } B), P(E_A \cap (\Omega \setminus E_B)) = \mu(A \text{ and } B'), P((\Omega \setminus E_A) \cap E_B) = \mu(A' \text{ and } B)$ and $P((\Omega \setminus E_A) \cap (\Omega \setminus E_B)) = \mu(A' \text{ and } B')$. From the general properties of a Kolmogorovian probability space it follows that (1), (2), (3), (4), (5), (6), (7) and (8) are satisfied.

Now suppose that x is such that its membership weights $\mu(A), \mu(B), \mu(A'), \mu(B'), \mu(A \text{ and } B), \mu(A \text{ and } B')$ and $\mu(A' \text{ and } B)$ and $\mu(A' \text{ and } B')$ with respect to the concepts A, B, A', B', "A and B," "A and B'," "A' and B" and "A' and B'," respectively, satisfy (1), (2), (3), (4), (5), (6), (7) and (8). We will prove that as a consequence $\mu(A), \mu(B), \mu(A'), \mu(B'), \mu(A \text{ and } B), \mu(A \text{ and } B')$ and $\mu(A' \text{ and } B)$ and $\mu(A' \text{ and } B')$ are classical conjunction and negation data, in the sense that "there exists a classical Kolmogorovian probability space, such that we can represent all of them as measures on event sets of this space". We make our proof by explicitly constructing a Kolmogorovian probability space that models these data. Consider the set $\Omega = \{1, 2, 3, 4\}$ and $\sigma(\Omega) = \mathcal{P}(\Omega)$, the set of all subsets of Ω. We define

$$P(\{1\}) = \mu(A \text{ and } B) \tag{A6}$$
$$P(\{2\}) = \mu(A \text{ and } B') = \mu(A) - \mu(A \text{ and } B) \tag{A7}$$
$$P(\{3\}) = \mu(A' \text{ and } B) = \mu(A') - \mu(A' \text{ and } B') \tag{A8}$$
$$P(\{4\}) = \mu(A' \text{ and } B') \tag{A9}$$

and further for an arbitrary subset $S \subseteq \{1, 2, 3, 4\}$ we define

TABLE A1 | Representation of the membership weights in the case of the concepts *Home Furnishing* and *Furniture*.

Exemplar	$\mu(A)$	$\mu(B)$	$\mu(A')$	$\mu(B')$	$\mu(A$ and $B)$	$\mu(A$ and $B')$	$\mu(A'$ and $B)$	$\mu(A'$ and $B')$	Δ_{AB}	$\Delta_{AB'}$	$\Delta_{A'B}$	$\Delta_{A'B'}$	$I_{ABA'B'}$	I_A	I_B	$I_{A'}$	$I_{B'}$
Mantelpiece	0.9	0.61	0.12	0.5	0.71	0.75	0.21	0.21	0.1	0.25	0.09	0.09	−0.89	−0.56	−0.31	−0.31	−0.46
Window Seat	0.5	0.48	0.47	0.55	0.45	0.49	0.39	0.41	−0.03	−0.01	−0.08	−0.06	−0.74	−0.44	−0.36	−0.33	−0.35
Painting	0.8	0.49	0.35	0.64	0.64	0.6	0.33	0.38	0.15	−0.04	−0.03	0.03	−0.94	−0.44	−0.48	−0.35	−0.33
Light Fixture	0.88	0.6	0.16	0.51	0.73	0.63	0.33	0.16	0.13	0.11	0.16	0	−0.84	−0.48	−0.45	−0.33	−0.28
Kitchen Counter	0.67	0.49	0.31	0.62	0.55	0.54	0.38	0.33	0.06	−0.08	0.06	0.01	−0.79	−0.42	−0.44	−0.39	−0.24
Bath Tub	0.73	0.51	0.28	0.46	0.59	0.59	0.36	0.29	0.08	0.13	0.08	0.01	−0.83	−0.45	−0.44	−0.37	−0.41
Deck Chair	0.73	0.9	0.27	0.2	0.74	0.41	0.54	0.18	0.01	0.21	0.27	−0.03	−0.86	−0.42	−0.38	−0.44	−0.39
Shelves	0.85	0.93	0.24	0.13	0.84	0.39	0.53	0.08	−0.01	0.26	0.29	−0.05	−0.83	−0.38	−0.43	−0.36	−0.34
Rug	0.89	0.58	0.18	0.61	0.7	0.68	0.41	0.21	0.13	0.07	0.24	0.04	−1	−0.48	−0.54	−0.45	−0.28
Bed	0.76	0.93	0.26	0.11	0.79	0.36	0.61	0.14	0.03	0.26	0.35	0.03	−0.9	−0.39	−0.48	−0.49	−0.39
Wall-Hangings	0.87	0.46	0.21	0.68	0.55	0.71	0.35	0.24	0.09	0.03	0.14	0.03	−0.85	−0.39	−0.44	−0.38	−0.27
Space Rack	0.38	0.43	0.63	0.62	0.41	0.49	0.43	0.58	0.04	0.11	0	−0.04	−0.9	−0.53	−0.41	−0.37	−0.44
Ashtray	0.74	0.4	0.32	0.64	0.49	0.6	0.36	0.39	0.09	−0.04	0.04	0.07	−0.84	−0.34	−0.45	−0.43	−0.35
Bar	0.72	0.63	0.37	0.51	0.61	0.61	0.4	0.4	−0.01	0.11	0.03	0.03	−1.03	−0.51	−0.39	−0.43	−0.51
Lamp	0.94	0.64	0.15	0.49	0.75	0.7	0.4	0.2	0.11	0.21	0.25	0.05	−1.05	−0.51	−0.51	−0.45	−0.41
Wall Mirror	0.91	0.76	0.13	0.45	0.83	0.66	0.44	0.14	0.07	0.21	0.31	0.01	−1.06	−0.58	−0.51	−0.45	−0.35
Door Bell	0.75	0.33	0.32	0.79	0.5	0.64	0.34	0.51	0.17	−0.11	0.02	0.19	−0.99	−0.39	−0.51	−0.53	−0.36
Hammock	0.62	0.66	0.41	0.41	0.6	0.5	0.56	0.31	−0.02	0.09	0.16	−0.09	−0.98	−0.48	−0.5	−0.47	−0.41
Desk	0.78	0.95	0.31	0.09	0.78	0.33	0.75	0.15	−0.01	0.24	0.44	0.06	−1	−0.32	−0.58	−0.59	−0.39
Refrigerator	0.74	0.73	0.26	0.41	0.66	0.55	0.46	0.25	−0.06	0.14	0.21	−0.01	−0.93	−0.47	−0.4	−0.46	−0.39
Park Bench	0.53	0.66	0.59	0.46	0.55	0.29	0.56	0.39	0.02	−0.17	−0.03	−0.07	−0.79	−0.31	−0.45	−0.36	−0.22
Waste Paper Basket	0.69	0.54	0.36	0.63	0.59	0.41	0.46	0.49	0.04	−0.22	0.1	0.13	−0.95	−0.31	−0.51	−0.59	−0.27
Sculpture	0.83	0.46	0.34	0.66	0.58	0.73	0.46	0.36	0.11	0.07	0.13	0.03	−1.13	−0.48	−0.58	−0.49	−0.43
Sink Unit	0.71	0.57	0.34	0.58	0.6	0.56	0.38	0.38	0.03	−0.01	0.04	0.04	−0.91	−0.46	−0.41	−0.41	−0.36

A = Home Furnishing, B = Furniture.

$$P(S) = \sum_{a \in S} P(\{a\}) \qquad (A10)$$

Let us prove that $P : \sigma(\Omega) \to [0, 1]$ is a probability measure. For this purpose, we need to prove that $P(S) \in [0, 1]$ for an arbitrary subset $S \subseteq \Omega$, and that the "sum formula" for a probability measure is satisfied. The sum formula for a probability measure is satisfied because of definition (A10). What remains to be proved is that $P(S) \in [0, 1]$ for an arbitrary subset $S \subseteq \Omega$, and that all different subsets that can be formed are contained in $\sigma(\Omega)$. $P(\{1\})$, $P(\{2\})$, $P(\{3\})$ and $P(\{4\})$ are contained in $[0, 1]$ as a consequence of equations (1), (3), (5) and (6). Using (5) we have that $P(\{1, 2\}) = \mu(A$ and $B) + \mu(A$ and $B') = \mu(A$ and $B) + \mu(A) - \mu(A$ and $B) = \mu(A)$. Using (6) we have that $P(\{3, 4\}) = \mu(A'$ and $B') + \mu(A'$ and $B) = \mu(A'$ and $B') + \mu(A') - \mu(A'$ and $B') = \mu(A')$. Again using (6) we have that $P(\{1, 3\}) = \mu(A$ and $B) + \mu(A'$ and $B) = \mu(A$ and $B) + \mu(B) - \mu(A$ and $B) = \mu(B)$, and using again (5) we have that $P(\{2, 4\}) = \mu(A$ and $B') + \mu(A'$ and $B') = \mu(B') - \mu(A'$ and $B') + \mu(A'$ and $B') = \mu(B')$. Moreover, $P(\{1, 2\})$, $P(\{3, 4\})$, $P(\{1, 3\})$ and $P(\{2, 4\})$ are all contained in $[0, 1]$ as a consequence of equations (1), (2), (3) and (4). We have already found the representatives of all elements and their conjunctions in $\sigma(\Omega)$. But we have not yet considered all subsets of Ω. Indeed, let us consider $\mu(\{1, 2, 3\}) = \mu(A$ and $B) + \mu(A$ and $B') + \mu(A'$ and $B) = 1 - \mu(A'$ and $B')$. And from (7) it follows that this is contained in $[0.1]$. In an analogous way we prove that $\mu(\{1, 2, 4\}) = 1 - \mu(A'$ and $B)$, $\mu(\{1, 3, 4\}) = 1 - \mu(A$ and $B')$, and $\mu(\{2, 3, 4\}) = 1 - \mu(A$ and $B)$. We almost have all subsets of Ω. Let us consider $\{1, 4\}$ and $\{2, 3\}$. Since by construction we have $\mu(\{1\}) \leq \mu(\{1, 4\}) \leq \mu(\{1, 2, 4\})$ and $\mu(\{2\}) \leq \mu(\{2, 3\}) \leq \mu(\{2, 3, 4\})$, it follows that both $\mu(\{1, 4\})$ and $\mu(\{2, 3\})$ are contained in $[0, 1]$. The last subset to control is Ω itself. We have $P(\Omega) = P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) = 1$, following the calculation we made above. We have verified all subsets $S \subseteq \Omega$, and hence proved that P is a probability measure. All subsets for which we have gathered data are represented in this σ-algebra, which completes our proof. $\qquad \square$

Proof of Theorem 2. Let us consider Theorem 1. In its proof, we did not use (8),„ which means that inequalities (5), (6), (7) and (8) are not independent. By using, for example, (5), (6), and then (7) we get

$$1 - \mu(A') - \mu(B') + \mu(A' \text{ and } B')$$
$$= 1 - \mu(B) - \mu(B') + \mu(A \text{ and } B)$$

TABLE A2 | Representation of the membership weights in the case of the concepts *Spices* and *Herbs*.

Exemplar	$\mu(A)$	$\mu(B)$	$\mu(A')$	$\mu(B')$	$\mu(A$ and $B)$	$\mu(A$ and $B')$	$\mu(A'$ and $B)$	$\mu(A'$ and $B')$	Δ_{AB}	$\Delta_{AB'}$	$\Delta_{A'B}$	$\Delta_{A'B'}$	$I_{ABA'B'}$	I_A	I_B	$I_{A'}$	$I_{B'}$
Molasses	0.36	0.13	0.67	0.84	0.24	0.54	0.25	0.73	0.11	0.18	0.12	0.06	−0.75	−0.41	−0.36	−0.31	−0.43
Salt	0.67	0.04	0.36	0.92	0.24	0.69	0.09	0.6	0.19	0.02	0.04	0.24	−0.61	−0.26	−0.28	−0.33	−0.37
Peppermint	0.67	0.93	0.38	0.1	0.7	0.38	0.55	0.15	0.03	0.28	0.18	0.05	−0.78	−0.41	−0.33	−0.33	−0.43
Curry	0.96	0.28	0.04	0.78	0.54	0.88	0.16	0.21	0.26	0.1	0.13	0.18	−0.79	−0.45	−0.42	−0.34	−0.31
Oregano	0.81	0.86	0.21	0.13	0.79	0.4	0.5	0.08	−0.03	0.28	0.29	−0.05	−0.76	−0.38	−0.43	−0.36	−0.35
MSG	0.44	0.12	0.59	0.85	0.23	0.58	0.24	0.73	0.11	0.13	0.12	0.13	−0.76	−0.36	−0.34	−0.37	−0.45
Chili Pepper	0.98	0.53	0.05	0.56	0.8	0.9	0.28	0.13	0.27	0.34	0.23	0.08	−1.1	−0.73	−0.54	−0.35	−0.46
Mustard	0.65	0.28	0.39	0.71	0.49	0.65	0.23	0.46	0.21	0	−0.05	0.08	−0.83	−0.49	−0.44	−0.3	−0.41
Mint	0.64	0.96	0.43	0.09	0.79	0.31	0.64	0.11	0.14	0.23	0.21	0.03	−0.85	−0.46	−0.47	−0.32	−0.34
Cinnamon	1	0.49	0.02	0.51	0.69	0.79	0.21	0.15	0.19	0.28	0.19	0.13	−0.84	−0.48	−0.41	−0.34	−0.43
Parsley	0.54	0.9	0.54	0.09	0.68	0.26	0.73	0.18	0.14	0.18	0.19	0.09	−0.84	−0.4	−0.5	−0.36	−0.35
Saccarin	0.34	0.14	0.68	0.88	0.24	0.54	0.24	0.8	0.1	0.19	0.1	0.12	−0.81	−0.43	−0.34	−0.36	−0.46
Poppy Seeds	0.82	0.47	0.29	0.54	0.59	0.66	0.31	0.28	0.12	0.13	0.02	−0.02	−0.84	−0.43	−0.43	−0.29	−0.4
Pepper	0.99	0.47	0.1	0.58	0.7	0.9	0.18	0.14	0.23	0.32	0.08	0.04	−0.91	−0.61	−0.41	−0.21	−0.46
Turmeric	0.88	0.53	0.11	0.43	0.74	0.69	0.28	0.21	0.21	0.26	0.16	0.1	−0.91	−0.54	−0.49	−0.38	−0.47
Sugar	0.45	0.34	0.59	0.77	0.35	0.56	0.25	0.65	0.01	0.11	−0.09	0.06	−0.81	−0.46	−0.26	−0.31	−0.44
Vinegar	0.3	0.11	0.76	0.88	0.15	0.41	0.26	0.83	0.04	0.11	0.16	0.07	−0.65	−0.26	−0.31	−0.33	−0.36
Sesame Seeds	0.8	0.49	0.3	0.59	0.59	0.7	0.34	0.29	0.1	0.11	0.04	−0.01	−0.91	−0.49	−0.44	−0.33	−0.4
Lemon Juice	0.28	0.2	0.74	0.81	0.15	0.43	0.39	0.81	−0.05	0.15	0.19	0.07	−0.78	−0.3	−0.34	−0.46	−0.43
Chocolate	0.27	0.21	0.78	0.8	0.2	0.46	0.38	0.78	−0.01	0.19	0.16	−0.01	−0.81	−0.39	−0.36	−0.37	−0.44
Horseradish	0.61	0.67	0.48	0.28	0.61	0.4	0.53	0.33	0	0.12	0.04	0.04	−0.86	−0.4	−0.47	−0.37	−0.44
Vanilla	0.76	0.51	0.3	0.49	0.63	0.61	0.33	0.35	0.11	0.13	0.03	0.05	−0.91	−0.48	−0.44	−0.38	−0.48
Chives	0.66	0.89	0.43	0.26	0.76	0.28	0.64	0.31	0.1	0.02	0.21	0.06	−0.99	−0.38	−0.51	−0.53	−0.33
Root Ginger	0.84	0.56	0.23	0.44	0.69	0.59	0.41	0.23	0.13	0.14	0.18	−0.01	−0.91	−0.43	−0.54	−0.41	−0.37

A = Spices, B = Herbs.

$$= 1 - \mu(B) - \mu(A' \text{ and } B') - \mu(A) + \mu(A \text{ and } B)$$
$$+ \mu(A \text{ and } B)$$
$$= \mu(A' \text{ and } B') - \mu(A' \text{ and } B') + \mu(A \text{ and } B)$$
$$= \mu(A \text{ and } B)$$

which proves indeed that (8) can be derived from (5), (6) and then (7), and can be left out as a condition.

Let us now prove a result which is useful for our purposes. Following (5) and (6) we have that $\mu(A \text{ and } B') + \mu(A \text{ and } B) + \mu(A' \text{ and } B) = \mu(A) + \mu(B) - \mu(A \text{ and } B)$. Moreover, by using (7), we get $\mu(A \text{ and } B') + \mu(A \text{ and } B) + \mu(A' \text{ and } B) + \mu(A' \text{ and } B') = \mu(A) + \mu(B) - \mu(A \text{ and } B) + 1 - \mu(A) - \mu(B) + \mu(A \text{ and } B) = 1$.

The equality

$$\mu(A \text{ and } B') + \mu(A \text{ and } B) + \mu(A' \text{ and } B) + \mu(A' \text{ and } B') = 1 \tag{A11}$$

can be used, together with (5) and (6), as follows.

$$\mu(A) + \mu(A') = \mu(A \text{ and } B) + \mu(A \text{ and } B') + \mu(A' \text{ and } B)$$
$$+ \mu(A' \text{ and } B') = 1$$
$$\mu(B) + \mu(B') = \mu(A \text{ and } B) + \mu(A' \text{ and } B) + \mu(A \text{ and } B')$$
$$+ \mu(A' \text{ and } B') = 1$$

This means that from (1), and hence $0 \leq \mu(A) \leq 1$, follows that $0 \leq 1 - \mu(A) \leq 1$, and hence $0 \leq \mu(A') \leq 1$. And from (2),

and hence $0 \leq \mu(B) \leq 1$, follows that $0 \leq 1 - \mu(B) \leq 1$, and hence $0 \leq \mu(B') \leq 1$. Suppose now that (1) and (2) are satisfied. This means that $0 \leq \mu(A) - \mu(A \text{ and } B) = \mu(B') - \mu(A' \text{ and } B')$ and hence $\mu(A' \text{ and } B') \leq \mu(B')$, and $0 \leq \mu(B) - \mu(A \text{ and } B) = \mu(A') - \mu(A' \text{ and } B')$ and hence $\mu(A' \text{ and } B') \leq \mu(A')$. The only condition that lacks to have derived (3) and (4) from (1) and (2), is that $0 \leq \mu(A' \text{ and } B')$. We can add this as a requirement to (7).

Hence, we have proved the Theorem 2. □

Proof of Lemma 1. That (14), (15), (16) and (17) follow from (5) and (6) follows from a simple reshuffling of the terms. Suppose now that (14), (15), (16) and (17) are satisfied. The inverse reshuffling of the same terms proves that (5) and (6) are satisfied. This completes the proof of Lemma 1. □

Proof of Theorem 3. Lemma 1 entails that we can substitute (5) and (6) by the four equations expressing the marginal law to be satisfied.

A further simplification is possible. Indeed, (5), (6) and (7) are equivalent with (14), (15), (16), (17) and (A11). We have proved above that (5), (6) and (7) imply (A11). Let us prove the inverse. Hence suppose that (14), (15), (16), (17) and (A11) are satisfied, and let us proof (7). We have

$$1 - \mu(A) - \mu(B) + \mu(A \text{ and } B)$$
$$= 1 - \mu(A \text{ and } B) - \mu(A \text{ and } B') - \mu(A \text{ and } B) - \mu(A' \text{ and } B)$$
$$+ \mu(A \text{ and } B)$$

TABLE A3 | Representation of the membership weights in the case of the concepts *Pets* and *Farmyard Animals*.

Exemplar	$\mu(A)$	$\mu(B)$	$\mu(A')$	$\mu(B')$	$\mu(A \text{ and } B)$	$\mu(A \text{ and } B')$	$\mu(A' \text{ and } B)$	$\mu(A' \text{ and } B')$	Δ_{AB}	$\Delta_{AB'}$	$\Delta_{A'B}$	$\Delta_{A'B'}$	$I_{ABA'B'}$	I_A	I_B	$I_{A'}$	$I_{B'}$
Goldfish	0.93	0.17	0.12	0.81	0.43	0.91	0.18	0.43	0.26	0.1	0.06	0.31	−0.94	−0.41	−0.43	−0.48	−0.53
Robin	0.28	0.36	0.71	0.64	0.31	0.35	0.46	0.46	0.04	0.08	0.1	−0.18	−0.59	−0.39	−0.41	−0.22	−0.18
Blue-tit	0.25	0.31	0.76	0.71	0.18	0.39	0.44	0.56	−0.08	0.14	0.13	−0.15	−0.56	−0.31	−0.3	−0.24	−0.24
Collie Dog	0.95	0.77	0.03	0.35	0.86	0.56	0.25	0.11	0.09	0.21	0.23	0.09	−0.79	−0.48	−0.34	−0.34	−0.33
Camel	0.16	0.26	0.89	0.75	0.2	0.31	0.51	0.68	0.04	0.16	0.26	−0.08	−0.7	−0.36	−0.46	−0.3	−0.24
Squirrel	0.3	0.39	0.74	0.65	0.28	0.26	0.46	0.59	−0.03	−0.04	0.07	−0.06	−0.59	−0.24	−0.34	−0.31	−0.2
Guide Dog for Blind	0.93	0.33	0.13	0.69	0.55	0.73	0.16	0.33	0.23	0.03	0.04	0.2	−0.76	−0.35	−0.39	−0.36	−0.36
Spider	0.31	0.39	0.73	0.63	0.31	0.31	0.44	0.51	0	0	0.05	−0.12	−0.58	−0.31	−0.36	−0.23	−0.19
Homing Pigeon	0.41	0.71	0.61	0.34	0.56	0.25	0.59	0.34	0.16	−0.09	−0.03	0	−0.74	−0.41	−0.44	−0.31	−0.25
Monkey	0.39	0.18	0.65	0.79	0.2	0.49	0.29	0.61	0.03	0.09	0.11	−0.04	−0.59	−0.29	−0.31	−0.25	−0.31
Circus Horse	0.3	0.48	0.74	0.6	0.34	0.35	0.53	0.48	0.04	0.05	0.04	−0.13	−0.69	−0.39	−0.38	−0.26	−0.23
Prize Bull	0.13	0.76	0.88	0.26	0.43	0.28	0.83	0.34	0.29	0.14	0.06	0.08	−0.86	−0.57	−0.49	−0.28	−0.35
Rat	0.2	0.36	0.85	0.68	0.21	0.28	0.54	0.63	0.01	0.08	0.18	−0.05	−0.65	−0.29	−0.39	−0.31	−0.23
Badger	0.16	0.28	0.88	0.73	0.14	0.26	0.44	0.66	−0.03	0.1	0.16	−0.07	−0.5	−0.24	−0.3	−0.23	−0.19
Siamese Cat	0.99	0.5	0.05	0.53	0.74	0.75	0.18	0.24	0.24	0.23	0.13	0.19	−0.9	−0.5	−0.41	−0.36	−0.46
Race Horse	0.29	0.7	0.71	0.39	0.51	0.31	0.65	0.31	0.23	0.03	−0.05	−0.08	−0.79	−0.54	−0.46	−0.26	−0.24
Fox	0.13	0.3	0.86	0.68	0.18	0.29	0.46	0.59	0.04	0.16	0.16	−0.09	−0.51	−0.33	−0.34	−0.19	−0.19
Donkey	0.29	0.9	0.78	0.15	0.56	0.18	0.81	0.23	0.28	0.03	0.04	0.08	−0.78	−0.45	−0.48	−0.26	−0.25
Field Mouse	0.16	0.41	0.83	0.59	0.23	0.24	0.43	0.58	0.06	0.08	0.02	−0.01	−0.46	−0.3	−0.24	−0.18	−0.23
Ginger Tom-cat	0.82	0.51	0.21	0.54	0.59	0.58	0.26	0.29	0.08	0.03	0.05	0.08	−0.71	−0.34	−0.34	−0.34	−0.32
Husky in Slead team	0.64	0.51	0.37	0.53	0.56	0.51	0.44	0.29	0.06	−0.01	0.07	−0.08	−0.8	−0.43	−0.49	−0.36	−0.28
Cart Horse	0.27	0.86	0.76	0.15	0.53	0.2	0.84	0.23	0.26	0.05	0.08	0.08	−0.79	−0.46	−0.5	−0.31	−0.28
Chicken	0.23	0.95	0.8	0.06	0.58	0.11	0.81	0.18	0.34	0.05	0.01	0.11	−0.68	−0.46	−0.44	−0.19	−0.23
Doberman Guard Dog	0.88	0.76	0.14	0.27	0.8	0.55	0.45	0.23	0.04	0.28	0.31	0.09	−1.03	−0.47	−0.49	−0.54	−0.51

A = Pets, B = Farmyard Animals.

$$= 1 - \mu(A \text{ and } B) - \mu(A \text{ and } B') - \mu(A' \text{ and } B)$$

$$= \mu(A' \text{ and } B')$$

which proves that (7) holds.

We have thus proved Theorem 3, stating a new and more symmetric set of classicality conditions. □

Proof of Theorem 4. Suppose that the theoretical values for a modeling only in second sector of Fock space are given. This means that $|C\rangle$ and M are given, and the values of $\mu(A), \mu(B), \mu(A'), \mu(B')$, $\mu(A \text{ and } B), \mu(A \text{ and } B'), \mu(A' \text{ and } B)$ and $\mu(A' \text{ and } B')$ are given respectively by $\langle C|M \otimes \mathbb{1}|C\rangle, \langle C|\mathbb{1} \otimes M|C\rangle, \langle C|(\mathbb{1} - M) \otimes \mathbb{1}|C\rangle$, $\langle C|\mathbb{1} \otimes (\mathbb{1} - M)|C\rangle, \langle C|M \otimes M|C\rangle, \langle C|M \otimes (\mathbb{1} - M)|C\rangle, \langle C|(\mathbb{1} - M) \otimes M|C\rangle$ and $\langle C|(\mathbb{1} - M) \otimes (\mathbb{1} - M)|C\rangle$. If we use the results calculated in (76), (77), (78), (79), (80), (81), (82) and (83), we can easely prove all the classicality conditions (26)–(30) to be satisfied. Let us prove one of them explicitly. For example, $\mu(A' \text{ and } B) + \mu(A' \text{ and } B') = \langle C|(\mathbb{1} - M) \otimes M|C\rangle + \langle C|(\mathbb{1} - M) \otimes (\mathbb{1} - M)|C\rangle = \langle C|(\mathbb{1} - M) \otimes M + (\mathbb{1} - M)|C\rangle = \langle C|(\mathbb{1} - M) \otimes \mathbb{1}|C\rangle = \mu(A')$.

Let us prove the other implication. Hence suppose that we have available data satisfying the classicality conditions (26)–(30), let us prove that we can find a state $|C\rangle$ and an orthogonal projector M, such that second sector models these data. It is a straightforward verification that an entangled vector $|C\rangle$, such that $c_{ij} = \frac{1}{16}\sqrt{\mu(A \text{ and } B)}$ for $5 \leq i \leq 8$ and $5 \leq j \leq 8$, $c_{ij} = \frac{1}{16}\sqrt{\mu(A \text{ and } B')}$ for $5 \leq i \leq 8$ and $1 \leq j \leq 4$, $c_{ij} = \frac{1}{16}\sqrt{\mu(A' \text{ and } B)}$ for $1 \leq i \leq 4$ and $5 \leq j \leq 8$ and $c_{ij} = \frac{1}{16}\sqrt{\mu(A' \text{ and } B')}$ for $1 \leq i \leq 4$ and $1 \leq j \leq 4$, is a solution. □

A5. Conditions for the Existence of Solutions in First and Second Sector of Fock Space

In this section we make explicit the conditions that should be satisfied by experimental data in order to be represented in Fock space. In our analysis, we distinguish between the first sector representation in $\mathbb{C}^8$ and the complete two-sector representation in $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$.

Let us start from the $\mathbb{C}^8$ representation. We first analyze whether or not the solution of (52)–(65) is compatible with (70)–(73). To this end note that the right hand side of (70)–(73) correspond to the average of the probabilities of the former concepts, plus the so called "interference term," which depends on (i) how the

TABLE A4 | Representation of the membership weights in the case of the concepts *Fruits* and *Vegetables*.

Exemplar	$\mu(A)$	$\mu(B)$	$\mu(A')$	$\mu(B')$	$\mu(A$ and $B)$	$\mu(A$ and $B')$	$\mu(A'$ and $B)$	$\mu(A'$ and $B')$	Δ_{AB}	$\Delta_{AB'}$	$\Delta_{A'B}$	$\Delta_{A'B'}$	$I_{ABA'B'}$	I_A	I_B	$I_{A'}$	$I_{B'}$
Apple	1	0.23	0	0.82	0.6	0.89	0.13	0.18	0.38	0.07	0.13	0.18	−0.79	−0.49	−0.5	−0.3	−0.24
Parsley	0.02	0.78	0.99	0.25	0.45	0.1	0.84	0.44	0.43	0.08	0.06	0.19	−0.83	−0.53	−0.51	−0.29	−0.29
Olive	0.53	0.63	0.47	0.44	0.65	0.34	0.51	0.36	0.12	−0.11	0.04	−0.08	−0.86	−0.46	−0.53	−0.41	−0.26
Chili Pepper	0.19	0.73	0.83	0.35	0.51	0.2	0.68	0.44	0.33	0.01	−0.06	0.09	−0.83	−0.53	−0.46	−0.29	−0.29
Broccoli	0.09	1	0.94	0.06	0.59	0.09	0.9	0.25	0.49	0.03	−0.04	0.19	−0.83	−0.58	−0.49	−0.21	−0.28
Root Ginger	0.14	0.71	0.81	0.33	0.46	0.14	0.71	0.43	0.33	0	0	0.1	−0.74	−0.46	−0.46	−0.33	−0.24
Pumpkin	0.45	0.78	0.51	0.26	0.66	0.21	0.63	0.18	0.21	−0.05	0.11	−0.09	−0.68	−0.43	−0.51	−0.29	−0.13
Raisin	0.88	0.27	0.13	0.76	0.53	0.75	0.25	0.34	0.26	−0.01	0.12	0.21	−0.86	−0.39	−0.51	−0.46	−0.33
Acorn	0.59	0.4	0.49	0.64	0.46	0.49	0.38	0.51	0.06	−0.1	−0.03	0.02	−0.84	−0.36	−0.44	−0.39	−0.36
Mustard	0.07	0.39	0.87	0.6	0.29	0.23	0.55	0.75	0.22	0.16	0.16	0.15	−0.81	−0.44	−0.45	−0.43	−0.38
Rice	0.12	0.46	0.9	0.52	0.21	0.23	0.59	0.59	0.09	0.11	0.13	0.07	−0.61	−0.32	−0.34	−0.28	−0.29
Tomato	0.34	0.89	0.64	0.19	0.7	0.2	0.74	0.23	0.36	0.01	0.1	0.04	−0.86	−0.56	−0.55	−0.33	−0.24
Coconut	0.93	0.32	0.17	0.7	0.56	0.69	0.2	0.34	0.24	−0.01	0.03	0.17	−0.79	−0.33	−0.44	−0.37	−0.33
Mushroom	0.12	0.66	0.9	0.38	0.33	0.13	0.66	0.5	0.21	0.01	0	0.12	−0.61	−0.33	−0.33	−0.26	−0.24
Wheat	0.17	0.51	0.8	0.52	0.34	0.21	0.61	0.56	0.17	0.04	0.11	0.04	−0.73	−0.38	−0.44	−0.38	−0.26
Green Pepper	0.23	0.61	0.81	0.41	0.49	0.24	0.61	0.43	0.26	0.01	0	0.02	−0.76	−0.5	−0.49	−0.23	−0.26
Watercress	0.14	0.76	0.89	0.25	0.49	0.1	0.79	0.35	0.35	−0.04	0.03	0.1	−0.73	−0.45	−0.51	−0.24	−0.2
Peanut	0.62	0.29	0.48	0.75	0.48	0.55	0.25	0.53	0.18	−0.07	−0.04	0.05	−0.8	−0.41	−0.43	−0.3	−0.33
Black Pepper	0.21	0.41	0.81	0.61	0.38	0.21	0.5	0.63	0.17	0.01	0.09	0.01	−0.71	−0.38	−0.46	−0.31	−0.23
Garlic	0.13	0.79	0.88	0.24	0.53	0.1	0.75	0.45	0.4	−0.03	−0.04	0.21	−0.83	−0.5	−0.49	−0.33	−0.31
Yam	0.38	0.66	0.71	0.43	0.59	0.24	0.65	0.44	0.21	−0.14	−0.01	0.01	−0.91	−0.45	−0.58	−0.38	−0.24
Elderberry	0.51	0.39	0.54	0.61	0.45	0.41	0.46	0.48	0.06	−0.09	0.07	−0.07	−0.8	−0.36	−0.52	−0.39	−0.28
Almond	0.76	0.29	0.28	0.72	0.48	0.61	0.24	0.48	0.18	−0.11	−0.04	0.19	−0.8	−0.33	−0.42	−0.43	−0.37
Lentils	0.11	0.66	0.89	0.38	0.38	0.11	0.7	0.53	0.26	0	0.04	0.15	−0.71	−0.38	−0.41	−0.33	−0.26

A = Fruits, B = Vegetables.

vectors representing the former concepts in the combination, when restricted to the subspace determined by M, project into each other, and (ii) on the phase angles of the vectors (Aerts, 2009). Note that the configuration of the phase angles $\phi_A, \phi_B, \phi_{A'}, \phi_{B'}$ allow to model a variety of interference situations. In fact, when the difference between these phase angles is close to 0 or π, we have a maximal amount of interference while, when the difference between these phase angles is close to $\frac{\pi}{2}$ or $\frac{3\pi}{2}$, we have a minimal amount of interference. We can then characterize a set of "compatibility intervals" for the solution of (52)–(65) and (70)–(73) which determines the modeling capacity of this Hilbert space model. Let,

$$i(A, B)_1 = \frac{1}{2}(\mu(A) + \mu(B)) - |a_5 b_5 + a_6 b_6 \\ + a_7 b_7 + a_8 b_8| \tag{A12}$$

$$i(A, B)_2 = \frac{1}{2}(\mu(A) + \mu(B)) + |a_5 b_5 + a_6 b_6 \\ + a_7 b_7 + a_8 b_8| \tag{A13}$$

$$i(A, B')_1 = \frac{1}{2}(\mu(A) + \mu(B')) - |a_5 b'_5 + a_6 b'_6 \\ + a_7 b'_7 + a_8 b'_8| \tag{A14}$$

$$i(A, B')_2 = \frac{1}{2}(\mu(A) + \mu(B')) + |a_5 b'_5 + a_6 b'_6 \\ + a_7 b'_7 + a_8 b'_8| \tag{A15}$$

$$i(A', B)_1 = \frac{1}{2}(\mu(A') + \mu(B)) - |a'_5 b_5 + a'_6 b_6 \\ + a'_7 b_7 + a'_8 b_8| \tag{A16}$$

$$i(A', B)_2 = \frac{1}{2}(\mu(A') + \mu(B)) + |a'_5 b_5 + a'_6 b_6 \\ + a'_7 b_7 + a'_8 b_8| \tag{A17}$$

$$i(A', B')_1 = \frac{1}{2}(\mu(A') + \mu(B')) - |a'_5 b'_5 + a'_6 b'_6 \\ + a'_7 b'_7 + a'_8 b'_8| \tag{A18}$$

$$i(A', B')_2 = \frac{1}{2}(\mu(A') + \mu(B')) + |a'_5 b'_5 + a'_6 b'_6 \\ + a'_7 b'_7 + a'_8 b'_8| \tag{A19}$$

and let us define the following "solution intervals"

$$\mathscr{I}_{AB} = [i(A, B)_1, i(A, B)_2] \tag{A20}$$

$$\mathscr{I}_{AB'} = [i(A, B')_1, i(A, B')_2] \tag{A21}$$

$$\mathscr{I}_{A'B} = [i(A', B)_1, i(A', B)_2] \tag{A22}$$

$$\mathscr{I}_{A'B'} = [i(A', B')_1, i(A', B')_2] \tag{A23}$$

A solution of (52)-(73) exists if and only if the membership weights $\mu(A$ and $B)$, $\mu(A'$ and $B)$, $\mu(A$ and $B')$ and $\mu(A$ and $B)$ are respectively contained in the intervals $\mathscr{I}_{AB}$, $\mathscr{I}_{A'B}$, $\mathscr{I}_{AB'}$ and $\mathscr{I}_{A'B'}$.

TABLE A5A | Calculation of the p-values corresponding to the deviation I_A between $\mu(A)$ and $\mu(A$ and $B) + \mu(A$ and $B')$.

Deviation of $\mu(A)$ from $\mu(A$ and $B) + \mu(A$ and $B')$

(Home Furnishing, Furniture)	p-value	(Spices, Herbs)	p-value	(Pets, Farmyard Animals)	p-value	(Fruits, Vegetables)	p-value
Mantelpiece	5.61E-09	Molasses	9.73E-07	Goldfish	2.52E-05	Apple	1.78E-08
Window Seat	1.08E-05	Salt	7.01E-04	Robin	1.78E-06	Parsley	5.14E-07
Painting	4.19E-07	Peppermint	5.57E-06	Blue-tit	3.56E-06	Olive	1.98E-05
Light Fixture	4.20E-06	Curry	6.51E-05	Collie Dog	7.90E-06	Chili Pepper	1.85E-07
Kitchen Counter	6.10E-05	Oregano	1.78E-06	Camel	5.92E-05	Broccoli	4.35E-09
Bath Tub	4.17E-06	MSG	5.93E-05	Squirrel	5.23E-04	Root Ginger	1.44E-06
Deck Chair	5.35E-06	Chili Pepper	4.00E-12	Guide Dog for Blind	6.72E-04	Pumpkin	9.35E-06
Shelves	2.20E-06	Mustard	2.26E-05	Spider	4.19E-05	Raisin	1.13E-06
Rug	1.81E-09	Mint	1.56E-05	Homing Pigeon	4.87E-05	Acorn	1.04E-05
Bed	5.81E-07	Cinnamon	8.90E-08	Monkey	7.21E-04	Mustard	6.05E-07
Wall-Hangings	7.75E-07	Parsley	2.05E-05	Circus Horse	3.71E-07	Rice	7.87E-05
Space Rack	2.02E-08	Saccarin	1.64E-06	Prize Bull	2.02E-08	Tomato	1.77E-07
Ashtray	2.73E-06	Poppy Seeds	5.63E-05	Rat	1.05E-03	Coconut	1.61E-03
Bar	3.07E-08	Pepper	2.70E-07	Badger	3.79E-04	Mushroom	3.40E-05
Lamp	4.80E-08	Turmeric	1.62E-08	Siamese Cat	4.20E-06	Wheat	3.76E-06
Wall Mirror	1.95E-10	Sugar	1.52E-07	Race Horse	1.47E-07	Green Pepper	3.96E-07
Door Bell	5.94E-07	Vinegar	5.93E-04	Fox	2.26E-05	Watercress	2.60E-07
Hammock	2.35E-06	Sesame Seeds	5.49E-07	Donkey	1.79E-06	Peanut	7.62E-05
Desk	2.94E-05	Lemon Juice	2.79E-05	Field Mouse	1.20E-05	Black Pepper	6.54E-06
Refrigerator	2.41E-07	Chocolate	3.06E-06	Ginger Tom-cat	9.79E-06	Garlic	2.38E-07
Park Bench	1.09E-06	Horseradish	1.55E-06	Husky in Slead team	1.56E-05	Yam	3.68E-08
Waste Paper Basket	2.41E-06	Vanilla	4.28E-07	Cart Horse	2.62E-08	Elderberry	8.82E-05
Sculpture	1.43E-06	Chives	7.42E-06	Chicken	8.62E-08	Almond	1.04E-04
Sink Unit	3.97E-07	Root Ginger	5.02E-06	Doberman Guard Dog	1.29E-04	Lentils	9.59E-07

By applying a Bonferroni correction procedure, the null hypothesis can be rejected for a p-value less than $\frac{0.05}{24} = 2.08 \cdot 10^{-3}$.

Let us now come to the complete representation in $\mathbb{C}^8 \oplus (\mathbb{C}^8 \otimes \mathbb{C}^8)$, and let us consider the data collected in the experiments e_{XY}, $X = A, A'$, $Y = B, B'$, of Section 2. These data are explicitly reported in **Tables A1–A4**. Since the existence of solutions for (86)–(92) depends, for a given experiment e_{XY}, on the exemplar x and the pair (A, B) of concepts that are considered, we explicitly report such dependence for all the relevant variables that are considered in this section. Hence, for each considered exemplar x, we collected the eight membership weights $\mu_x(A)$, $\mu_x(B)$, $\mu_x(A')$, $\mu_x(B')$, $\mu_x(A$ and $B)$, $\mu_x(A$ and $B')$, $\mu_x(A'$ and $B)$, and $\mu_x(A'$ and $B')$.

The analysis we made in the foregoing sections makes it possible for us to propose a general modeling procedure. For what concerns solutions that can be found on first sector alone, we determined the intervals of solutions as explained in (A20), (A21), (A22) and (A23). We can now easily determine the general intervals of solutions, including the extra solutions made possible by second sector. Therefore, we need to consider the following quantities $\sum_{i,j=5}^{8} c_{ij}^2$, $\sum_{i=5}^{8} \sum_{j=1}^{4} c_{ij}^2$, $\sum_{i=1}^{4} \sum_{j=5}^{8} c_{ij}^2$ and $\sum_{i=5}^{8} \sum_{j=1}^{4} c_{ij}^2$, respectively for the combinations "A and B," "A and not B," "not A and B" and "not A and not B". To be able to express the intervals of Fock space solutions, we introduce the following quantities.

$$s(A, B, x) = \min\left(\sum_{i,j=5}^{8} c_{ij}^2, i(A, B)_1 \right) \quad (A24)$$

$$t(A, B, x) = \max\left(\sum_{i,j=5}^{8} c_{ij}^2, i(A, B)_2 \right) \quad (A25)$$

Then, the interval

$$U_{sol}(AB, x) = [s(A, B, x), t(A, B, x)] \quad (A26)$$

is the solution interval for the general Fock space model. Hence, in case the experimental value $\mu(A$ and $B)$ is contained in this interval, a solution exists. As a second step we can then see whether a solution in first sector alone exists, which consists of veryfing whether the experimental value $\mu(A$ and $B)$ is contained in I_{AB}. Suppose that the anwer is "yes," then we can caculate the angle $\phi_B - \phi_A$ that gives rise to this solution in first sector. This angle is then an indication of which angle to choose for the general Fock space solution. Usually different choices are possible. If there is no solution in first sector, we anyhow can choose an angle $\phi_B - \phi_A$, such that a choice of this angle $\phi_B - \phi_A$, and a choice of m_{AB} and n_{AB} gives a solution. The possible values of the angle and the coefficients m_{AB} and n_{AB} are calculated by solving (86).

We can analyze the other combinations in an equivalent way. Let us start with the combination "A and not B". We have:

$$s(A, B', x) = \min\left(\sum_{i=5}^{8} \sum_{j=1}^{4} c_{ij}^2, i(A, B')_1 \right) \quad (A27)$$

TABLE A5B | Calculation of the *p*-values corresponding to the deviation I_B between $\mu(B)$ and $\mu(A$ and $B) + \mu(A'$ and $B)$.

Deviation of $\mu(B)$ from $\mu(A$ and $B) + \mu(A'$ and $B)$

(Home Furnishing, Furniture)	*p*-value	(Spices, Herbs)	*p*-value	(Pets, Farmyard Animals)	*p*-value	(Fruits, Vegetables)	*p*-value
Mantelpiece	1.09E-05	Molasses	4.21E-06	Goldfish	1.89E-07	Apple	9.00E-07
Window Seat	1.89E-05	Salt	9.37E-05	Robin	8.27E-07	Parsley	9.14E-08
Painting	2.20E-07	Peppermint	2.50E-04	Blue-tit	2.90E-06	Olive	6.81E-08
Light Fixture	2.99E-06	Curry	8.55E-06	Collie Dog	2.00E-06	Chili Pepper	1.60E-07
Kitchen Counter	5.12E-06	Oregano	1.25E-06	Camel	1.35E-07	Broccoli	1.22E-06
Bath Tub	1.43E-06	MSG	1.79E-06	Squirrel	5.85E-06	Root Ginger	7.24E-06
Deck Chair	1.07E-04	Chili Pepper	2.20E-07	Guide Dog for Blind	4.55E-06	Pumpkin	1.17E-07
Shelves	7.84E-07	Mustard	5.57E-06	Spider	6.15E-05	Raisin	4.75E-08
Rug	4.99E-09	Mint	1.80E-06	Homing Pigeon	1.18E-06	Acorn	9.32E-08
Bed	2.73E-06	Cinnamon	5.03E-06	Monkey	2.95E-05	Mustard	4.61E-08
Wall-Hangings	3.93E-06	Parsley	7.52E-06	Circus Horse	2.27E-05	Rice	9.76E-06
Space Rack	2.16E-08	Saccarin	2.63E-06	Prize Bull	2.17E-06	Tomato	4.63E-07
Ashtray	8.83E-07	Poppy Seeds	3.05E-06	Rat	4.08E-06	Coconut	3.58E-06
Bar	3.07E-05	Pepper	4.70E-06	Badger	8.59E-05	Mushroom	1.39E-04
Lamp	8.94E-07	Turmeric	2.06E-07	Siamese Cat	1.98E-05	Wheat	5.03E-07
Wall Mirror	1.05E-06	Sugar	2.45E-06	Race Horse	5.03E-06	Green Pepper	8.15E-07
Door Bell	6.27E-07	Vinegar	1.89E-05	Fox	9.66E-05	Watercress	1.26E-07
Hammock	4.82E-06	Sesame Seeds	6.89E-07	Donkey	2.41E-06	Peanut	3.58E-06
Desk	5.97E-06	Lemon Juice	1.64E-06	Field Mouse	1.29E-03	Black Pepper	3.97E-07
Refrigerator	3.82E-05	Chocolate	9.12E-06	Ginger Tom-cat	1.64E-06	Garlic	1.15E-05
Park Bench	2.16E-08	Horseradish	1.31E-05	Husky in Slead team	1.01E-07	Yam	1.62E-09
Waste Paper Basket	7.21E-08	Vanilla	1.22E-07	Cart Horse	2.26E-07	Elderberry	1.98E-08
Sculpture	4.46E-08	Chives	3.81E-07	Chicken	4.98E-05	Almond	2.17E-06
Sink Unit	3.64E-06	Root Ginger	1.02E-08	Doberman Guard Dog	2.40E-06	Lentils	7.69E-06

By applying a Bonferroni correction procedure, the null hypothesis can be rejected for a p-value less than $\frac{0.05}{24} = 2.08 \cdot 10^{-3}$.

$$t(A, B', x) = \max\left(\sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2, i(A, B')_2\right) \quad (A28)$$

Then, the interval

$$U_{sol}(AB', x) = [s(A, B', x), t(A, B', x)] \quad (A29)$$

is the solution interval for the general Fock space model. The equation to be used to calculate the angle $\phi_{B'} - \phi_A$, and the coefficients $m_{AB'}$ and $n_{AB'}$ is (88).

For the combination "A' and B," we have:

$$s(A', B, x) = \min\left(\sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2, i(A', B)_1\right) \quad (A30)$$

$$t(A', B, x) = \max\left(\sum_{i=1}^{4}\sum_{j=5}^{8} c_{ij}^2, i(A', B)_2\right) \quad (A31)$$

Then, the interval

$$U_{sol}(A'B, x) = [s(A', B, x), t(A', B, x)] \quad (A32)$$

is the solution interval for the general Fock space model. The equation to be used to calculate the angle $\phi_B - \phi_{A'}$, and the coefficients $m_{A'B}$ and $n_{A'B}$ is (90).

Finally, for the combination "not A and not B," we have:

$$s(A', B', x) = \min\left(\sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2, i(A', B')_1\right) \quad (A33)$$

$$t(A, B', x) = \max\left(\sum_{i=5}^{8}\sum_{j=1}^{4} c_{ij}^2, i(A', B')_2\right) \quad (A34)$$

Then, the interval

$$U_{sol}(A'B', x) = [s(A', B', x), t(A', B', x)] \quad (A35)$$

is the solution interval for the general Fock space model. The formula to be used to calculate the angle $\phi_{B'} - \phi_{A'}$, and the coefficients $m_{A'B'}$ and $n_{A'B'}$ is Equation (92).

TABLE A5C | Calculation of the *p*-values corresponding to the deviation $I_{A'}$ between $\mu(A')$ and $\mu(A'$ and $B) + \mu(A'$ and $B')$.

Deviation of $\mu(A')$ from $\mu(A'$ and $B) + \mu(A'$ and $B')$

(Home Furnishing, Furniture)	*p*-value	(Spices, Herbs)	*p*-value	(Pets, Farmyard Animals)	*p*-value	(Fruits, Vegetables)	*p*-value
Mantelpiece	2.38E-05	Molasses	6.35E-05	Goldfish	1.51E-05	Apple	9.93E-05
Window Seat	4.88E-04	Salt	1.63E-04	Robin	2.85E-03	Parsley	3.03E-05
Painting	6.06E-05	Peppermint	3.01E-03	Blue-tit	4.54E-04	Olive	1.33E-06
Light Fixture	5.64E-04	Curry	7.06E-04	Collie Dog	1.39E-05	Chili Pepper	3.47E-05
Kitchen Counter	1.55E-05	Oregano	1.33E-04	Camel	5.17E-04	Broccoli	7.60E-03
Bath Tub	9.61E-05	MSG	2.91E-05	Squirrel	5.10E-05	Root Ginger	1.71E-04
Deck Chair	2.96E-04	Chili Pepper	8.75E-05	Guide Dog for Blind	2.75E-05	Pumpkin	2.55E-04
Shelves	6.06E-05	Mustard	2.14E-03	Spider	1.06E-02	Raisin	7.20E-06
Rug	1.09E-05	Mint	2.62E-03	Homing Pigeon	8.10E-04	Acorn	1.68E-05
Bed	3.00E-05	Cinnamon	3.65E-04	Monkey	5.59E-03	Mustard	1.75E-06
Wall-Hangings	1.12E-04	Parsley	1.33E-04	Circus Horse	5.80E-04	Rice	4.90E-03
Space Rack	3.28E-06	Saccarin	6.42E-06	Prize Bull	3.59E-04	Tomato	4.41E-04
Ashtray	1.55E-05	Poppy Seeds	2.05E-03	Rat	1.29E-03	Coconut	1.86E-04
Bar	2.34E-05	Pepper	7.60E-03	Badger	2.48E-02	Mushroom	2.62E-03
Lamp	2.26E-05	Turmeric	5.96E-04	Siamese Cat	2.71E-04	Wheat	2.76E-05
Wall Mirror	5.58E-06	Sugar	2.81E-04	Race Horse	1.28E-03	Green Pepper	2.75E-02
Door Bell	1.64E-05	Vinegar	5.31E-05	Fox	9.46E-02	Watercress	1.41E-03
Hammock	2.04E-04	Sesame Seeds	1.17E-03	Donkey	6.94E-03	Peanut	1.20E-05
Desk	1.36E-05	Lemon Juice	1.35E-07	Field Mouse	2.22E-02	Black Pepper	1.10E-03
Refrigerator	2.55E-05	Chocolate	3.03E-05	Ginger Tom-cat	2.79E-04	Garlic	6.14E-05
Park Bench	2.93E-05	Horseradish	2.80E-06	Husky in Slead team	3.22E-05	Yam	1.13E-06
Waste Paper Basket	7.87E-09	Vanilla	2.07E-06	Cart Horse	2.98E-04	Elderberry	1.58E-05
Sculpture	2.41E-06	Chives	1.08E-06	Chicken	2.66E-02	Almond	6.63E-06
Sink Unit	7.90E-06	Root Ginger	7.32E-04	Doberman Guard Dog	2.64E-07	Lentils	6.54E-05

By applying a Bonferroni correction procedure, the null hypothesis can be rejected for a p-value less than $\frac{0.05}{24} = 2.08 \cdot 10^{-3}$.

TABLE A5D | Calculation of the *p*-values corresponding to the deviation $I_{B'}$ between $\mu(B')$ and $\mu(A$ and $B') + \mu(A'$ and $B')$.

Deviation of $\mu(B')$ from $\mu(A$ and $B') + \mu(A'$ and $B')$

(Home Furnishing, Furniture)	*p*-value	(Spices, Herbs)	*p*-value	(Pets, Farmyard Animals)	*p*-value	(Fruits, Vegetables)	*p*-value
Mantelpiece	1.09E-06	Molasses	3.14E-07	Goldfish	2.89E-07	Apple	1.46E-05
Window Seat	8.03E-05	Salt	1.56E-05	Robin	2.23E-02	Parsley	5.26E-05
Painting	5.38E-05	Peppermint	1.56E-05	Blue-tit	1.63E-04	Olive	4.77E-04
Light Fixture	1.16E-03	Curry	1.83E-04	Collie Dog	2.06E-04	Chili Pepper	5.16E-05
Kitchen Counter	3.02E-03	Oregano	1.11E-03	Camel	5.93E-03	Broccoli	1.83E-04
Bath Tub	4.21E-06	MSG	2.60E-07	Squirrel	2.64E-03	Root Ginger	2.14E-03
Deck Chair	4.45E-06	Chili Pepper	1.71E-07	Guide Dog for Blind	5.83E-04	Pumpkin	5.09E-03
Shelves	5.06E-04	Mustard	4.08E-06	Spider	2.57E-02	Raisin	2.38E-05
Rug	1.17E-05	Mint	1.13E-04	Homing Pigeon	9.81E-03	Acorn	1.37E-05
Bed	4.44E-05	Cinnamon	5.03E-07	Monkey	9.47E-03	Mustard	3.07E-05
Wall-Hangings	7.21E-04	Parsley	6.14E-05	Circus Horse	1.87E-03	Rice	6.15E-04
Space Rack	1.38E-06	Saccarin	6.59E-08	Prize Bull	4.22E-04	Tomato	5.09E-04
Ashtray	5.03E-06	Poppy Seeds	1.87E-04	Rat	1.14E-02	Coconut	2.81E-04
Bar	1.74E-08	Pepper	3.52E-04	Badger	1.18E-02	Mushroom	4.84E-04
Lamp	2.24E-07	Turmeric	1.32E-06	Siamese Cat	3.69E-05	Wheat	1.45E-04
Wall Mirror	1.52E-04	Sugar	2.60E-06	Race Horse	7.53E-03	Green Pepper	4.63E-03
Door Bell	5.64E-06	Vinegar	1.50E-05	Fox	1.97E-02	Watercress	1.68E-03
Hammock	2.32E-04	Sesame Seeds	3.41E-05	Donkey	4.90E-03	Peanut	1.16E-04
Desk	6.37E-05	Lemon Juice	1.16E-06	Field Mouse	1.18E-02	Black Pepper	3.76E-03
Refrigerator	1.71E-05	Chocolate	8.26E-08	Ginger Tom-cat	1.20E-03	Garlic	4.62E-05
Park Bench	1.04E-04	Horseradish	7.46E-08	Husky in Slead team	2.62E-03	Yam	3.20E-05
Waste Paper Basket	2.88E-06	Vanilla	1.36E-06	Cart Horse	3.42E-04	Elderberry	3.72E-04
Sculpture	1.28E-04	Chives	6.15E-05	Chicken	1.92E-03	Almond	2.73E-06
Sink Unit	1.89E-04	Root Ginger	6.54E-05	Doberman Guard Dog	1.15E-05	Lentils	6.94E-05

By applying a Bonferroni correction procedure, the null hypothesis can be rejected for a p-value less than $\frac{0.05}{24} = 2.08 \cdot 10^{-3}$.

TABLE A5E | Calculation of the p-values corresponding to the deviation $I_{ABA'B'}$ between $\mu(A$ and $B) + \mu(A$ and $B') + \mu(A'$ and $B) + \mu(A'$ and $B')$ and 1.

Deviation of $\mu(A$ and $B) + \mu(A$ and $B') + \mu(A'$ and $B) + \mu(A'$ and $B')$ from 1

(Home Furnishing, Furniture)	p-value	(Spices, Herbs)	p-value	(Pets, Farmyard Animals)	p-value	(Fruits, Vegetables)	p-value
Mantelpiece	4.34E-09	Molasses	4.33E-07	Goldfish	3.98E-09	Apple	1.24E-08
Window Seat	5.12E-06	Salt	4.04E-05	Robin	1.09E-05	Parsley	1.03E-08
Painting	1.07E-07	Peppermint	7.51E-07	Blue-tit	3.54E-07	Olive	1.44E-08
Light Fixture	5.01E-07	Curry	3.31E-06	Collie Dog	1.55E-07	Chili Pepper	2.81E-09
Kitchen Counter	4.63E-06	Oregano	1.95E-07	Camel	1.19E-05	Broccoli	6.15E-09
Bath Tub	1.13E-07	MSG	5.67E-07	Squirrel	1.99E-05	Root Ginger	1.79E-06
Deck Chair	3.04E-07	Chili Pepper	2.19E-10	Guide Dog for Blind	2.43E-05	Pumpkin	6.60E-07
Shelves	7.84E-08	Mustard	5.67E-07	Spider	5.24E-05	Raisin	7.37E-09
Rug	8.18E-09	Mint	9.21E-08	Homing Pigeon	5.23E-06	Acorn	8.26E-09
Bed	3.69E-08	Cinnamon	2.42E-08	Monkey	5.05E-04	Mustard	9.61E-08
Wall-Hangings	1.09E-06	Parsley	2.61E-07	Circus Horse	7.10E-07	Rice	3.68E-06
Space Rack	3.77E-08	Saccarin	1.20E-07	Prize Bull	1.27E-08	Tomato	1.82E-09
Ashtray	9.08E-08	Poppy Seeds	1.24E-06	Rat	5.07E-05	Coconut	6.54E-07
Bar	2.27E-09	Pepper	8.33E-07	Badger	3.01E-04	Mushroom	5.77E-06
Lamp	1.87E-08	Turmeric	5.34E-08	Siamese Cat	1.34E-07	Wheat	9.21E-08
Wall Mirror	2.20E-09	Sugar	5.03E-07	Race Horse	5.77E-08	Green Pepper	6.54E-07
Door Bell	1.62E-07	Vinegar	3.40E-05	Fox	9.66E-04	Watercress	6.57E-08
Hammock	7.17E-07	Sesame Seeds	1.07E-07	Donkey	2.38E-06	Peanut	3.51E-07
Desk	7.94E-07	Lemon Juice	4.30E-07	Field Mouse	8.23E-04	Black Pepper	9.33E-07
Refrigerator	5.49E-07	Chocolate	5.18E-08	Ginger Tom-cat	6.79E-07	Garlic	2.51E-07
Park Bench	1.39E-08	Horseradish	5.03E-08	Husky in Slead team	3.99E-08	Yam	1.60E-10
Waste Paper Basket	1.38E-09	Vanilla	6.49E-08	Cart Horse	8.26E-09	Elderberry	1.60E-07
Sculpture	7.78E-09	Chives	1.80E-08	Chicken	1.53E-06	Almond	9.09E-08
Sink Unit	2.66E-07	Root Ginger	6.10E-08	Doberman Guard Dog	3.86E-08	Lentils	3.47E-07

By applying a Bonferroni correction procedure, the null hypothesis can be rejected for a p-value less than $\frac{0.05}{24} = 2.08 \cdot 10^{-3}$.

Reconstruction of a Real World Social Network Using the Potts Model and Loopy Belief Propagation

Cristian Bisconti[1], Angelo Corallo[1], Laura Fortunato[1], Antonio A. Gentile[1,2]*, Andrea Massafra[1] and Piergiuseppe Pellè[1,3]

[1] *CoSSNA Group, cPDM Lab, Department for Innovation Engineering, University of Salento, Lecce, Italy,* [2] *EKA srl, Lecce, Italy,* [3] *Advantech srl, Lecce, Italy*

The scope of this paper is to test the adoption of a statistical model derived from Condensed Matter Physics, for the reconstruction of the structure of a social network. The inverse Potts model, traditionally applied to recursive observations of quantum states in an ensemble of particles, is here addressed to observations of the members' states in an organization and their (anti)correlations, thus inferring interactions as links among the members. Adopting proper (Bethe) approximations, such an inverse problem is showed to be tractable. Within an operational framework, this network-reconstruction method is tested for a small real-world social network, the Italian parliament. In this study case, it is easy to track statuses of the parliament members, using (co)sponsorships of law proposals as the initial dataset. In previous studies of similar activity-based networks, the graph structure was inferred directly from activity co-occurrences: here we compare our statistical reconstruction with such standard methods, outlining discrepancies and advantages.

Keywords: social network analysis, Potts model, network reconstruction, community detection, loopy belief propagation, inverse problem, quantum structures

1. INTRODUCTION

A growing interest raised in recent years about policy networks in social and organizational studies: the concept has flourished even in the absence of a widely agreed definition. Among the most successful ones, we may quote (Börzel, 1997) and the concept of horizontal networks linking a variety of actors, who share common interests about a policy, and cooperate toward its adoption. Now, such a broad idea withstood critiques considering the policy network a mere metaphor, more than a model capable of understanding the process of genesis and evolution of policies (Dowding, 1995). A rich literature has adopted both qualitative and quantitative methods to analyse the *network* paradigm. In fact, in most study cases, the relations between the actors involved are depicted as links between the corresponding nodes of a graph (the actors). Most discussions are also driven by network analysis tools and methods (Besussi, 2006).

Among quantitative methods, for our case study we focused on the collaborative nature of policy networks, dealing with *vote behavior*, an idea originally dating back to the "socio-structural and interactional effects," investigated since (Lazarsfeld et al., 1968). Here, however, following a recent but well developed approach, sponsorships and endorsements of law *proposals* are tracked in the dataset, rather than proper voting behavior when these proposals are approved or rejected. Social Network Analysis (SNA) performed with co-sponsorships and other

similar data in legislative bodies was started by Fowler (2006), where the network structure and proximity measures among the US Senate members were obtained. Further analyses were also based on roll calls, but they focused upon the creation and evolution of communities in the network (Porter et al., 2007; Zhang et al., 2008). Retrieving the policy network in a legislative body, and communities therein, using roll-calls and co-sponsorship data, instead of final votes, is a method embedding both advantages and disadvantages. A useful discussion about the point can be found in Chiru and Neamu (2012).

Main interest of this paper is the problem of preliminary reconstruction of member-member networks, starting from the member-activity affiliations, i.e., the roll-call votes data. Indeed, when it comes to the network reconstruction, the preliminary step of most SNA approaches is affected by the simplistic assumption that: two people are related to each other if, and only if, they perform simultaneously a (sub)set of activities, and the strength of their interaction is measured directly counting (and weighting) these co-occurrences.

Evidently, this standard method for network reconstruction can be improved in its capability of finding *hidden* links, or removing those due to noise and bearing no useful information. Various strategies may contribute significantly in this improvement: both those originated from SNA realm itself (e.g., adopting homophily for the study of the network structure); or from other fields (e.g., analysing covariates generated by different observations of the network, making use of random/mixed effects models, or checking covariance data against pseudo-randomization in the samples, etc.). The interested reader may find more details in specific papers. For example, in Newman and Leicht (2007) it is performed the reconstruction of the clusters inside a large-scale network via mixture models, investigating similar structural connections among the nodes. A mixture model in random graphs is used also in Daudin et al. (2008), but this time enriching it with a Bayesian approach, with the purpose to infer unknown classes (Nowicki and Snijders, 2001). Finally, in Jedidi et al. (1997) a general finite mixture structural equation model is built, capable of dealing with heterogeneities in the network's structural equation models, and based upon a set of observed variables (measured with error). In general, these approaches may adopt finite mixture simultaneous equation models, finite mixture confirmatory factor analysis, and finite mixture second-order factor analysis.

Most statistical methods outlined above are a way to relax the strong assumption that filters only those interactions due to co-occurrences. Here, instead, it is discussed an approach adopting the *inverse Potts* model, originated from Condensed Matter Physics. Inverse models aim to infer and model the interactions in an unknown network structure, starting from recursive observations of the nodes' states. As such, these models are adapt to capture underlying quantum structures in a decision making process, whenever the final decision state can be deduced in terms of the observed actions (this argument will be discussed in Section 2). Moreover, this paper also envisages how a Q-states Potts model enables a much

better understanding and mimicking of the statistical features of complex network structures, compared for example to a more basic Ising modeling[1]. The approach is tested against a policy network reconstruction, starting from co-sponsorship data collected from the Italian Senate[2].

It is worth to notice how Ising and Potts (*direct*) models have already found a large number of applications also in the realm of social sciences (Phani et al., 2004; Bordogna and Albano, 2007), including policy networks (Liu et al., 2010), but always applied to networks whose structure had been inferred previously by other strategies. However, the inverse problem formulation has been confined to the Ising model alone, and most of its interest for non-physical problems has involved so far only biological and neural sciences (Yamanishi et al., 2004; Ricci-Tersenghi, 2012), or image reconstruction tasks (Kiwata, 2012). To authors' knowledge, this paper is the first using the inverse Potts problem to *reconstruct* a network in social sciences, and it is in general the first to apply a moment-based Loopy Belief Propagation (LBP) method[3] to solve the Potts inverse problem in the real world.

The paper will be structured as follows. In Section 2, we will present how the Q-states Potts model intervenes in network reconstruction, and our approach to solve it. Then, in Section 3, a reconstruction of the Italian Senate network is reported, starting from data tracking co-sponsorships of law-proposals and inferring interactions among the senators, according to their decision patterns. Finally, in the Conclusions we will compare the results with traditional SNA methods, i.e., not employing statistical inference.

2. MODEL AND METHODOLOGY

The principle behind the approach described in this paragraph is that (co-participation in) activities of an organization lead(s) to *two-body interactions* among the organization members, and these interactions can be captured by a networked structure. In other words, a complete approach handling relations between different realms (e.g., users and activities) must be able also to examine relations within each realm, separately. Using typical SNA nomenclature, this means computing a *one-mode* network (represented by an *adjacency* matrix), starting from a *two-mode* network (represented by an *affiliation* matrix, that reports participations in the activities, by different organization members). Currently, the standard approach to deduce the one-mode matrix is based upon a mere counting and normalization of co-occurrences, according to some schemes: these include matches-counting, covariance and correlation measures, cross-products, up to Bonacich and Jaccard indexes (Hanneman and Riddle, 2005). Each of these methods brings along some peculiar features, and the Jaccard index in particular is widely adopted (Borgatti, 2012), being well-suited for sparse affiliation matrices that are very common in the real world.

However, none of these standard approaches resembles probabilistic features, capable of taking into account noisy

---

[1] Adopted elsewhere in SNA literature for the same task.
[2] Publicly available at http://www.senato.it/leg/16/BGT/Schede/Attsen/Sena.html.
[3] See Section 2.

signals, anti-correlations and co-occurrences of *idle states*[4]. This issue highlights the chance to improve the reconstruction of the corresponding one-mode networks, by a mapping to an inverse statistical problem for pairwise Markov Random Fields (MRF), as discussed in detail later. Especially for large systems, inverse statistical problems are computationally expensive, and approximate methods must be used. For the inverse Ising problem are known: expansions in *correlations* and *clusters* (Sessak and Monasson, 2009; Cocco and Monasson, 2011), methods based upon the *Bethe* approximation (Ricci-Tersenghi, 2012), and *pseudo-likelihood* methods (Ekeberg et al., 2013). Here, we will refer to the moment representation of the LBP approach (MR-LBP), considered particularly advantageous for solving this task (Horiguchi, 1981).

However, Ising models pose severe limitations for SNA applications[5], and it would be advantageous to switch to a more general Q-state Potts modeling. A theoretical extension of MR-LBP for this general inverse problem has been provided already by Yasuda et al. (2012), making use of an expansion in *Chebyshev polynomials*. This approach is briefly outlined in this paragraph, before explaining how to match it with the specific needs of our case. As the first, however, it is important to discuss at an introductory level why inverse Potts (Ising) modeling are considered adapt to deal with affiliation matrices, that may well derive from quantum features of decision processes.

The starting point of the inverse Q-state Potts problem is a set of M observations: $\mathcal{D} = \{d^\mu \in \{0, 1, ...Q-1\}^n | \mu = 1, 2, ...M\}$. The task of the inverse problem is to reconstruct the (Potts) model subtended to the observations[6]. In other words, each observation in $\mathcal{D}$ can be considered a "snapshot" of the network at a certain moment in time, where the (positive integer) *state* of each node x_i is observed as d_i^μ at the μ-th observation. Pairwise states are indicated as $x_{(i,j)} := \{x_i, x_j\}$, meaning that, at the time of the same observation, the states of nodes x_i and x_j were found to be as in $x_{(i,j)}$. In this study case, the allowed Q-states can be interpreted as the possible decisions and thus positions (both active or not), about a law proposal, which can be held by the Senate members.

Now, among the fundamental principles of Quantum Mechanics, there is the possibility that if an object can be in either of two generic *orthogonal*[7] states $|\phi\rangle$ and $|\psi\rangle$, then, in general it is also allowed to be in any linear superposition of the two: $\alpha|\phi\rangle + \beta|\psi\rangle$. Intuitively, however, when a measurement of the object's state is performed, the state must collapse into either one *or* the other. This is also at the core of many models exploiting quantumness in the cognitive realm (Haven and Khrennikov, 2013). Mapping this general statement into our specific study case is equivalent to supposing that policy network agents perform decisions according to the same scheme of a quantum state measurement. Intuitively, this means that these agents do not

already "embed" a decision about what to do, before being asked support for a roll-call. Only when they are confronted with the decision making, they *contextually* choose one of the possible alternatives to act: before that moment, it is possible to suppose they were in a superposition of some (all) possible decisions. I.e., they were considering also alternatives, before finalizing their choice.

More formally, the generic decision state of each senator can be mapped as a superposition state $|X_i\rangle$, in (some of) the Q-states $|\chi\rangle$ of the Potts model:

$$|X_i\rangle = \sum_{\chi=1}^{Q} \beta_{i,\chi} |\chi\rangle \tag{1}$$

and each observation of a node's state can be understood as a POVM of $|X_i\rangle$ in the basis of the states $|\chi\rangle$, that are mutually orthogonal. This underlies the plausible assumption that a single member may *desire*—but not *intend*—more than one decision at once, toward a certain law proposal: for example they cannot simultaneously support and ignore the same roll call. Non-classical effects of this superposition of states guiding the final decision have already been discussed, e.g., in Aerts et al. (2012), and a more complex quantum modeling of decision making has been proposed in Bisconti et al. (2015).

It may be noticed that, when introducing at first the Potts model in this paragraph, no explicit reference to quantum states was made. In fact, this is because an effective treatment of the quantum Potts model can be done within a classical formalism: a more technical justification follows in the rest of this paragraph. Indeed, a quantum Potts model introduces a Hamiltonian characterized by two-body[8] interactions as:

$$\mathcal{H}_{Potts} = - \sum_{\{i,j\}} H_{(i,j)} \sum_{\chi} P_i^\chi P_j^\chi \tag{2}$$

where P_i^χ are projectors onto the $|\chi\rangle$ state of the local space for the i-th node. H is instead called the *ferromagnetic coupling*, and it captures the intensity of interaction among the nodes.

It is known how any classical (finite-dimensional) spin model on a lattice can be associated to a quantum model (Somma and Ortiz, 2010), defined on the same lattice, by mapping every classical state x_i into measurement outcomes of the state $|X_i\rangle$ and viceversa. Classically, the spin model has an energy functional that is:

$$\mathcal{E}_{Potts} = - \sum_{\{i,j\}} H_{(i,j)} x_{(i,j)} \tag{3}$$

Therefore, the energy functional maps into the eigenvalues of the Hamiltonian operator defined in Equation (2), and when performing statistical inference from the observations of the nodes' states in the network, this correspondence allows us to refer directly to the values of the classical variable x_i. In the following, therefore, the baseline assumption will be that a model subtending a statistical treatment of the network reconstruction

---

[4]To be intended as those states, that label nodes observed to be inactive, whereas certain other activities are being performed by other nodes.

[5]E.g., the maximum number of allowed states is intrinsically limited to 2, while a generic Q-states Potts model allows $Q \in \mathcal{N}$, $Q \geq 2$. Indeed, in Section 3, it will be shown how an inverse Ising problem fails for our case study.

[6]To be specific, the observations are supposed to be sampled from a certain MRF.

[7]I.e., they cannot be observed simultaneously for the same object.

[8]Here and in the following single-node terms are skipped for simplicity.

problem, inspired by a quantum-mechanical counterpart, can be far more efficient in revealing hidden links and patterns from observations, inferring even those interactions that standard methods are not capable of detecting.

2.1. The Inverse Potts Model

It has been seen how a statistical approach to the Potts problem, dealing with classical variables x_i, still implicitly underlines an intrinsically quantum process of decision making, because the likelihood of observing a certain value d_i for x_i can be interpreted in terms of projecting the generic quantum state $|X_i\rangle$, onto the corresponding basis state $|\chi\rangle$, where each of the orthogonal basis states identifies a single possible decision. This paragraph is devoted to a detailed explanation of the algorithm inferring relationships among nodes, from the set of observations performed: non-technical readers may skip it and move to the considerations in Section 2.2.

It can be observed how the probability distribution—for observations of the node states x—is clearly connected with the energy functional in Equation (3):

$$\mathcal{P}(x) \propto \exp\left[-\sum_{(i,j) \in E}^{n} H_{(i,j)}(x_{(i,j)}) \right] \tag{4}$$

and this closely resembles the probability distribution in general pairwise MRF formalism. E defines here the set of connections expected in the model, and therefore the condition $(i,j) \in E$ set in the summation can be understood as an explicit network constraint, whereas in Equation (3) we had the generic $\{i,j\}$. Now, in the inverse problem, the $H_{(i,j)}$ setting up the network model are unknown[9] and must be inferred by the probabilities in Equation (4). In terms of the orthogonal set of Chebyshev polynomials $\Phi_k(x_i)$ and appropriate constants $J_{(i,j)}^{(k,l)}$, it is possible to write the two-body potential function H as:

$$H_{(i,j)}(x_{(i,j)}) = \frac{1}{\sqrt{Q}} \sum_{k=1}^{Q-1} \left[J_{(i,j)}^{(k,0)} \Phi_k(x_i) + J_{(i,j)}^{(0,k)} \Phi_k(x_j) \right] +$$
$$+ \sum_{k=0}^{Q-1} \sum_{l=0}^{Q-1} J_{(i,j)}^{(k,l)} \Phi_k(x_i) \Phi_l(x_j) + \text{constant} \tag{5}$$

where constant terms in the expansion (e.g., $\Phi_0(x_i)$) have been all included in the last constant term. Starting from Equation (5), Yasuda et al. (2012) applied a moment representation of the LBP scheme and *message-passing* rules to the MRF described so far. Within the *Bethe* approximation, it was shown how it is possible to approximately find the constants J from marginal probabilities of the observations:

$$J_{(i,j)}^{(k,l)} = -\sum_{x_i=0}^{Q-1} \sum_{x_j=0}^{Q-1} \Phi_k(x_i) \Phi_l(x_j) \ln \mathcal{P}_{(i,j)}(x_{(i,j)}|\mathcal{D}) \tag{6}$$

thus minimizing the (Bethe) approximate entropy of the model: the $\mathcal{P}$ probability values are used to reconstruct the parameters of the Potts model.

The probabilities $\mathcal{P}$, for observing in $\mathcal{D}$, respectively values x_i and $x_{(i,j)}$, can also be expressed as sums of Chebyshev polynomials:

$$\mathcal{P}_i(x_i|\mathcal{D}) = \frac{1}{Q} + \sum_{k=1}^{Q-1} \langle \Phi_k(x_i) \rangle_{\mathcal{D}} \Phi_k(x_i) \tag{7}$$

$$\mathcal{P}_{(i,j)}(x_{(i,j)}|\mathcal{D}) =$$
$$\frac{1}{Q^2} + \frac{1}{Q} \sum_{k=1}^{Q-1} [\langle \Phi_k(x_i) \rangle_{\mathcal{D}} \Phi_k(x_i) + \langle \Phi_k(x_j) \rangle_{\mathcal{D}} \Phi_k(x_j)]$$
$$+ \sum_{k=1}^{Q-1} \sum_{l=1}^{Q-1} \langle \Phi_k(x_i) \Phi_l(x_j) \rangle_{\mathcal{D}} \Phi_k(x_i) \Phi_l(x_j) \tag{8}$$

Here, the interesting advantage of using the LBP moment representation is that all the quantities $\langle ... \rangle_{\mathcal{D}}$ can be derived by averaging over an appropriate number of M observations $\mathcal{D}$ of the network.

It can be both intuitively predicted, and numerical experiments in Yasuda et al. (2012) confirmed it, that the number of observations used is correlated with the quality of the final network reconstruction obtained. It shall be observed how in the original paper, numerical experiments were limited to the case when the network structure underlying the inverse problem was a non-periodic lattice (i.e., $|i-j| \notin \{\theta[\min(i,j) \mod p], p\} \Rightarrow (i,j) \notin E \Leftrightarrow J_{(i,j)} = 0$, where θ the step function and p the lattice period).

Considering that the main specific interest of this paper is the reconstruction of the network, i.e., the pairwise interactions among the nodes, here the key parameter in the Potts model is indeed $H_{(i,j)}$, measuring the intensity of connection between users i and j in the network. Equation (5) shows that $H_{(i,j)}$ is directly related to the set of constants $J_{(i,j)}^{(k,l)}$.

An interesting feature, that contributes to the sensitivity of this approach compared to standard ones listed above, is that $\Phi_k(x_i)\Phi_l(x_j)$—used in Equation (6) for calculating J—is in general different from 0, even when $k \neq l$. Therefore, interactions are inferred also when simultaneous participation in the same activity plays no role. The interpretation is that, even if one expects no interaction to occur among users because they tended to perform different activities[10] in the observation snapshots, this assumption is actually tested by the reconstruction method against the observations, and indirect ("out-of-diagonal") correlations may be detected.

As better explained in Section 2.2, in most cases data collected from social networks require caution before being used as "observational data" in a Q-state inverse Potts problem.

---

[9]And in particular, it is unknown which nodes in the interaction model E are truly linked to each other, i.e., have a non-negligible interaction: $(i,j) \in E \iff H_{(i,j)} \ngeqq 0$.

[10]I.e., assuming that $J_{(i,j)}^{(k,l)} \propto \delta(k,l)$, where $\delta(k,l)$ is the *Kronecker delta*. Indeed, also in the *pseudo-observational* models defined below, the parameters $\alpha_{(i,j)}$ act as an initial guess for the interactions, based upon the assumption that interactions shall be inferred only when simultaneous participation in activities occurs, but this is tested against the observations.

Therefore, it is interesting to mention the possibility to simulate observations, whenever data about the probability distributions are known to depend upon some parameter(s). For example, observation samples may be reconstructed using one only parameter α in a *generative* model: in this case probabilities of observing a certain collective state x are computed according to α, and it is possible to write down averaged functions (such as the averages required by Equations 7 and 8):

$$\langle f(x)\rangle_{\mathcal{D}} = \sum_{x\in\mathcal{D}} f(x)P_{GenMod}(x|\alpha) \qquad (9)$$

Clearly, because of the assumptions underlying the LBP inverse problem approach, the most general choice for the generative model (*GenMod*) must be a Q-state Potts model.

2.2. Data and Observations

It is left to explain how to employ an inverse Q-state Potts model for the reconstruction of the Italian Senate network of members, starting from data tracking law co-sponsorships by senators. In the Italian legislative system, a law undergoes a few preliminary steps before being discussed in the Senate. As the first, one or more[11] senators are responsible for writing it down, signing and proposing it; these are very similar to the *sponsors* in US legislative system. After that, other senators who are aware that this specific law is being proposed for discussion, may co-sign it, as an act of endorsement. They act as the US *co-sponsors*. According to the Senate's schedule, the law is then discussed in detail and subjected (eventually) to a final vote. Therefore, collecting (co)sponsorships' data brings along a considerable insight about patterns of collaboration and support among the senators, and can be considered equivalent to other studies performed with similar legislative bodies in other countries, as cited in the introduction.

Our case study focuses on the first *part*[12] of the XVI Italian legislature, using co-sponsorship data for Senate roll calls in the same period. We chose this period for two reasons. As the first, one of the intents is to find communities (and their members) in the network by automatic community detection algorithms, and compare the resulting groups with the "official memberships in political parties" of the senators. For this purpose, the beginning of a legislature is ideal, because senators have just been elected as members of a certain political party[13]. This makes it easy to refer to these parties as their *true memberships*, whereas at later points in time, several senators may have moved to different political parties (e.g., because some parties have been dismantled), and tracking these changes in a mindful way turns extremely difficult. Moreover, the dataset of this study case is the most recent (thus

eventually more interesting from a policy network point of view), while referring at the same time to a past Legislature. This renders available data "crystallized," with less risk of updates to occur.

Usage of minimization procedures in Potts-like models for legislative bodies is not fully new in the literature: for example, in Liu et al. (2010), an Ising model had been used to model the US Senate network starting from bill cosponsorships. However, compared with this previous study, there are here a few important differences.

- In Liu et al. (2010), the quantum Ising model is *not* used for the network reconstruction, achieved by a simple weighted interaction counts procedure. The Ising model intervenes merely in a second phase, for the influence maximization analysis.
- Because of the intrinsic political nature of the Italian VS the US Senate, whereas a 2-state Ising model is perfectly adapted to the strongly bipartite US case, it is rather limiting when used to describe the multi-partite structure of its Italian counterpart, that requires a more generic Q-states approach.
- Observing more closely the available data, US co-sponsorship data of the 108th Congress (used for the network analysis in Fowler, 2006 and derived ones) had in average 285 bills (co)sponsored per legislator—against 62 bills/legislator for Italian co-sponsorship data from the XVI legislation. Each US bill was (co)sponsored in average by 4 legislators—while 8 legislators per bill was the average in the Italian case. The total is of 4630 bills for 100 senators in the US case, and 3100 (M) law proposals for 338 (N) senators in the Italian case. Summarizing, the US Senate was much more active in sponsoring bills, and still proportionally more active in co-sponsoring, when compared to the Italian counterpart.

The connectedness of the US legislative network, given by the ratio cosponsorships/senators, and the reduced number of communities therein, make it adapt of being treated with an Ising 2-state model. Also standard methods may reproduce the structure of that network in an acceptable way, given that its high density may well represent[14] the absence of hidden or evolving links. This considerations, however, suggest that the same approach may provide poor results for the Italian situation.

Here we intend to use a Q-state Potts model *directly* for the network reconstruction, as outlined in Section **??**. A naive application of the model may involve two only possible states for the nodes (senators).

1. An *active* state ($x_i = 1$), corresponding to nodes sponsoring or co-sponsoring a bill, when this is being proposed or introduced. It is indeed intuitive to consider the request for cosponsorships an observational event, measuring the behavior of the nodes, and therefore the state they are in.
2. A *passive* state ($x_i = 0$), when the nodes do not act as (co)sponsors when a bill is introduced (i.e., they are detected as inactive when they undergo "measurement"), and therefore no endorsement is tracked in the data.

---

[11] Usually one or just a few. In some special cases, the law undergoes a peculiar path where no initial senator is quoted for sponsoring the proposal.

[12] Corresponding to the first Cabinet.

[13] Indeed, when in the following there will be references to the "true" memberships of the senators, these have been deduced by the participation of the senators to political *groups*: a specificity of the Italian Parliament, that enables the tracking of a senator's loyalty to a party or group of parties. See http://www.senato.it/leg/16/BGT/Schede/GruppiStorici/Grp.html. When, along with the period under observation, a senator belonged to multiple groups, he/she was assigned to the group where he spent most of the observation time.

[14] Actually, a specular interpretation is that the mechanism used to track interaction was poorly efficient, and therefore links in excess shall be excluded by an inference scheme.

This would be equivalent to an Ising model. In order to better explain what follows, it is worth a parenthesis about the Italian case. It was highlighted how each *activity*, i.e., the proposal for each law in the Senate, was participated in average by about 8 people (N_A). That is, for each of the M bills, the active community was in average 2.5% of the whole Senate. Now, using a 2-state Potts model as above implicitly generates correlations also among senators often detected in passive states: Indeed, it is evident how co-occurrences of inactive states would be assigned the same importance, in principle, as co-occurrences of active states.

Is this meaningful? Consider the underlying phenomenon: co-endorsing a law proposal presumes a much more intensive link between two senators (as it obviously brings along the sharing of the same political point of view, as well as some sort of acquaintance with the senator who conceived the law itself), compared to simultaneous abstaining from the endorsement (which may be due to lack of chance to discuss and share the law proposal; or to early abundance of cosponsors, making worthless for other senators to join the cosponsoring group; etc.). A simple abstinence from action is an ambiguous behavior, as it supposes no direct opposition or lack of interest. Therefore, it is intuitively necessary to find a mechanism that keeps these *inactive correlations*[15] less significant, compared to those due to simultaneous observation of the same active state in two nodes[16].

A first approach may be to still use $Q = 2$, while explicitly ignoring inactive correlations when computing the interaction parameters. This can be done by replacing:

$$x_{i(j)} \in \{0, ..., Q-1\} \rightarrow x_{i(j)} \in \{1, ..., Q-1\} \qquad (10)$$

in the sums of Equation (6), where we supposed that $x_i = 0$ corresponds to the only inactive state. However, this choice will miss the chance of capturing hidden connections, due to simultaneous occurrence of inactive states for some specific reason, and particularly the hostility against the law proposal under discussion.

An effective solution, but computationally expensive, is to pick a high enough Q-value for the model, assigning different $x_i \neq 1$ to members in inactive states. In particular, to avoid aprioristic considerations about the level of interaction of people belonging to the same faction, the random probability of assigning two nodes to the same *inactive* state (p_{ina}) shall not be bigger than the average empirical probability of two nodes being assigned to the same *active* state (p_{act}). Now: $p_{ina} = \frac{N-N_A}{(Q+1)N} \leq p_{act} \cong 0.024$, which gives in turn: $Q \geq 40$. Because of the computational complexity of the procedure ($O(Q^2)$), here for demonstrative purposes it will be shown how the performance of the method can change moving from $Q = 2$ up to $Q = 10$. That is, we start by assigning to inactive correlations the same importance as active correlations, then we progressively reduce the importance of the second compared to the first ones. The case with $Q = 5$ has a specific underlying reason: community detection algorithms revealed 5 clusters in the Senate network, when run against the network, reconstructed with the standard Jaccard approach, see

Section 3. The intent is therefore to try using this information as an initial guess for the LBP approach, introducing a number of possible states corresponding to community membership (under the reasonable assumption that such a membership strongly influences the co-sponsorship decisions). However, it should be emphasized here that partitioning the network in 5 communities may be non-optimal. Indeed, along the period of the analysis performed, it is true that the Parliament involved 4 major parties, plus senators being independent, or belonging to small[17] parties, but the 4 major parties were actually joint in 2 different alliances, thus reducing the number of effective communities to only 3. This is an important consideration, therefore it will be discussed again in the following.

There is still another feature in the procedure, left to discuss: cleaning and eventually generating the observation samples. This feature can be tuned as well, in order to introduce aprioristic knowledge about the network structure. In general, there are at least three different strategies to use properly the collected data:

1. a full *generative* model, where at first some standard method is applied to reconstruct the network, this network is used as an initial guess for the interactions among the nodes, allowing to sample a set of observations;
2. a *semi-observational* model, where the observations collected are used directly as samples, but sample averages are adjusted against the network reconstructed via standard methods;
3. a pure *observational* model, that is agnostic of any coarse-grain network structure, and applies directly the LBP procedure to the data: here sample averages are computed directly from the data (that can thus be confused with the observation samples).

It is worth to notice how the first strategy replaces real data with samples obtained according to some reasonable[18] assumptions about the strength of relationship among nodes, summarized as α_{ij} elements[19] of a preliminary adjacency matrix. For example, in Liu et al. (2010) a count of co-occurrences of active states was used:

$$\alpha_{ij} = \sum_\mu \frac{\delta(x_i^\mu, x_j^\mu, 1)}{n_\mu} \qquad (11)$$

weighted with the number of cosponsors n_μ, for each bill μ.

Even if a standard method is used for the preliminary calculation of the interaction among the nodes, the LBP procedure still intervenes in allowing to infer hidden connections, not evident from the first step. The generative approach is particularly useful whenever only a few or only aggregate[20] data are available for the analysis. However, this

---

[15]As we will call them in the following for simplicity.
[16]In the following for simplicity: *active correlations*.
[17]Here by *small*, we intend parties with a number of senators below the threshold of 10, because this is the minimum number to constitute an official *group* in the Italian Senate. Smaller parties are obliged to group in the so called *mixed group*.
[18]E.g., frequency considerations, as those used in the following Equations (11) and (13) for the semi-observational model.
[19]Used in a second step to calculate the probability of observations in the sample, see Equation (12).
[20]I.e., there is no temporal allocation of the single observations, but only a global count.

strategy still introduces a manipulation of original data, in order to make the network reconstruction possible or less noisy[21].

The semi-observational strategy can be seen as a compromise between the adoption of a generative model, and the direct usage of data with no further adjustments. In this case, observation data are used directly for each step of the network reconstruction, except the calculation of averages. More technically, in this case $\langle \Phi_k(x_i) \rangle_{\mathcal{D}}$ and $\langle \Phi_k(x_i)\Phi_l(x_j) \rangle_{\mathcal{D}}$ are not simple averages from the samples' set, but they are adjusted according to Equation (9). As an example, the P_{GenMod} probabilities of occurrence of the state x may be chosen as:

$$ P_{Potts}(x|\alpha) \propto \exp\left[\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{ij}\delta(x_i, x_j) \right] \qquad (12) $$

Generally speaking, the introduction of a subtending model as in Equation (12) favors a reconstruction similar to the output of the standard preliminary reconstruction method, because the probabilities of observing configurations (not) matching the standard reconstructions are increased (decreased), compared to the probabilities calculated directly from the observations.

The α parameters were evaluated here in terms of frequencies of matching activities, within the set of observations, according to different approaches. One possibility is a pure frequentist probability, for the two nodes i and j to be observed in the same *active* state:

$$ \alpha_{ij} = \frac{1}{M} \sum_{\mu} \frac{\delta(d_i^{\mu}, d_j^{\mu}, 1)}{n_{\mu}} \qquad (13) $$

with generalized Kronecker $\delta(i, j, k) = 1 \iff i = j = k$ and null otherwise, and the same weighting of Equation (11). This strategy penalizes the interactions of those nodes having a poor participation rate.

A second derivation for α, instead, was adjusted against the number of times the two users were active:

$$ \alpha_{ij} = \frac{1}{\sum_{\mu}\left(\delta(d_i^{\mu}, 1) + \delta(d_j^{\mu}, 1)\right)} \sum_{\mu} \frac{\delta(d_i^{\mu}, d_j^{\mu}, 1)}{n_{\mu}} \qquad (14) $$

thus reducing the bias of the previous formula toward active nodes.

Finally, when a pure observational method is used, the α parameter should play no role[22], because no generative model needs to be provided and all the averaged quantities are computed as from the original set of data. Unfortunately, a pure observational method with the considered dataset (characterized by $Q = 2$, because of lacking information) intuitively requires to omit the contribution of inactive correlations, such as in Equation (10), in order not to overestimate their contribution.

Whatever the strategy chosen to derive observation samples from original data, the interactions $H_{(i,j)}$ will be calculated replacing in Equation (5) the pairwise interactions J from Equation (6).

3. RESULTS AND DISCUSSION

It was envisaged the importance of the LBP inference method, for discovering non-evident links and connections among the network members, as compared to traditional methods not employing statistical inference. This paragraph illustrates the first numerical application of a LBP procedure, to reconstruct a generic graph Potts model. Previous simulations (Yasuda et al., 2012), indeed, dealt only with lattice-like Potts models: the sums in Equation (12) had a constant α instead of $\alpha_{i,j}$, and the allowed indexes were only those compatible with the lattice structure $(i, j) \in E$.

The first and most important results to be observed are in **Table 1** and in **Figure 1**. In the table are reported the main network parameters for the various methods listed in Section 2. As a comparison, the network was also reconstructed via the *Jaccard index*, a standard method particularly adapt to sparse networks (Borgatti, 2009), such as the one analyzed in this paper (the calculated density is indeed smaller than 0.01). For LBP-reconstructed networks, we introduced an additional parameter, the *threshold* (t_m). In fact, after normalizing the intensity of connections (i.e., $0 \leq H_{(i,j)} \leq 1$), the density of these networks was close to 1 in most approaches. This is an effect of the sensitivity of the LBP method, prone to reproduce in the final *adjacency* matrix also links due to noise. In order to exclude the weakest links, we set a threshold value $t_m = 0.5$, thus comparing the residual links with the standard network. It is evident how in all cases, also the LBP-reconstructed social network displays many more connections compared to the Jaccard one. These *hidden* connections would be hard to identify without referring to an inference statistical method, and this is a novelty of the approach. In the pure observational case, because *off-diagonal* interactions were neglected (i.e., the case in Equation 10), also noisy connections tend to occur in a small range, thus producing still a very high density at $t_m = 0.5$. Because of the increased difficulty to filter properly this noise, the pure observational model will be omitted from analyses in the following. Also the average strength of *all* the links detected "Avg. $H_{(i,j)}$" has an interesting behavior: it is strongly affected by the initial guess for the network structure, that the LBP method tries to reproduce, and considerably less, instead, by the value chosen for Q.

In **Figure 1**, instead, it is performed a more systematic analysis of the relation between the number of links in the network, against the threshold parameter[23]. A few interesting features are evident. As the first, all statistical methods tend to saturate the network at low values of t_m. Moreover, a smoothing effect in the dependency of the number of links on the threshold value is observed, both when increasing Q or decreasing the average α_{ij}. The higher the smoothing, the closest are the data to the expected

---

[21]Indeed, once the preliminary model has been decided, samples can be drawn from it in abundance, whereas a real sampling of a social network is clearly bound to pragmatic constraints.

[22]For the case $Q = 2$, actually, it is advised to adopt a fictitious $\alpha =$ constant $\ll 1$, because the critical value for the 2-state Potts model is $\alpha = 0.88$, therefore the calculation may turn unstable if averages are computed directly.

[23]Clearly, the density and thus the number of links detected in the Jaccard network is independent from any threshold chosen.

TABLE 1 | Collection of fundamental model and network parameters, for a set of network reconstruction methods, and basic metrics resulting from the analysis.

Method	# Samples	Q	Avg. α	Density of reconstructed network ($t_m = 0.5$)	Avg. $H_{(i,j)}$ (normalized)
Jaccard index	3100	NA	NA	0.00202	0.04901
Generative	100,000	5	0.65984^{-04} (Equation 13)	0.01554	0.16336
	100,000	10	"	0.01907	0.16938
Semi-observational	3100	2	0.65984^{-04} (Equation 13)	0.00540	0.17612
	3100	5	"	0.00936	0.13385
	3100	5	0.01650 (Equation 14)	0.48623	0.50464
Pure observational	3100	2	0.01*	0.99275	0.98301

*The analysis with Q = 10 required to artificially generate more samples than the direct observations from data, for the results to be reliable, therefore it is listed only within the Generative case. *For numerical convergence reasons, in the Pure observational case, it was set $\alpha_{ij} = constant = 0.01$.*
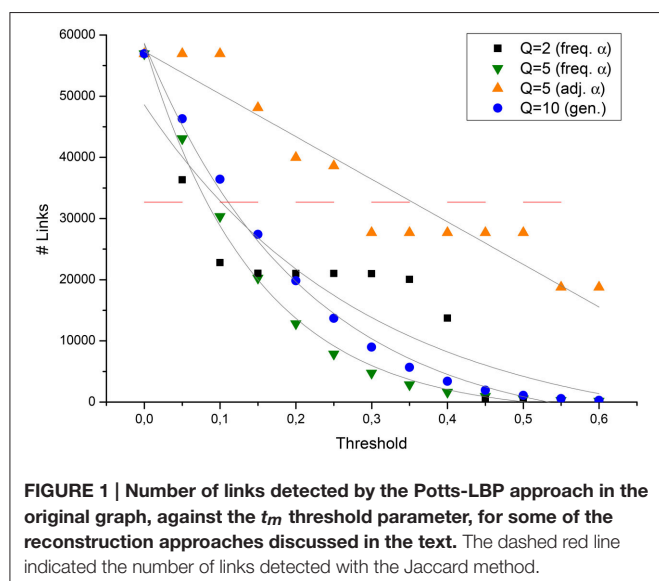
FIGURE 1 | Number of links detected by the Potts-LBP approach in the original graph, against the t_m threshold parameter, for some of the reconstruction approaches discussed in the text. The dashed red line indicated the number of links detected with the Jaccard method.
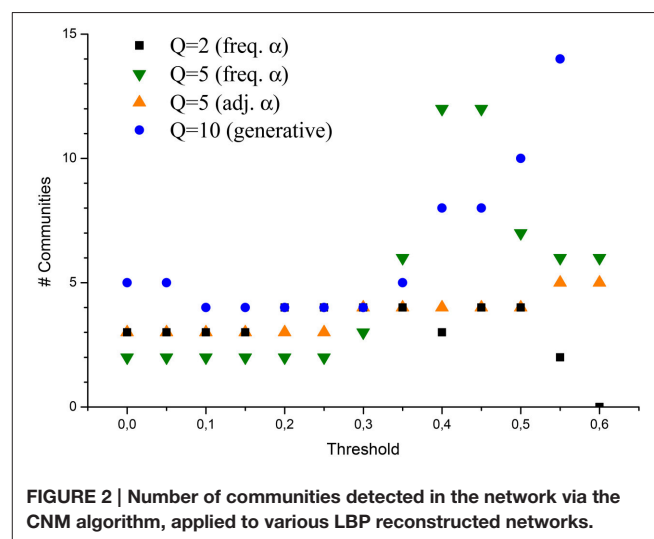
FIGURE 2 | Number of communities detected in the network via the CNM algorithm, applied to various LBP reconstructed networks.

exponential decay in the number of detected links[24]. Several possible explanations for this conclusion may be proposed. As the first, preliminary community detection analyses with networks reconstructed via standard methods identified 5 groups[25] in the Italian Senate network. This suggests how the observation of only 2 states with the roll calls tends to produce distortions and artifacts. In fact, results are improved also by randomly introducing states other than the observed (non)sponsoring. Smaller values of α, instead, allow the method to compute the $J_{(i,j)}^{(k,l)}$ *not* in proximity of critical values $ln(1 + \sqrt{Q})$ of the Potts model, thus improving the stability of the results.

A different analysis was focused about the capability of the method not only to reconstruct pairwise interactions, but also to better identify the clusters inside the network, to be interpreted as communities of members. In **Figure 2** it is investigated how

the number of such communities depends on t_m: the plot shows that when a small t_m is taken into account[26], LBP reconstructed networks have a cluster structure involving 2–3 groups. A plausible interpretation is that, if hidden links are considered, slight differences in the policy approach by the Senate members are swiped out in the analysis, and the CNM algorithm tends to detect only the fundamental communities: the ones related with the party(ies) participating in the Cabinet, and the group of parties opposing the first ones (plus eventually a third group which may be considered as composed by *neutral* senators). As the threshold is increased, and the graph becomes more disconnected, also clustering features are emphasized, and the number of detected communities increases. In particular, when a high number of possible states is allowed (high Q), and at the same time weak interactions are hypothesized (α_{ij} is small in average), the number of communities tends to "explode." However, excluding this extreme case, detected communities are otherwise stable, ranging between 3 and 6. It is also evident how, when links in the network are filtered and the cluster structure emerges, the network assuming $Q = 2$ totally fails to reproduce a plausible number of communities: this is clearly due to the

---

[24]This can be inferred by recalling that the probability to observe a certain collective state x has the form in Equation (4).

[25]Here and in the following, communities are always detected with a very successful method based upon random graph theory, the Clauset-Newman-Moore (CNM) method (Clauset et al., 2004). Only communities whose sizes are bigger than 3 nodes will be considered, while isolated nodes and dyads will be omitted.

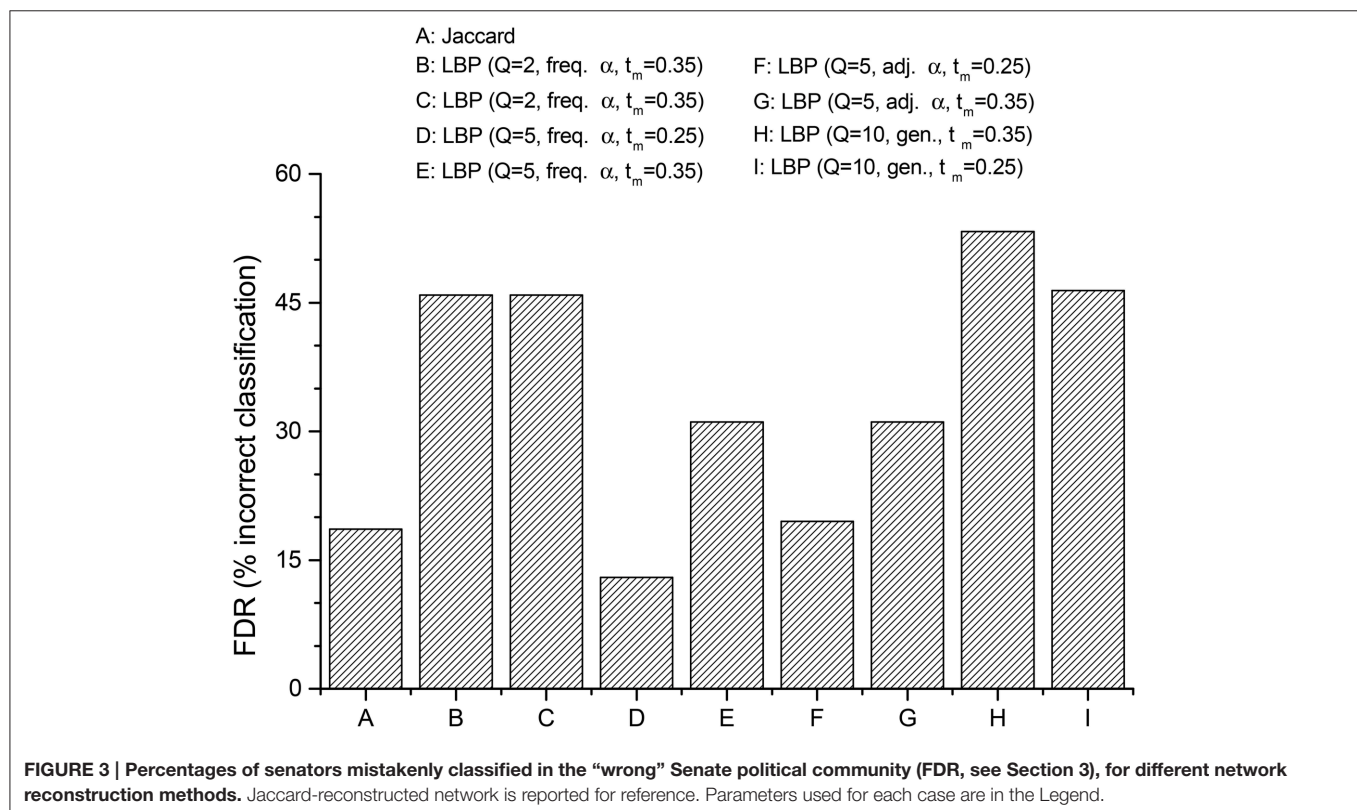[26]Preserving more links, indeed, leads to the discovery of weak interactions.

artifact of imposing a naturally bipartite network in the model, which does not correspond, though, to the expected network structure.

Finally, we investigated if—and how much—the LBP algorithm is able to improve the assignment of senator-nodes to the "right" political community: i.e., the one identified by the same political party, the senator officially belonged to. The figure of merit will be a sort of a *false discovery rate* (FDR), that is, the ratio between the number of senators assigned by the CNM algorithm to wrong communities, and the total number of senators analyzed. In order to emphasize the role of hidden links, the focus will be the capability of the algorithm to classify the senators as belonging to the group supporting the government, the opposition group, or the mixed independent group.

Referring to **Figure 3**, the reconstruction of the network via a simple Jaccard coefficient in this case is already capable of reproducing accurately the true membership of the nodes, scoring only about 15% of nodes classified. The case of LBP with $Q = 2$, instead, is very inefficient: almost half of the nodes is misclassifiedm, even if the target number of communities is close enough to allowed values of Q. This result shows the importance of extending the Ising model used elsewhere in analyses of policy networks: even when a subtended bipartite interaction is tracked (i.e., sponsoring VS abstaining in a roll call vote), in the end this is a projection of a more complex state, each agent in the network is before performing the voting action. From a modeling perspective, such a gap between model and reality can

be reduced by a full quantum Ising model (as showed by Liu et al., 2010), or by semi-classical approaches with a Q-state Potts model. In fact, results with $Q = 5$ display a great improvement compared to the case with $Q = 2$. Especially when t_m has a value in the range where the number of detected communities is stable, the percentage of nodes classified in the wrong group almost matches, or even outperforms the Jaccard one (13%), without assuming a-priori that indirect correlations among the network members are negligible. Interestingly, the best results are achieved for values of the threshold, corresponding to intervals where the number of comunities detected is stable (compare with **Figure 2**).

Moreover, it must be remembered how this analysis is affected by an important bias. States other than the active state (i.e., $x_i = 1$ for node i) are assigned randomly, therefore favoring communities of homogeneous cardinality: some misclassified nodes originally belonged to mid-size communities, but at their expense, these nodes where assigned to smaller groups. Networks obtained with very low thresholds are particularly prone to this effect. Some other misclassifications are due to a specular effect: similarly to the "*rich gets richer*" phenomenon, discovery of hidden links increases the size of the major communities at the expense of the smallest ones, as it is expected when modularity-based community detection algorithms are applied to very dense networks. Indeed, in all LBP-reconstructed networks the group of senators members of the party leading the Cabinet was always (mistakenly?) bigger than expected. This effect is evident comparing **Figure 4** with **Figure 5**, where the last one has
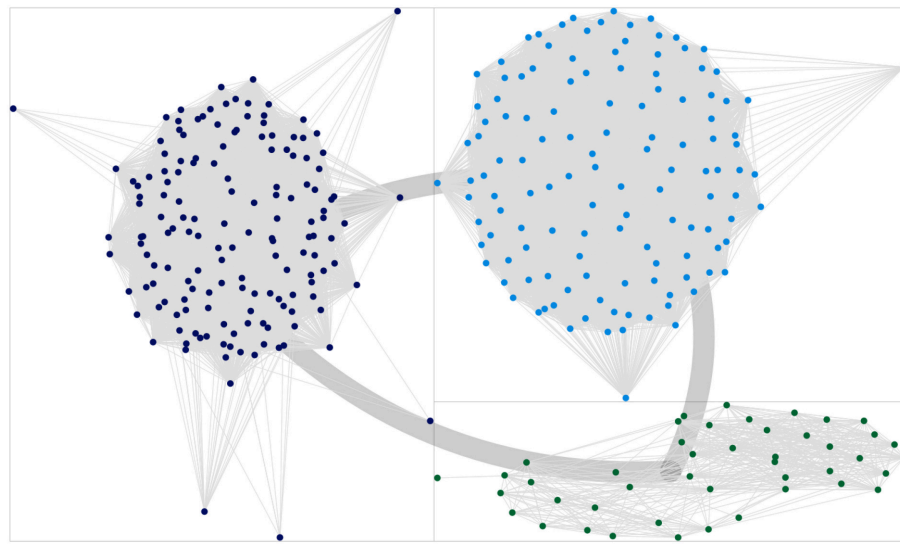

FIGURE 3 | Percentages of senators mistakenly classified in the "wrong" Senate political community (FDR, see Section 3), for different network reconstruction methods. Jaccard-reconstructed network is reported for reference. Parameters used for each case are in the Legend.

FIGURE 4 | Plot of the clusters obtained via the CNM algorithm, for the network reconstructed using the standard Jaccard method. Thick lines connecting the quadrants indicate the global cumulative strength of inter-community links. Dark blue dots indicate the community interpreted as "loyal to the Cabinet," light blue dots are connected with the "opposition," green dots are to be interpreted as "indipendent senators."
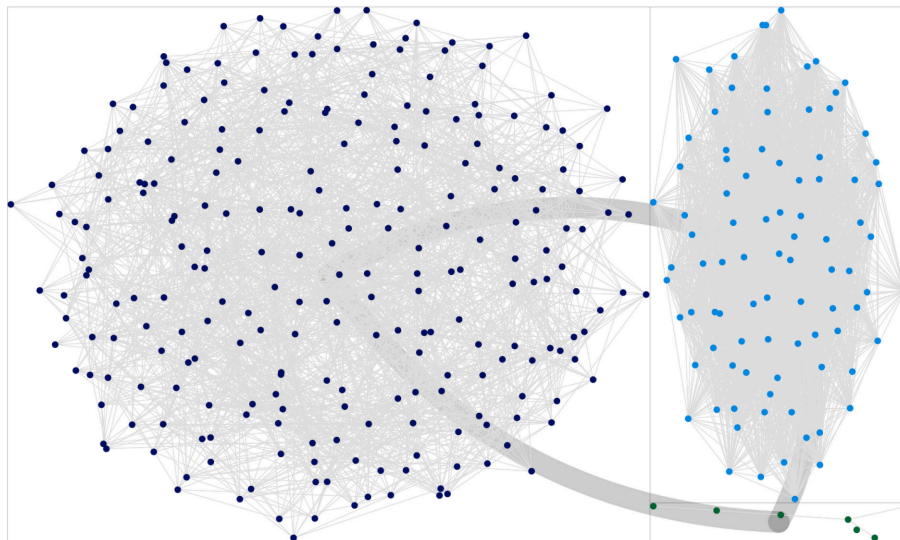

FIGURE 5 | Plot of the clusters obtained via the CNM algorithm, for the network reconstructed using a semi-observational LBP approach, with parameters $Q = 5$ and $t_m = 0.35$. It is evident how the bigger community (dark blue dots) is overestimated compared to standard approaches (see also **Figure 4**), at the expense of underestimating minor communities. As stated in the text, this effect can be reduced by lowering t_m.

been indeed obtained with LBP and a relatively high $t_m = 0.35$. The specular consideration above suggests how to compensate this artifact, by lowering opportunely the threshold t_m (the corresponding network graph is omitted for brevity). In any case, it shall be remembered how major Senate groups were actually bigger at the beginning of the legislature, compared to its end (when a few independent groups had been founded). By inferring weak links, it can be thus argued how LBP algorithms thus proved

more efficient in merging communities into a few principal components, compared to forcing modularity algorithms to split the network in 2–3 groups (i.e., forcing the CNM algorithm to merge further the 5 communities detected as optimal by its modularity maximization procedure, leading to the results in **Figure 3**).

On the other side, further increasing the value of Q required to rely upon a Generative model, in order to have a number of

samples sufficient for the analysis (100 k samples), whereas the reduced number of original data was prone to cause difficulties[27] in the numerical simulations. As envisaged in Section 2.2, moving from a (Pseudo-)Observational to a Generative approach produced a degradation in the results, because of losing the temporal information of the available data. In conclusion, an approach with higher Q, but still low enough to be based upon observational data, seems to produce the best and more stable results for both hidden links and community detection purposes.

4. CONCLUSIONS

Along the paper, a method based upon Q-state Potts inverse problem and Bethe-LBP approximation for network reconstruction was elucidated. Several possible ways were disclosed, to use the method for inferring links among the nodes of a generic networked social structure, under the hypotheses that: (i) actions like roll call sponsorships resemble decision-making processes and (ii) that these processes can be modeled efficiently by methods used for inferring the structure of an ensemble of quantum states, observed repeatedly over time.

The LBP-based resolution of the inverse problem was applied for the first time to reconstruct a generic graph structure. More specifically, in the Social Sciences realm, this work has been the first to use a Q-state model (instead of Ising model) to infer the structure of a real network. The study case chosen was the Italian Senate, analyzed starting from a dataset tracking law proposal co-sponsorships. This allowed to evaluate the power of the method in detecting those links, that cannot be retrieved via standard reconstruction methods. Also the role of the diverse modeling choices—and peculiar parameters employed—was thoroughly

---

[27]The simulation for high values of Q requires limited precision in the intermediate values calculated, to reduce the memory space required.

---

discussed, finding how the maximal value of Q permitted by the Potts model can introduce crucial differences in the quality of the results, alongside with aprioristic knowledge about the network structure.

It was investigated, as well, the capability of the model to reproduce the community structure of the network and the single memberships of the senators: it was found that the present method must be carefully reviewed, compared to standard ones, in order to produce a reliable output. In fact, a naive application without any further assumption may lead to completely wrong conclusions. The reason is that the Potts-LBP method is much closer to an ab-initio approach, therefore it originally embeds no information such as the weight to be assigned to inactive vs. active states, or direct vs. indirect correlations, or how weak connections shall be considered *noisy*, ... In turn, this higher flexibility allows to explore the role (and therefore the plausibility) of several assumptions made when reconstructing the network.

The authors envisage how interesting directions for further investigation may be the adoption of a full quantum treatment of the Potts model, as well as the possibility to apply this extended method to cases where data retrieved for the network do exhibit natively non-bipartite features, thus allowing a more direct application of generic Q-state Potts models.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Aerts, D., Sozzo, S., and Tapia, J. (2012). "A quantum model for the Ellsberg and Machina paradoxes," in *Quantum Interaction*, Vol. 7620, eds J. R. Busemeyer, F. Dubois, A. Lambert-Mogiliansky, and M. Melucci (Berlin; Heidelberg: Springer), 48–59. doi: 10.1007/978-3-642-35659-9_5

Besussi, E. (2006). *Policy Networks: Conceptual Developments and Their European Applications.* London: Centre for Advanced Spatial Analysis, University College London.

Bisconti, C., Corallo, A., Fortunato, L., and Gentile, A. A. (2015). A quantum-BDI model for information processing and decision making. *Int. J. Theor. Phys.* 54, 710–726. doi: 10.1007/s10773-014-2263-x

Bordogna, C. M., and Albano, E. V. (2007). Dynamic behavior of a social model for opinion formation. *Phys. Rev. E* 76:061125. doi: 10.1103/PhysRevE.76.061125

Borgatti, S. P. (2009). "2-mode concepts in social network analysis," in *Encyclopedia of Complexity and System Science*, ed R. A. Meyers (Heidelberg: Springer), 8279–8291.

Borgatti, S. P. (2012). "Social network analysis, two-mode concepts in," in *Computational Complexity*, ed R. A. Meyers (New York, NY: Springer), 2912–2924.

Börzel, T. (1997). What's so special about policy networks? An exploration of the concept and its usefulness in studying European governance. *Eur. Integr. Online Pap.* 1, 1–28.

Chiru, M., and Neamu, S. (2012). "Parliamentary representation under changing electoral rules: Co-sponsorship in the Romanian parliament," in *Inaugural General Conference of the ECPR Standing Group on Parliaments, Parliaments in Changing Times* (Dublin), 24–27.

Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Phys. Rev. E* 70:066111. doi: 10.1103/PhysRevE.70.066111

Cocco, S., and Monasson, R. (2011). Adaptive cluster expansion for inferring Boltzmann machines with noisy data. *Phys. Rev. Lett.* 106:090601. doi: 10.1103/PhysRevLett.106.090601

Daudin, J.-J., Picard, F., and Robin, S. (2008). A mixture model for random graphs. *Stat. Comput.* 18, 173–183. doi: 10.1007/s11222-007-9046-7

Dowding, K. (1995). Model or metaphor? a critical review of the policy network approach. *Polit. Stud.* 43, 136–158. doi: 10.1111/j.1467-9248.1995.tb01705.x

Ekeberg, M., Lövkvist, C., Lan, Y., Weigt, M., and Aurell, E. (2013). Improved contact prediction in proteins: using pseudolikelihoods to infer Potts models. *Phys. Rev. E* 87:012707. doi: 10.1103/PhysRevE.87.012707

Fowler, J. H. (2006). Legislative cosponsorship networks in the US House and Senate. *Soc. Netw.* 28, 454–465. doi: 10.1016/j.socnet.2005.11.003

Hanneman, R. A., and Riddle, M. (2005). *Introduction to Social Network Methods.* Riverside, CA: University of California.

Haven, E., and Khrennikov, A. (2013). *Quantum Social Science.* Cambridge: Cambridge University Press.

Horiguchi, T. (1981). On the Bethe approximation for the random bond Ising model. *Physica A* 107, 360–370. doi: 10.1016/0378-4371(81)90095-9

Jedidi, K., Jagpal, H. S., and DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Mark. Sci.* 16, 39–59. doi: 10.1287/mksc.16.1.39

Kiwata, H. (2012). Physical consideration of an image in image restoration using Bayes formula. *Physica A* 391, 2215–2224. doi: 10.1016/j.physa.2011.11.025

Lazarsfeld, P. F., Berelson, B., and Gaudet, H. (1968). *The Peoples Choice: How the Voter Makes Up His Mind in a Presidential Campaign.* New York, NY: Columbia University Press.

Liu, S., Ying, L., and Shakkottai, S. (2010). "Influence maximization in social networks: an Ising-model-based approach," in *Communication, Control, and Computing (Allerton), 48th Annual Allerton Conference on* (Cambridge: IEEE), 570–576.

Newman, M. E., and Leicht, E. A. (2007). Mixture models and exploratory analysis in networks. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9564–9569. doi: 10.1073/pnas.0610537104

Nowicki, K., and Snijders, T. A. B. (2001). Estimation and prediction for stochastic blockstructures. *J. Am. Stat. Assoc.* 96, 1077–1087. doi: 10.1198/016214501753208735

Phani, D., Gordon, M. B., and Nadal, J.-P. (2004). "20 social interactions in economic theory: an insight from statistical mechanics," *in Cognitive Economics: An Interdisciplinary Approach*, eds P. Bourgine and J.-P. Nadal (Berlin; Heidelberg: Springer-Verlag), 335.

Porter, M. A., Mucha, P. J., Newman, M. E., and Friend, A. J. (2007). Community structure in the United States house of representatives. *Physica A* 386, 414–438. doi: 10.1016/j.physa.2007.07.039

Ricci-Tersenghi, F. (2012). The Bethe approximation for solving the inverse Ising problem: a comparison with other inference methods. *J. Stat. Mech. Theory Exp.* 2012:P08015. doi: 10.1088/1742-5468/2012/08/P08015

Sessak, V., and Monasson, R. (2009). Small-correlation expansions for the inverse Ising problem. *J. Phys. A* 42:055001. doi: 10.1088/1751-8113/42/5/055001

Somma, R., and Ortiz, G. (2010). "Quantum approach to classical thermodynamics and optimization," in *Quantum Quenching, Annealing and Computation*, eds A. K. Chandra, A. Das, and B. K. Chakrabarti (Berlin; Heidelberg: Springer), 1–20.

Yamanishi, Y., Vert, J.-P., and Kanehisa, M. (2004). Protein network inference from multiple genomic data: a supervised approach. *Bioinformatics* 20(Suppl. 1), i363–i370. doi: 10.1093/bioinformatics/bth910

Yasuda, M., Kataoka, S., and Tanaka, K. (2012). Inverse problem in pairwise markov random fields using loopy belief propagation. *J. Phys. Soc. Jpn.* 81, 044801–044808. doi: 10.1143/JPSJ.81.044801

Zhang, Y., Friend, A., Traud, A. L., Porter, M. A., Fowler, J. H., and Mucha, P. J. (2008). Community structure in congressional cosponsorship networks. *Physica A* 387, 1705–1712. doi: 10.1016/j.physa.2007.11.004

CrossMark

Quantum information, cognition, and music

Maria L. Dalla Chiara[1]*, Roberto Giuntini[2], Roberto Leporini[3], Eleonora Negri[4] and Giuseppe Sergioli[2]

[1] *Dipartimento di Lettere e Filosofia, Università di Firenze, Firenze, Italy,* [2] *Dipartimento di Pedagogia, Psicologia, Filosofia, Università di Cagliari, Cagliari, Italy,* [3] *Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università di Bergamo, Dalmine, Italy,* [4] *Scuola di Musica di Fiesole, San Domenico di Fiesole, Fiesole, Italy*

Parallelism represents an essential aspect of human mind/brain activities. One can recognize some common features between psychological parallelism and the characteristic parallel structures that arise in quantum theory and in quantum computation. The article is devoted to a discussion of the following questions:

1. a comparison between classical probabilistic Turing machines and quantum Turing machines.
2. possible applications of the quantum computational semantics to cognitive problems.
3. parallelism in music.

Keywords: Turing machines, quantum computation, quantum information, semantics, music cognition

1. INTRODUCTION

Parallelism represents an essential aspect of the activities of human brain and mind. One can recognize some common features between psychological parallelism and the characteristic parallel structures that arise in quantum theory and in quantum computation, being responsible for the extraordinary efficiency and speed of quantum computers.

Quantum parallelism and classical parallelism are deeply different, although it is sometimes claimed that quantum Turing machines are nothing but special examples of classical probabilistic Turing machines[1]. But what exactly are quantum Turing machines? So far, the literature has not provided a rigorous "institutional" concept of *quantum Turing machine*. Some definitions seem to be based on a kind of "imitation" of the classical definition of *Turing machine*, by referring to a *tape* (where the symbols are written) and to a *moving head* (which changes its position on the tape)[2]. These concepts, however, seem to be hardly applicable to physical quantum computers. Both in the classical and in the quantum case, it is expedient to consider a more abstract concept: the notion of *state machine*, which neglects both tapes and moving heads. Every finite computational task realized in different computational models proposed in the literature can be simulated by a state machine[3]. In order to compare classical and quantum parallelism, we will analyze the concepts of (classical) *deterministic state machine*, (classical) *probabilistic state machine*, and *quantum state machine*. On this basis we will discuss the question: to what extent can quantum state machines be simulated by probabilistic state machines? (Sections 2, 3).

---

[1]See, for instance, Penrose (1994).

[2]See, for instance, Fouché et al. (2007).

[3]See, for instance, Savage (1998) and Gudder (1999).

In the investigation about possible links between *quantum structures* and *psychological structures* a useful tool is represented by a special form of quantum logical semantics (called *quantum computational semantics*) that has been inspired by the theory of quantum computation. We will see how this semantics can be naturally applied to a formal analysis of musical compositions, where *parallel structures*, *ambiguity*, *holism*, and *contextuality* play an essential role (Sections 4, 5)[4].

Our analysis seems to confirm a general conjecture that has been defended and discussed in different research-fields: the basic concepts of the quantum-theoretic formalism (which had for a long time been regarded as mysterious and potentially paradoxical) seem to have a *universal* interest that goes beyond the domain of microphysical phenomena.

2. CLASSICAL DETERMINISTIC AND PROBABILISTIC MACHINES

We will first introduce a formal definition for the notion of *deterministic state machine*. On this basis, *probabilistic state machines* will be represented as stochastic variants of deterministic machines, which are able to calculate different outputs with different probability-values.

Definition 1. *Deterministic state machine.*
A *deterministic state machine* is an abstract system **M** based on the following elements:

1. A finite set $\mathcal{S}$ of *internal states*, which contains an *initial state* s_{in} and includes a set of *halting states* $\mathcal{S}_{halt} = \{s_{halt_j} | j \in J\}$.
2. A finite alphabet, which can be identified with the set $\{0, 1\}$ of the two classical bits. Any *register* represented by a bit-sequence $w = (x_1, \ldots, x_n)$ is a *word* (of length n). Any pair (s, w) consisting of an internal state s and of a word w represents a possible *configuration* of **M**, which is interpreted as follows: **M** is in the internal state s and w is the word written on an ideal tape.
3. A set of words that represent possible *word-inputs* for **M**.
4. A *program*, which is identified with a finite sequence of *rules*:

$$(R_0, \ldots, R_t).$$

Each R_i is a partial function that transforms configurations into configurations. We may have: $R_i = R_j$ with $i \neq j$. The number i, corresponding to the rule R_i, represents the i-th step of the program. The following conditions are required:

4.1 The rule R_0 is defined for any configuration (s_0, w_0), where s_0 is the initial state s_{in} and w_0 is a possible word-input. We have:

$$R_0 : (s_0, w_0) \mapsto (s_1, w_1),$$

---

where s_1 is different from the initial state and from all halting states (if $t \neq 0$).

4.2 For any i $(0 < i < t)$,

$$R_i : (s_i, w_i) \mapsto (s_{i+1}, w_{i+1}),$$

where s_{i+1} is different from all $s_i, \ldots, s_0$ and from all halting states.

4.3 $R_t : (s_t, w_t) \mapsto (s_{t+1}, w_{t+1})$,
where s_{t+1} is a halting state.

Each configuration (s_{i+1}, w_{i+1}) represents the *output* for the step i and the *input* for the step $i + 1$.

The concept of *computation* of a deterministic state machine can be now defined as follows.

Definition 2. *Computation of a deterministic state machine.*
A *computation* of a deterministic state machine **M** is a finite sequence of configurations

$$((s_0, w_0), \ldots, (s_{t+1}, w_{t+1})),$$

where:

1. w_0 is a possible word-input of **M**.
2. $s_0, \ldots, s_{t+1}$ are different internal states of **M** such that: $s_0 = s_{in}$ and s_{t+1} is a halting state.
3. For any i $(0 \leq i \leq t)$,

$$(s_{i+1}, w_{i+1}) = R_i((s_i, w_i)),$$

where R_i is the i-th rule of the program.

The configurations (s_0, w_0) and (s_{t+1}, w_{t+1}) represent, respectively, the *input* and the *output* of the computation; while the words w_0 and w_{t+1} represent, respectively, the *word-input* and the *word-output* of the computation.

Apparently, each deterministic state machine is devoted to a single task that is determined by its program.

Let us now turn to the concept of *probabilistic state machine*. The only difference between deterministic and probabilistic state machines concerns the program, which may be stochastic in the case of a probabilistic state machine (**PM**). In such a case, instead of a sequence of rules, we will have a sequence $(Seq_0, \ldots, Seq_t)$ of sequences of rules such that:

$$Seq_0 = (R_{0_1}, \ldots, R_{0_r})$$

$$\ldots \ldots \ldots$$

$$Seq_t = (R_{t_1}, \ldots, R_{t_l}).$$

Each rule R_{i_j} (occurring in the sequence Seq_i) is associated to a probability-value p_{i_j} such that:

$$\sum_j p_{i_j} = 1.$$

From an intuitive point of view, p_{i_j} represents the probability that the rule R_{i_j} be applied at the i-th step. A deterministic

state machine is, of course, a special case of a probabilistic state machine characterized by the following property: each sequence Seq_i consists of a single rule R_i.

Any probabilistic state machine naturally gives rise to a graph-structure for any choice of an input-configuration $conf_0 = (s_0, w_0)$. As an example, consider the following simple case: a probabilistic state machine **PM** whose program consists of two sequences, each consisting of two rules:

$$Seq_0 = (R_{0_1}, R_{0_2})$$

$$Seq_1 = (R_{1_1}, R_{1_2}).$$

The graph associated to **PM** for the configuration $conf_0$ is illustrated by **Figure 1**.

How do probabilistic machines compute? In order to define the concept of *computation* of a probabilistic machine, let us first introduce the notions of *program-path* and of *computation-path* of a given probabilistic machine.

Definition 3. *Program-path and computation-path.*
Let **PM** be a probabilistic state machine with program $(Seq_0, \ldots, Seq_t)$.

- A *program-path* of **PM** is a sequence

$$\mathcal{P} = (R_{0_h}, \ldots, R_{i_j}, \ldots, R_{t_k}),$$

consisting of t rules, where each R_{i_j} is a rule from Seq_i (probabilistically independent of all other rules of $\mathcal{P}$).
- For any choice of an input (s_0, w_0), any program-path $\mathcal{P}$ determines a sequence of configurations

$$\mathcal{CP} = ((s_0, w_0), \ldots, (s_i, w_i), \ldots, (s_{t+1}, w_{t+1})),$$

where $(s_{i+1}, w_{i+1}) = R_{i_j}(s_i, w_i)$ and R_{i_j} is the i-th element of $\mathcal{P}$. This sequence is called the *computation-path* of **PM** determined by the program-path $\mathcal{P}$ and by the input (s_0, w_0). The configuration (s_{t+1}, w_{t+1}) represents the output of $\mathcal{CP}$.

Any program-path $\mathcal{P} = (R_{0_h}, \ldots, R_{i_j}, \ldots, R_{t_k})$ has a well-determined probability-value $p(\mathcal{P})$, which is defined as follows (in terms of the probability-values of its rules):

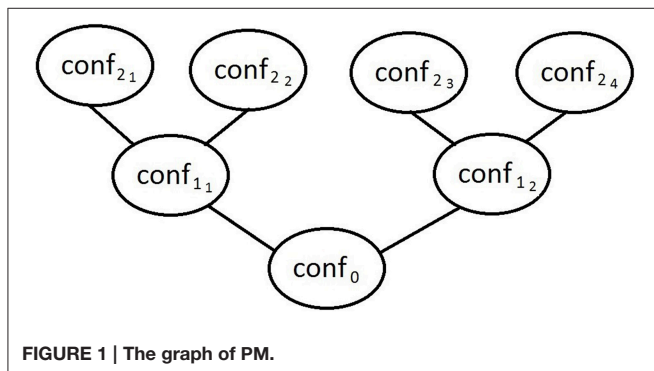$$p(\mathcal{P}) := p_{0_h} \cdot \ldots \cdot p_{i_j} \cdot \ldots \cdot p_{t_k}.$$

FIGURE 1 | The graph of PM.

As expected, the probability-value of a program-path $\mathcal{P}$ naturally determines the probability-values of all corresponding computation-paths. It is sufficient to put:

$$p(\mathcal{CP}) := p(\mathcal{P}).$$

Consider now the set $\mathbf{P_{PM}}$ of all program-paths and the set $\mathbf{CP_{PM}}$ of all computation-paths of a probabilistic machine **PM**. One can easily show that:

$$\sum_i \left\{ p(\mathcal{P}_i) | \mathcal{P}_i \in \mathbf{P_{PM}} \right\} = \sum_i \left\{ p(\mathcal{CP}_i) | \mathcal{CP}_i \in \mathbf{CP_{PM}} \right\} = 1.$$

On this basis the concept of *computation* of a probabilistic state machine can be defined as follows.

Definition 4. *Computation of a probabilistic state machine.*
A *computation* of a probabilistic state machine **PM** with input (s_0, w_0) is the system of all computation-paths of **PM** with input (s_0, w_0).

Unlike the case of deterministic state machines, a computation of a probabilistic state machine does not yield a unique output. For any choice of a configuration-input (s_0, w_0), the computation-output is a system of possible configuration-outputs (s_{t+1}^i, w_{t+1}^i), where each (s_{t+1}^i, w_{t+1}^i) corresponds to a computation-path $\mathcal{CP}_i$. As expected, each (s_{t+1}^i, w_{t+1}^i) has a well-determined probability-value that is defined as follows:

$$p((s_{t+1}^i, w_{t+1}^i)) := \sum_i \left\{ p(\mathcal{CP}_i) | \text{the configuration-output of } \mathcal{CP}_i \text{ is } (s_{t+1}^i, w_{t+1}^i) \right\}.$$

One can easily show that the sum of the probability-values of all configuration-outputs of any machine **PM** is 1.

3. QUANTUM STATE MACHINES

The strong parallelism that characterizes quantum computers is based on two quantum-theoretic notions that have been often described as mysterious and potentially paradoxical: *superposition* and *entanglement*. For the readers who are not expert of quantum theory it is expedient to recall some concepts of the quantum formalism that are used in quantum computation[5]. The basic idea is that any *piece of quantum information* is mathematically represented as a possible *state* of a quantum system that can store and transmit the information in question. In the simplest situations one is dealing with a single particle S (say, an electron or a photon), whose "mathematical environment" is a special example of a vector space: the two-dimensional Hilbert space $\mathbb{C}^2$, based on the set of all ordered pairs of complex numbers. The canonical (orthonormal) basis of $\mathbb{C}^2$ consists of the two following unit-vectors:

$$|0\rangle = (1, 0); \quad |1\rangle = (0, 1),$$

---

[5]A survey of quantum computation theory can be found, for instance, in Nielsen and Chuang (2000).

which represent, in this framework, the two classical bits (0 and 1), or (equivalently) the two classical truth-values (*Falsity* and *Truth*). A *pure state* corresponds to a *maximal piece of information* that cannot be consistently extended to a richer knowledge. Such state is represented as a unit-vector $|\psi\rangle$ that can be expressed as a *superposition* of the two elements of the canonical basis of $\mathbb{C}^2$:

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle,$$

where c_0 and c_1 (also called *amplitudes*) are complex numbers such that $|c_0|^2 + |c_1|^2 = 1$.

The physical interpretation of $|\psi\rangle$ (also called *qubit-state* or, briefly, *qubit*) is the following: the physical system S in state $|\psi\rangle$ might satisfy the physical properties that are *certain* for the bit $|0\rangle$ with probability $|c_0|^2$ and might satisfy the physical properties that are *certain* for the bit $|1\rangle$ with probability $|c_1|^2$. Due to the characteristic indeterminism of quantum theory, the pure state $|\psi\rangle$ is at the same time a *maximal and logically incomplete piece of information* that cannot *decide* some important physical properties of the system S. Accordingly, from an intuitive point of view, one can say that $|\psi\rangle$ describes a kind of *cloud of potential properties* that might become *actual* when a measurement is performed. Measuring a physical quantity (by means of an apparatus associated to the canonical basis) determines a sudden transformation of the qubit $|\psi\rangle$ either into the bit $|0\rangle$ or into the bit $|1\rangle$. Such transformation is usually called *collapse of the wave-function*.

Not all states associated to a physical system S are pure. Non-maximal pieces of information can be represented as *mixtures of pure states* (special examples of operators called *density operators*). In the space $\mathbb{C}^2$ a density operator ρ can be represented as a convenient finite sum of projection-operators:

$$\rho = \sum_i w_i P_{|\psi_i\rangle},$$

where w_i are real numbers such that $\sum_i w_i = 1$, while each $P_{|\psi_i\rangle}$ is a projection-operator that projects along the direction of $|\psi_i\rangle$. Notice that such representation is not generally unique. A density operator that cannot be represented as a projection $P_{|\psi\rangle}$ is called a *proper mixture*. While pure states codify an essential indetermination of some relevant properties of the quantum system under investigation, mixtures may correspond to an *epistemic uncertainty* of the observer. Unlike pure states (which always satisfy some well-determined properties), there are mixtures that cannot decide any (non-trivial) property of the associated system. An example of this kind is the state $\rho = \frac{1}{2}\mathbb{I}$, where $\mathbb{I}$ is the identity operator of the space $\mathbb{C}^2$.

As happens in classical information theory, quantum computation also needs complex pieces of information, which are supposed to be stored by composite quantum systems (generally consisting of n subsystems). Accordingly, one can naturally adopt the quantum-theoretic formalism for the mathematical representation of composite physical systems, based on the use of *tensor products* (special examples of products)[6]. While a single

qubit is a unit-vector of the space $\mathbb{C}^2$, a pure state representing a complex piece of information can be identified with a unit-vector of the n-fold tensor product of $\mathbb{C}^2$:

$$\otimes^n\mathbb{C}^2 = \underbrace{\mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2}_{n-times} \text{ (with } n \geq 1).$$

Such vectors are called *quregisters*. The canonical basis of the space $\otimes^n\mathbb{C}^2$ consists of all *registers*, products of bits that have the following form:

$$|x_1\rangle \otimes \ldots \otimes |x_n\rangle \text{ (where any } x_i \text{ is either } 0 \text{ or } 1).$$

Instead of $|x_1\rangle \otimes \ldots \otimes |x_n\rangle$, it is customary to write $|x_1, \ldots, x_n\rangle$. Any quregister can be represented as a superposition of registers:

$$|\psi\rangle = \sum_i c_i|x_{i_1}, \ldots, x_{i_n}\rangle,$$

where c_i are complex numbers such that $\sum_i |c_i|^2 = 1$. A tensor product $|\psi_1\rangle \otimes \ldots \otimes |\psi_n\rangle$ (of n quregisters) is often briefly indicated by: $|\psi_1\rangle \ldots |\psi_n\rangle$.

Quantum computation makes essential use of some characteristic quantum states that are called *entangled*. In order to illustrate the concept of entanglement from an intuitive point of view, let us refer to a simple paradigmatic case. We are concerned with a composite physical system S consisting of two subsystems S_1 and S_2 (say, a two-electron system). By the quantum-theoretic rules that concern the mathematical description of composite systems, all states of S shall live in the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, where $\mathcal{H}_1$ and $\mathcal{H}_2$ are the Hilbert spaces associated to the systems S_1 and S_2, respectively. The observer has a *maximal information* about S: a *pure state* $|\psi\rangle$ of $\mathcal{H}$. What can be said about the states of the two subsystems? Due to the form of $|\psi\rangle$, such states cannot be pure: they are represented by two identical *mixtures*, which codify a "maximal degree of uncertainty." A typical possible form of $|\psi\rangle$ is the following *Bell-state*:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0, 0\rangle + |1, 1\rangle),$$

which lives in the space $\mathbb{C}^2 \otimes \mathbb{C}^2$, whose canonical basis consists of the four vectors $|0, 0\rangle, |0, 1\rangle, |1, 0\rangle, |1, 1\rangle$.

This gives rise to the following physical interpretation: the global system S might satisfy the properties that are *certain* either for the state $|0, 0\rangle$ or for the state $|1, 1\rangle$ with probability-value $\frac{1}{2}$. At the same time, $|\psi\rangle$ determines that the *reduced state* of both subsystems (S_1 and S_2) is the mixture $\frac{1}{2}\mathbb{I}$. Although it is not determined whether the state of the global system S is $|0, 0\rangle$ or $|1, 1\rangle$, the two subsystems S_1 and S_2 can be described

---

[6]The basic property of the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ of two (finite-dimensional) Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is the following: $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a Hilbert space that properly includes an isomorphic image of the Cartesian product $\mathcal{H}_1 \times \mathcal{H}_2$ (consisting of all ordered pairs of vectors that belong to the spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively). Furthermore, $\mathcal{H}_1 \otimes \mathcal{H}_2$ contains all possible superpositions of its elements. A vector $|\psi\rangle$ of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is called *factorized* iff $|\psi\rangle$ corresponds to a pair $(|\psi_1\rangle, |\psi_2\rangle) \in \mathcal{H}_1 \times \mathcal{H}_2$. In such a case, it is customary to write: $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$. Of course, not all vectors of $\mathcal{H}_1 \otimes \mathcal{H}_2$ are factorized.

as "entangled," because in both possible cases they would satisfy the same properties, turning out to be *indistinguishable*. As a consequence, any measurement performed by an observer either on system S_1 or on system S_2 would instantaneously transform the *potential* properties of both subsystems into *actual* properties (by *collapse of the wave-function*).

The celebrated "Einstein–Podolsky–Rosen paradox"(*EPR*) is based on a similar physical situation. As is well-known, what mainly worried Einstein was the possibility of "non-local effects:" the subjective decision of an observer (who may choose among different *incompatible observables* to be measured on the system S_1) seems to determine the instantaneous emergence of an actual property for the system S_2, which might be very "far" from S_1 (possibly inaccessible by means of a light-signal). Interestingly enough, in the framework of quantum computation, entangled states have been often used as a powerful resource, even from a technological point of view (for instance, in the applications to teleportation-phenomena and to quantum cryptography).

As expected, quantum computation cannot be identified with a "static" representation of pieces of information. What is important is the dynamic *process* of information that gives rise to quantum computations. Such process is mathematically performed by *quantum logical gates* (briefly, *gates*): special examples of *unitary operators* that transform quregisters into quregisters in a reversible way. Since in quantum theory the time-evolution of all physical systems is mathematically described by unitary operators, one can say that quantum computations can be regarded as the time-evolution of some special quantum objects.

We will now introduce the definition of *quantum state machine*, which represents a quantum counterpart of the classical notion of *deterministic state machine*. From an intuitive point of view, any quantum state machine can be regarded as a kind of quantum superposition of many classical deterministic state machines. Some definitions of *quantum Turing machine* discussed in the literature are based on a strong idealization: no limit is assumed for the length of the registers occurring in a computation. This corresponds to the classical assumption according to which a Turing machine is equipped with an infinite tape. We will consider a more realistic concept, closer to physical quantum computers, which are of course always bound to a limited memory.

Definition 5. *Quantum state machine.*
A *quantum state machine* is an abstract system **QM** associated to a (finite-dimensional) Hilbert space $\mathcal{H}^{\mathbf{QM}}$ whose unit-vectors $|\psi\rangle$ represent possible pure states of a quantum system that could physically implement the computations of the state machine. The space $\mathcal{H}^{\mathbf{QM}}$ has the following form:

$$\mathcal{H}^{\mathbf{QM}} = \mathcal{H}^H \otimes \mathcal{H}^S \otimes \mathcal{H}^W.$$

The following conditions are required:

1. $\mathcal{H}^H$ (which represents the halting-space) is the space $\mathbb{C}^2$, where the two elements of the canonical basis ($\{|0\rangle_H, |1\rangle_H\}$) correspond to the states "the machine does not halt" and "the machine halts," respectively.
2. $\mathcal{H}^S$ (which represents the internal-state space) is associated to a finite set $\mathcal{S}$ of classical internal states. We require that $\mathcal{H}^S =$

$\otimes^m \mathbb{C}^2$, where 2^m is the cardinal number of $\mathcal{S}$. Accordingly, the set $\mathcal{S}$ can be one-to-one associated to a basis of $\mathcal{H}^S$.
3. $\mathcal{H}^W$ (which represents the word-space) is identified with a Hilbert space $\otimes^n \mathbb{C}^2$ (for a given $n \geq 1$). The number n determines the length of the registers $|x_1, \ldots, x_n\rangle$ that may occur in a computation. Shorter registers $|x_1, \ldots, x_h\rangle$ (with $h < n$) can be represented in the space $\otimes^n \mathbb{C}^2$ by means of convenient *ancillary bits*.

Let $B^{\mathbf{QM}}$ be a basis of $\mathcal{H}^{\mathbf{QM}}$, whose elements are unit-vectors having the following form:

$$|\varphi_i\rangle = |h_i\rangle|s_i\rangle|x_{i_1}, \ldots, x_{i_n}\rangle,$$

where $|h_i\rangle$ belongs to the basis of $\mathcal{H}^H$, while $|s_i\rangle$ belongs to the basis of $\mathcal{H}^S$.

Any unit-vector $|\psi\rangle$ of $\mathcal{H}^{\mathbf{QM}}$ that is a superposition of basis-elements $|\varphi_i\rangle$ represents a possible *computational state* of **QM**. The expected interpretation of a computational state

$$|\psi\rangle = \sum_i c_i |h_i\rangle|s_i\rangle|x_{i_1}, \ldots, x_{i_n}\rangle$$

is the following:

- the machine in state $|\psi\rangle$ might halt with probability $|c_i|^2$ (if $|h_i\rangle = |1_H\rangle$) or with probability $1 - |c_i|^2$ (if $|h_i\rangle = |0_H\rangle$).
- the machine in state $|\psi\rangle$ might correspond to the classical configuration $(s_i, (x_{i_1}, \ldots, x_{i_n}))$ with probability $|c_i|^2$. Hence, the state $|\psi\rangle$ describes a kind of *quantum co-existence* of different classical deterministic configurations.

4. The set of *possible inputs* of **QM** is identified with the set of all computational states that have the following form:

$$|\psi\rangle = \sum_i |0_H\rangle|s_{in}\rangle|x_{i_1}, \ldots, x_{i_n}\rangle.$$

5. Like a deterministic state machine, a quantum state machine **QM** is characterized by a *program*. In the quantum case, a program is identified with a sequence of unitary operators of $\mathcal{H}^{\mathbf{QM}}$:

$$(U_0, \ldots, U_t),$$

where we may have: $U_i = U_j$ with $i \neq j$.
The following conditions are required:

(a) for any possible input $|\psi_0\rangle$, $U_0(|\psi_0\rangle) = |\psi_1\rangle$ is a superposition of basis-elements having the following form:

$$|h_i^1\rangle|s_i^1\rangle|x_{i_1}^1, \ldots, x_{i_n}^1\rangle,$$

where all s_i^1 are different from s_{in} and $|h_i^1\rangle = |0_H\rangle$, if $t \neq 0$.
(b) For any j ($0 < j < t$), $U_j(|\psi_j\rangle) = |\psi_{j+1}\rangle$ is a superposition of basis-elements having the following form:

$$|0_H\rangle|s_i^{j+1}\rangle|x_{i_1}^{j+1}, \ldots, x_{i_n}^{j+1}\rangle.$$

(c) $U_t(|\psi_t\rangle) = |\psi_{t+1}\rangle$ is a finite superposition of basis-elements having the following form:

$$|1_H\rangle|s_{halt_j}\rangle|x_{i_1}^{t+1},\ldots,x_{i_n}^{t+1}\rangle.$$

The concept of *computation* of a quantum state machine can be now defined in a natural way.

Definition 6. *Computation of a quantum state machine.*
Let **QM** be a quantum state machine, whose program is the operator-sequence $(U_0,\ldots,U_t)$ and let $|\psi_0\rangle$ be a possible input of **QM**. A *computation* of **QM** with input $|\psi_0\rangle$ is a sequence of computational states of **QM**

$$\mathcal{QC} = (|\psi_0\rangle,\ldots,|\psi_{t+1}\rangle),$$

such that: $|\psi_{i+1}\rangle = U_i(|\psi_i\rangle)$, for any i ($0 \leq i \leq t$).
The vector $|\psi_{t+1}\rangle$ represents the output of the computation, while the density operator $Red^3(|\psi_{t+1}\rangle)$ (the reduced state of $|\psi_{t+1}\rangle$ with respect to the third subsystem) represents the word-output of the computation.

Like all abstract notions of *quantum computer*, the concept of *quantum state machine* gives rise to some critical questions that have been often discussed in the literature. Two important problems (which cannot have any counterpart in the case of classical computation) are the following:

- How shall we interpret the operation of "reading the output" of a computation of a given machine? What is the role of the *collapse of the wave-function* during a reading-action?
- Is it possible to measure the halting state without disturbing the configuration-state?

Consider now a quantum state machine whose program is

$$(U_0,\ldots,U_t).$$

Each U_i naturally determines a corresponding word-operator U_i^W, defined on the word-space $\mathcal{H}^W$. Generally, it is not guaranteed that all word-operators are unitary. But it is convenient to refer to quantum state machines that satisfy this condition. In this way, any quantum state machine (whose word-space is $\otimes^n\mathbb{C}^2$) determines a *quantum circuit*, consisting of a sequence of unitary operators (*gates*):

$$(U_0^W,\ldots,U_t^W),$$

where n represents the *width*, while $t+1$ represents the *depth* of the circuit.

To what extent can quantum state machines be simulated by classical probabilistic state machines? In order to discuss this important question, let us refer to a celebrated quantum experiment, based on the *Mach–Zehnder interferometer* (represented by **Figure 2**).

The physical situation can be sketched as follows. Consider a photon-beam (possibly consisting of a single photon) and assume that $|0\rangle$ describes the state of photons moving along the x direction, while $|1\rangle$ describes the state of photons moving along the y direction. All photons go through a first *beam splitter*
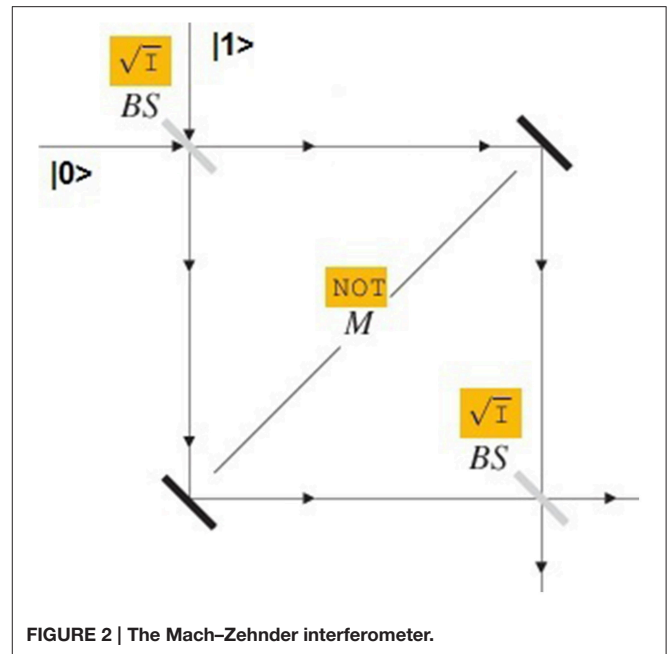

FIGURE 2 | The Mach–Zehnder interferometer.

that "splits" them giving rise to the following effect: within the box each photon follows a path corresponding either to the x-direction or to the y-direction with probability $\frac{1}{2}$. Soon after, on both paths, all photons are reflected by a *mirror* that inverts their direction. Finally, the photons pass through a second beam splitter that determines the output-state. Suppose that all photons entering into the interferometer-box are moving in the x-direction. According to a "classical way of thinking" we would expect that the photons detected at the end of the process will move either along the x-direction or along the y-direction with probability $\frac{1}{2}$. The result of the experiment is, instead, completely different: the Mach–Zehnder interferometer *always* transforms the input-state $|0\rangle$ into the output-state $|0\rangle$; while the input-state $|1\rangle$ is transformed into $|1\rangle$.

From a mathematical point of view, such a "surprising" result can be explained by using, in an essential way, the concept of superposition. The apparatuses (used in the Mach–Zehnder experiment) can be mathematically represented by two important gates. A beam splitter can be regarded as a physical implementation of the *Hadamard-gate* $\sqrt{I}$ (also called *square root of identity*), which is defined as follows (on the canonical basis of $\mathbb{C}^2$):

$$\sqrt{I}|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle); \quad \sqrt{I}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

Apparently, the Hadamard-gate transforms the two classical bits $|0\rangle$ and $|1\rangle$ into two (different) genuine superpositions. As a consequence, within the Mach–Zehnder box a photon in state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ turns out to satisfy at the same time two alternative properties: the property of moving along the x-direction and the property of moving along the y-direction. We have here a characteristic quantum parallelism: a single photon "goes along"

FIGURE 3 | The quantum skier.

FIGURE 4 | A word-graph for a "classical probabilistic Mach–Zehnder state machine."

two different paths at the same time! Metaphorically, situations of this kind have been sometimes compared to the puzzling behavior of a "quantum skier" who runs at the same time on the left and on the right side of a given tree (see **Figure 3**).

The second apparatus of the Mach–Zehnder interferometer (the mirror), can be regarded as a physical implementation of another important gate, the *negation* NOT (a quantum generalization of the classical negation), which is defined as follows:

$$\text{NOT}|0\rangle = |1\rangle; \quad \text{NOT}|1\rangle = |0\rangle.$$

Accordingly, the *Mach–Zehnder circuit* can be identified with the following sequence of three gates (all defined on the space $\mathbb{C}^2$):

$$(\sqrt{\text{I}}, \text{NOT}, \sqrt{\text{I}}).$$

Let us now apply the Mach–Zehnder circuit to the input $|0\rangle$. We obtain:

$$\sqrt{\text{I}} : |0\rangle \mapsto \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle); \quad \text{NOT} : \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \mapsto \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle); \quad \sqrt{\text{I}} : \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \mapsto |0\rangle.$$

We can see, in this way, how the Mach–Zehnder circuit transforms the input-state $|0\rangle$ into the output-state $|0\rangle$. In a similar way, the input-state $|1\rangle$ is transformed into the output-state $|1\rangle$.

Is there any natural "classical counterpart" for the Hadamard-gate? A natural candidate might be a particular example of a probabilistic state machine that we can conventionally call *the classical probabilistic NOT-state machine* (**PM**$^{\text{NOT}}$). Such machine can be defined as follows:

- The set of possible word-inputs of **PM**$^{\text{NOT}}$ is the set of words $\{(0), (1)\}$.
- The program of **PM**$^{\text{NOT}}$ consists of the following sequence of rules:

$$Seq_0 = (R_{0_1}, R_{0_2}),$$

where:

$R_{0_1} : (s_{in}, (x)) \mapsto (s_{halt_j}, (x))$ and $p(R_{0_1}) = \frac{1}{2}$;

$R_{0_2} : (s_{in}, (x)) \mapsto (s_{halt_j}, (1 - x))$ and $p(R_{0_2}) = \frac{1}{2}$.

Consider, for instance, the input $(s_{in}, (0))$. The output will be the following set:

$$\left\{ (s_{halt_j}, (0)), (s_{halt_j}, (1)) \right\}.$$

On this basis, a "classical probabilistic Mach–Zehnder state machine" would determine (for the word-input (0)) the word-graph illustrated by **Figure 4**.

Such a machine turns out to compute both the words (0) and (1) with probability $\frac{1}{2}$. Interestingly enough, this is the same probabilistic result that is obtained in the quantum case, when one performs a measurement inside the interferometer-box. In such a case, photons behave like "normal skiers," who pass *either* at the right *or* at the left side of a tree (where *or* represents here, of course, the exclusive disjunction).

The arguments we have developed seem to confirm the following conjecture: the characteristic superposition-patterns, that may occur during a quantum computation (when no measurement is performed during the computation-process), cannot be generally represented by probabilistic state machines. Quantum parallelism (based on superpositions) and classical parallelism are deeply different.

4. QUANTUM PARALLELISM, PSYCHOLOGICAL PARALLELISM, AND QUANTUM COMPUTATIONAL SEMANTICS

What kind of similarity can be recognized between quantum parallel structures and different forms of psychological parallelism? Trying to represent the human mind as a kind of system of quantum state machines would be, of course, naive and misleading. In spite of many important results in the framework of neurosciences, the complex network that connects

human conscious and unconscious thoughts is still quite mysterious[7]. Quantum-like superpositions can be reasonably applied to represent some aspects of such complex networks. Even quantum *interference phenomena* (with the characteristic constructive and destructive effects) can find some natural psychological interpretations.

According to an interesting hypothesis (discussed by the neuroscientist Boncinelli, 2012), the mysterious emergence of an *act of consciousness* can be represented as a sudden transition from a parallel structure to a linear one. Is it reasonable to conjecture that such transition could be described as a kind of "psychological collapse of the wave-function?"

In the investigations about possible links between *quantum structures* and *psychological structures* a useful tool is represented by a special form of quantum logical semantics (called *quantum computational semantics*) that has been naturally inspired by the theory of quantum computation[8].

Let us briefly recall the basic ideas of this semantics. We can refer a first-order language $\mathcal{L}$, whose non-logical alphabet contains individual terms (variables and names), predicates and sentential constants. Interpreting the language $\mathcal{L}$ means associating to any formula α a *meaning*, identified with a piece of quantum information that can be stored by a quantum system. Accordingly, any possible meaning of α is represented by a possible (pure or mixed) state of a quantum system: generally, a density operator ρ_α that lives in a Hilbert space $\mathcal{H}^\alpha$, whose dimension depends on the linguistic complexity of α.

The logical operators of $\mathcal{L}$ are associated to special examples of Hilbert-space operations that have a characteristic dynamic behavior, representing possible computation-actions. The logical connectives are interpreted as particular (reversible) gates, like the negation NOT, the Hadamard-gate $\sqrt{I}$, the Toffoli-gate T (which allows us to define a reversible conjunction AND). At the same time, the logical quantifiers ($\forall$, $\exists$) are interpreted as possibly irreversible quantum operations. Since the universe of discourse (which the language refers to) may be indeterminate, the use of quantum quantifiers may give rise to a reversibility-breaking, which is quite similar to what happens in the case of measurement-phenomena.

Due to the characteristic features of quantum holism, meanings turn out to behave in a *holistic* and *contextual* way: the density operator ρ_α (which represents the *global* meaning of a formula α) determines the *contextual meanings* of all parts of α (which can be obtained by applying the *reduced-state function* to ρ_α). As a consequence, it may happen that the meaning of a formula is an entangled pure state, while the meanings of its parts are proper mixtures. In such cases, the meaning of a global expression turns out to be more precise than the meanings of its parts. It is also admitted that one and the same formula receives different contextual meanings in different contexts.

As an example, consider the atomic sentence "Alice is pretty" (formalized as **Pa**). In order to store the information expressed by this sentence, we need three quantum objects whose states

represent the pieces of information corresponding, respectively, to the predicate **P**, to the name **a** and to the truth-degree according to which the individual denoted by the name **a** satisfies the property denoted by the predicate **P**. Accordingly, the meaning of the sentence **Pa** can be identified with a (pure or mixed) state $\rho_{\mathbf{Pa}}$ living in the tensor-product space $\mathcal{H}^{\mathbf{Pa}} = \otimes^3 \mathbb{C}^2$. In order to obtain the contextual meanings of the linguistic parts of **Pa** it is sufficient to consider the two reduced states $Red^1(\rho_{\mathbf{Pa}})$ and $Red^2(\rho_{\mathbf{Pa}})$, which describe (respectively) the states of the first and of the second subsystem of the quantum object that stores the information expressed by the sentence **Pa**. From a logical point of view, $Red^1(\rho_{\mathbf{Pa}})$ and $Red^2(\rho_{\mathbf{Pa}})$ can be regarded as two *intensional meanings*: a property-concept and an individual concept, respectively; while $\rho_{\mathbf{Pa}}$ represents a *propositional concept* (or *event*).

Like formulas, sequences of formulas also can be interpreted according to the quantum computational rules. As expected, a possible meaning of the sequence $(\alpha_1, \ldots, \alpha_n)$ will be a density operator $\rho_{(\alpha_1,\ldots,\alpha_n)}$ living in a Hilbert space $\mathcal{H}^{(\alpha_1,\ldots,\alpha_n)}$, whose dimension depends on the linguistic complexity of the formulas $\alpha_1, \ldots, \alpha_n$.

In this framework one can develop an abstract theory of *vague possible worlds*. Consider a pair

$$W = ((\alpha_1, \ldots, \alpha_n),\ \rho_{(\alpha_1,\ldots,\alpha_n)}),$$

consisting of a sequence of formulas and of a density operator that represents a possible meaning for our sequence. It seems reasonable to assume that W describes a *vague possible world*, a kind of *abstract scene* where most events are characterized by a "cloud of ambiguities," due to quantum uncertainties. In some cases W might be exemplified as a "real" scene of a theatrical play or as a vague situation that is described either in a novel or in a poem. And it is needless to recall how ambiguities play an essential role in literary works.

As an example, consider the following vague possible world:

$$W = ((\mathbf{Pab}), \rho_{(\mathbf{Pab})}),$$

where **Pab** is supposed to formalize the sentence "Alice is kissing Bob," while $\rho_{\mathbf{Pab}}$ corresponds to the pure state

$$|\Psi\rangle_{\mathbf{Pab}} = |\varphi\rangle \otimes \frac{1}{\sqrt{2}}(|0, 1\rangle) + |1, 0\rangle) \otimes |1\rangle,$$

where $|\varphi\rangle$ lives in the space $\mathbb{C}^2$, while $|\Psi\rangle_{\mathbf{Pab}}$ lives in the space $\otimes^4 \mathbb{C}^2$. Here the reduced state of $|\Psi\rangle_{\mathbf{Pab}}$ that describes the pair (Alice, Bob) has the typical form of an entangled state; consequently, the states describing the two individuals Alice and Bob are two identical mixed states. In the context $|\Psi\rangle_{\mathbf{Pab}}$ Alice and Bob turn out to be indistinguishable: it is not determined "who is who" and "who is kissing whom." It is not difficult to imagine some "real" theatrical scenes representing ambiguous situations of this kind.

5. A QUANTUM SEMANTICS FOR MUSIC

An abstract version of the quantum computational semantics can be applied to a formal analysis of musical compositions, where

---

[7] As is well-known, the literature devoted to the study of parallel structures in the mind/brain-behavior is very rich. As an example, one can refer to some important contributions of Damasio (see, for instance, Damasio, 1999).

[8] See (Dalla Chiara et al., 2005, 2010, in press).

both *musical ideas* and *extra-musical meanings* are generally characterized by some essentially vague and ambiguous features[9].

Any musical composition (say, a sonata, a symphony, an opera,...) is, generally, determined by three elements:

- a *score*;
- a set of *performances*;
- a set of musical *thoughts* (or *ideas*), which represent possible *meanings* for the *musical phrases* written in the score.

While scores represent the syntactical component of musical compositions, performances are physical events that occur in space and time. From a logical point of view, we could say that performances are, in a sense, similar to *extensional meanings*, i.e., well-determined systems of objects which the linguistic expressions refer to.

Musical thoughts (or ideas) represent, instead, a more mysterious element. Is it reasonable to assume the existence of such ideal objects that are, in a sense, similar to the *intensional meanings* investigated by logic? Is there any danger to adhere, in this way, to a form of *Platonism*? When discussing semantic questions, one should not be "afraid" of Platonism. In the particular case of music, a composition cannot be simply reduced to a score and to a system of sound-events. Between a score (which is a system of signs) and the sound-events created by a performance there is something intermediate, represented by the musical ideas that underlie the different performances. This is the abstract environment where normally live both composers and conductors, who are accustomed to study scores without any help of a material instrument.

Following the rules of the quantum semantics, *musical ideas* can be naturally represented as superpositions that ambiguously describe a variety of co-existent thoughts. Accordingly, we can write:

$$|\mu\rangle = \sum_i c_i |\mu_i\rangle,$$

where:

- $|\mu\rangle$ is an abstract object representing a musical idea that *alludes* to other ideas $|\mu_i\rangle$ (possible *variants* of $|\mu\rangle$ that are, in a sense, all co-existent);
- the number c_i measures the "weigth" of the component $|\mu_i\rangle$ in the context $|\mu\rangle$.

As happens in the case of composite quantum systems, musical ideas (which represent possible meanings of *musical phrases* written in a score) have an essential *holistic* behavior: the meaning of a global musical phrase determines the *contextual meanings* of all its parts (and not the other way around).

An important feature of music is the capacity of *evoking* extra-musical meanings: subjective feelings, situations that are vaguely imagined by the composer or by the interpreter or by the listener, real or virtual theatrical scenes (which play an essential role in the case of lyric operas and of *Lieder*). The interplay between musical ideas and extra-musical meanings can be naturally represented in the framework of our quantum semantics, where extra-musical

meanings can be dealt with as special examples of vague possible worlds.

We can refer to the abstract tensor product of two spaces

$$MSpace \otimes WSpace,$$

where:

- *MSpace* represents the space of musical ideas $|\mu\rangle$.
- *WSpace* represents the space of vague possible worlds, dealt with as special examples of abstract objects $|w\rangle$ that can be evoked by musical ideas.

Following the quantum-theoretic formalism, we can distinguish between *factorized* and *non-factorized* global musical ideas. A factorized global musical idea will have the form:

$$|M\rangle = |\mu\rangle \otimes |w\rangle.$$

But we might also meet entangled global musical ideas, having the form:

$$|M\rangle = c_1(|\mu_1\rangle \otimes |w_1\rangle) + c_2(|\mu_2\rangle \otimes |w_2\rangle).$$

As is well-known, music gives rise to a special kind of psychological experience, where some complex parallel structures are consciously grasped, in a way that may appear miraculous. Paradigmatic examples arise, for instance, in the case of trios or quartets of lyric operas. In such cases, the listener perceives a *global polyphonic structure*; at the same time, he/she is able to follow (at least to a certain extent) the different melodic lines and even the different thoughts and feelings of the characters who are singing. As an example, it may be interesting to consider three great masterpieces of the history of lyric operas: the quartet of Act 1 in Beethoven's *Fidelio*, the quartet of Act 3 in Verdi's *Rigoletto* and the trio of Act 3 of *Der Rosenkavalier* by Richard Strauss. The parallel structures that arise in these three examples have some significant differences both from the musical and from the semantic point of view.

In *Fidelio*'s quartet the psychological contraposition between the four characters (Marzelline, Leonore, Rocco, Jaquino) is realized by means of a single musical theme that is successively sung by the four singers (**Figure 5**).

It is amazing how Beethoven succeeds in expressing, by one and the same theme, different attitudes and emotions: the joyful hope of Marzelline, the doubts and the anguish of Leonore, the paternal satisfaction of Rocco, the jealous rage of Jaquino. The whole context is dominated by strong ambiguities and antagonistic elements: the contrast between an improbable family-portrait and the cruel jail-environment, the contradictions of Rocco (who is at the same time a fond father and an accomplice of the prison-system), the sexual ambiguity of Leonore, the loving heroin who has disguised herself as a man (Fidelio), in the attempt to save her husband, the prisoner Florestan. The musical result is an extraordinary and highly emotional polyphonic construction based on very simple musical components.

The structure of *Rigoletto*'s quartet is completely different. All characters are associated to specific musical themes that

---

[9]See (Dalla Chiara et al., 2012).

FIGURE 5 | The quartet-theme.

FIGURE 6 | The Duke.

are repeated with some variations. The leading musical idea is represented by the wonderful theme sung by the Duke of Mantova at the very beginning (**Figure 6**)[10].

Like Mozart's *Don Giovanni*, Verdi's Duke is a cynic seducer, who may appear sweet and sincere with his victims. And music often exalts a paradoxical co-existence of contradictory

psychological attitudes. All contrasts are emphasized in the quartet by the sordid environment, where a crime is going to be committed. Maddalena's answer to the Duke is based on a fully different theme, a *staccato*-sequence of sixteenth-notes (**Figure 7**)[11].

Both the music and the text reflect Maddalena's ambiguity: she is a prostitute who is playing a traditional seductive role;

---

[10] *Fairest daughter of love, I am a slave of your charms; with but a single word you could relieve my every pain. Come touch my breast and feel how my heart is racing. With but a single word you could relieve my every pain.*

[11] *Ah! Ah! That really makes me laugh, talk like that is cheap enough.*

FIGURE 7 | Maddalena.

FIGURE 8 | Gilda.

at the same time she is also instrumental to a murder-project. Gilda's entrance (soon after Maddalena's first phrase) determines a sudden dramatic change. What Gilda sings is a cry of sorrow, interrupted by some short pauses and *appoggiaturas* that seem to describe desperate sobs (**Figure 8**)[12].

One has often discussed the reasons that may have led Gilda to her unreasonable sacrifice for an unworthy man who had deceived her. Representing Gilda as a naive and modest girl is, however, misleading and in contrast with the greatness expressed by the music. Gilda's death-choice can be perhaps better understood as a suicide, caused by an unendurable disillusion. Rigoletto's role in the quartet is musically less "visible." His mind

is completely absorbed in the vengeance-project ("la vendetta") that shall be shortly accomplished. From a musical point of view, the quartet is constructed as a polyphonic structure, where the four voices are interlaced, each preserving its own musical, semantic and psychological autonomy.

Der Rosenkavalier by Strauss belongs to a musical and literary world that is somewhat far both from *Fidelio* and from *Rigoletto*. Different forms of ambiguity are exalted in this opera, which is characterized by an extraordinary unity of music and text, written by the great poet Hugo von Hofmannsthal. The theme of sexual ambiguity is here developed by the character of Octavian, the *Rosenkavalier* whose role is sung by a mezzo-soprano. Although Octavian may recall Mozart's Cherubino, ambiguities are in Strauss' opera more sophisticated: in two different situations Octavian disguises himself as a woman in order to make fun

---

[12]*Ah, these are the loving words the scoundrel spoke once to me! O wretched heart betrayed, do not break of sorrow.*

of the rude fiancé of the fascinating girl Sophie. Interestingly enough, some interpreters of the role of Octavian have told how often they have been puzzled by their "oscillating identity" during the opera's performance.

A different and deeper "identity-question" is evoked in a splendid aria sung by the *Marschallin* in Act 1. After a passionate night spent with her lover Octavian, the lady is troubled by some sad thoughts about the flowing of time and the mysterious co-existence of different identities of one and the same person in different stages of life. She sings:

> *Aber wie kann das wirklich sein,*
> *dass ich die kleine Resi war,*
> *und dass ich einmal die alte Frau sein werd'*
> ....
> *Wie kann denn das geschehen?*
> *Wie macht denn das der liebe Gott?*
> *Wo ich doch immer die gleiche bin.*
> *Und wenn er's schon so machen muss,*
> *warum lasst er mich denn zuschaun dabei*
> *mit gar so klarem Sinn?*
> *Warum versteckt er's nicht vor mir?*
> *Das alles ist geheim, so viel geheim* [13].

One is dealing with an extraordinary poetic and musical representation of a "hard" scientific and philosophical problem, that modern philosophers of science usually call "the genidentity-question"[14].

---

[13] *But how can it be that I was the little Resi and that I shall be the old lady. .... How can it come to pass? How can God decree it so? While, in fact, I am always the same. And if indeed it must be so, why does he let me look at it so clearly? Why does he not hide it to me? All this is a mystery, a great mystery.*

[14] The term "genidentity," which refers to the problematic *identity of individuals through time* has been introduced by Lewin (in his doctoral thesis in 1922) and

The trio performed at the end of the opera by three female voices (the Marschallin, Sophie, Octavian) is a wonderful polyphonic construction, where the three characters express different thoughts and feelings, which are not generally associated to some specific musical themes (unlike the case of *Rigoletto*'s quartet). The main theme is sung at the very beginning by the Marschallin (**Figure 9**)[15].

By this deeply moving musical phrase the Marschallin expresses her extreme act of love, which is to renounce love. Her choice might recall what Violetta Valery sings in Verdi's *La Traviata*:

> *Dite alla giovine sì bella e pura*[16]

although Violetta and the Marschallin are, of course, completely different characters.

Sophie's entrance in the trio is, in a sense, surprising. She joins in, in the final part of the Marschallin's first phrase, just upon the critical word "andern" ("other"). Her intervention creates a sudden brief dissonance (a minor-second chord), which immediately disappears when the two womem (who are both in love with Octavian) harmonically conclude the phrase at a distance of a minor-third. What Sophie perceives is a strange religious atmosphere that she cannot really understand, since she is not aware of the *liason* between Octavian and the Marschallin. The *incipit* of the main theme (the characteristic imprinting of the whole trio) is then immediately transposed to a different key (from *D* flat major to *A* major) by Octavian, whose initial attitude seems to be mainly dominated by embarassing doubts and questions. But finally the reasons of love prevail over all doubts. At the end of the trio, while

---

has been further investigated by Reichenbach and many other scholars. See, for instance, Reichenbach (1928).

[15] *I promised to love him in the right way, even to love his love for another woman.*

[16] *Tell the beautiful and pure girl.*

FIGURE 9 | The Marschallin.

the two young lovers sing an expected "dich habe ich lieb" ("I love you"), the Marschallin concludes with an enigmatic phrase:

als wie halt Männer das glücklich sein verstehen[17].

singing the last note alone over a perfect tonic chord.

The three examples of polyphonic constructions, created by Beethoven, Verdi, and Strauss, are all characterized by strong unitary conceptions, based on complex parallel networks of harmonic, melodic, timbric, and semantic relationships (which have been extensively analyzed in musicological literature[18]). At the same time, one can easily recognize some significant differences that distinguish the three cases, both from the musical and from the semantic point of view. The structure of Fidelio's quartet is very close to a *canon-form*, where the entrance of each voice is associated to a specific semantic connotation. Rigoletto's quartet is, instead, dominated by strong musical contrasts that reflect the conflicting feelings of four human beings, living in a highly dramatic situation. Finally, Strauss' trio seems to propose a kind of musical and semantic "peaceful resolution." The trio is perceived by the listener as a strongly unitary musical idea that evolves in time. The three female voices are in a sense "entangled," sometimes creating the illusion that a single voice is singing

---

[17] *as far as men can understand happiness.*

[18] See, for instance, Budden (1983), Solomon (1998), Principe (2004).

(as happens in the case of some entangled quantum objects, whose parts are indistinguishable). Such musical situations can be naturally represented in the framework of the quantum musical semantics, where *musical thoughts* are dealt with as holistic ideal objects that vaguely allude to a (possibly infinite) variety of co-existing ideas.

The analysis proposed in this article has concerned questions that belong to worlds apparently "far apart": the theory of quantum computers, psychology, logical semantics, and music. A common pattern that arises in all these fields is a frequent and sometimes essential emergence of some characteristic parallel structures. We have seen how the quantum-theoretic concepts of *superposition* and *entanglement* have inspired the development of a "bridge-theory" (based on the quantum computational semantics) that can be usefully applied to a formal representation of different kinds of phenomena where parallelism plays a relevant role.

ACKNOWLEDGMENTS

REFERENCES

Aerts, D., and Gabora, L. (2005a). A theory of concepts and their combinations I: the structure of the sets of contexts and properties. *Kybernetics* 34, 167–191. doi: 10.1108/03684920510575799

Aerts, D., and Gabora, L. (2005b). A theory of concepts and their combinations II: a Hilbert space representation. *Kybernetics* 34, 192–221. doi: 10.1108/03684920510575807

Aerts, D., and Sozzo, S. (2014). Quantum entanglement in conceptual combinations. *Int. J. Theor. Phys.* 53, 3587–3603. doi: 10.1007/s10773-013-1946-z

Boncinelli, E. (2012). *Mi Ritorno in Mente. Il Corpo, le Emozioni, la Coscienza.* Milano: Longanesi.

Budden, J. (1983). *The Operas of Verdi.* New York, NY: Oxford University Press.

Dalla Chiara, M. L., Giuntini, R., Ledda, A., Leporini, R., and Sergioli, G. (2010). Entanglement as a semantic resource. *Found. Phys.* 40, 1494–1518. doi: 10.1007/s10701-010-9407-5

Dalla Chiara, M. L., Giuntini, R., and Leporini, R. (2005). Logics from quantum computation. *Int. J. Quantum Inf.* 3, 293–337. doi: 10.1142/S0219749905000943

Dalla Chiara, M. L., Giuntini, R., Leporini, R., and Sergioli, G. (in press). Holistic logical arguments in quantum computation. *Math. Slov.*

Dalla Chiara, M. L., Giuntini, R., Luciani, A. R., and Negri, E. (2012). *From Quantum Information to Musical Semantics.* Illustrated by C. Seravalli. London: College Publications.

Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotions in Making of Consciousness.* New York, NY: Harcourt Brace & Company.

Fouché, W., Heidema, J., Jones, G., and Potgieter, P. H. (2007). Deutsch's universal quantum Turing machine (Revisited). arXiv:quant-ph/0701108. Available online at: http://arxiv.org/pdf/quant-ph/0701108v1.pdf

Gudder, S. (1999). Quantum automata: an overview. *Int. J. Theor. Phys.* 38, 2261–2282. doi: 10.1023/A:1026663432352

Nielsen, M. A., and Chuang, I. L. (2000). *Quantum Computation and Quantum Information.* New York, NY: Cambridge University Press.

Penrose, R. (1994). *Shadows of the Mind.* Oxford: Oxford University Press.

Principe, Q. (2004). *Strauss. La Musica Nello Specchio di Eros.* Milano: Bompiani.

Reichenbach, H. (1928). *Philosophie der Raum-Zeit-Lehre.* Berlin-Leipzig: De Gruyter.

Savage, J. E. (1998). *Models Of Computation: Exploring the Power of Computing.* Reading: Addison Wesley.

Solomon, M. (1998). *Beethoven.* London: Schirmer Trade Books.

CrossMark

Two types of potential functions and their use in the modeling of information: two applications from the social sciences

Emmanuel E. Haven *

School of Management, Institute for Quantum Social and Cognitive Science and Institute of Finance, University of Leicester, Leicester, UK

In this paper we consider how two types of potential functions, the real and quantum potential can be shown to be of use in a social science context. The real potential function is a key ingredient in the Hamiltonian framework used in both classical and quantum mechanics. The quantum potential however emerges in a different way in quantum mechanics. In this paper we consider both potentials and we attempt to give them a social science interpretation within the setting of two applications.

Keywords: potential functions, quantum mechanics

1. Introduction

Potential functions are to physicists what utility functions are to economists: they are both examples of fundamental workhorse tools. But can there exist some connection? Utility functions, u, are defined as: $u : C \to \mathbb{R}$, where C is a set of objects. A preference relation on two objects, x and y such that, say, $x \succ y$ will imply that $u(x) > u(y)$ under the necessary conditions that the preference relation is transitive and asymmetric. This paper will not pretend to be at a level of rigor which has been attained in economics based preference theory. Examples of such rigor abound in the various expected utility frameworks, and some papers in this special issue will be devoted to probing how deviations from central axioms like the sure-thing principle can be explained with the aid of quantum structures. Our objective in this paper is modest: we would like to inform (and maybe convince) the reader that with the help of two types of potential functions we can model, in a reasonably successful way, information. This information can include parameters which refer to attitudes toward risk (preferences for risk).

In the next section we introduce the basic structure where those two potential functions can exist. In the next two sections we consider two applications from the social sciences which will attempt to highlight the possible added value of using those potential functions in a non-physics setting. We also consider in the last section of the paper a brief discussion on the relevance of such potential functions in real world market settings.

2. Basic Structure where Two Potential Functions can Occur

The so called stochastic equivalent of the Hamilton-Jacobi equations provides for at least, according to the author of this paper, a sort of twilight state where we move from classical mechanics to stochastics and then to quantum mechanics. The hydrodynamic approach to quantum mechanics was developed by Edward Nelson, and in this section we will provide for the essentials of the basic

structure which we need, to develop the examples in Sections 3 and 4. We use elements from the set up in the paper by Haven (2015). For a lot more detail, the paper by Nelson (1966) is the essential reference. The book by Paul and Baschnagel (1999) is also an excellent source (see also Haven and Khrennikov, 2013).

We will follow, as in the paper by Haven (2015) and Paul and Baschnagel (1999) approach to Nelson's theory. We also use the same notation. Consider a position in space, x, indexed by time (which we denote here by the index n). Let time contract to zero, one can then write as in Nelson (1966) and Paul and Baschnagel (1999), that $\frac{dx_+}{dt} \simeq \frac{x_{n+1}-x_n}{\epsilon}$ and $\frac{dx_-}{dt} \simeq \frac{x_n-x_{n-1}}{\epsilon}$, where ϵ denotes the difference in time t, and d indicates the infinitesimal differential operator. In the area of finance, the so called Brownian motion is a very common way of describing the random evolution of an asset price over time. Brownian motions are well-known in physics, especially with the formalization Einstein gave of such motions. As we mention in Haven (2015), Nelson (1966), and Paul and Baschnagel (1999), define two Brownian motions of the following types:

$$dx(t) = b_+(x, t)dt + \sigma \, dW(t); \tag{1}$$

where $dW(t)$ is a Wiener process; σ is the diffusion coefficient and $b_+(x, t)$ is the so called drift function. They also define:

$$dx(t) = b_-(x, t)dt + \sigma \, dW(t). \tag{2}$$

What is now the difference between the two drift functions? Still following the set up in Haven (2015), Nelson (1966), and Paul and Baschnagel (1999) define:

$$D_+ x(t) = \lim_{\epsilon \to 0} E\left[\frac{x_{n+1}-x_n}{\epsilon}\right] = b_+(x, t), \tag{3}$$

and also:

$$D_- x(t) = \lim_{\epsilon \to 0} E\left[\frac{x_n - x_{n-1}}{\epsilon}\right] = b_-(x, t). \tag{4}$$

Note that E is the expectation operator. This paper has mentioned in its introduction that we consider two types of potentials. But what are they? The real potential is the first type, well-known from elementary classical mechanics. The real potential formalizes potential energy. The second type, is the so called quantum potential which emerges from inserting the polar form of a wave function into the Schrödinger partial differential equation. The uses of those potentials were first proposed by Khrennikov (1999) already more than 10 years ago. Haven (2015)[1], indicates that a key component of the so called quantum potential, can be written as follows:

$$\frac{\nabla^2 R'}{R'} = \frac{1}{\sigma^2} \frac{\partial}{\partial x}\left[\frac{1}{2}\left(b_+(x, t) - b_-(x, t)\right)\right]$$
$$+ \frac{1}{4\sigma^4}\left[b_+(x, t)^2 - 2b_+(x, t)b_-(x, t) + b_-(x, t)^2\right], \tag{5}$$

---

[1]This is Equation (25) in that paper.

where R' is a scalar field. We then argue in Haven (2015)[2] that if we set $b_+(x, t) = b$ and $b_-(x, t) = c$ where $b \neq c; b, c \in \mathbb{R}$:

$$\frac{\nabla^2 R'}{R'} = \frac{1}{4\sigma^4}\left[b^2 - 2bc + c^2\right] = \frac{1}{4\sigma^4}\left(b - c\right)^2. \tag{6}$$

If one consider the classical mechanical equivalent of the quantum potential, one multiplies $\frac{\nabla^2 R'}{R'}$ with $\frac{-m\sigma^4}{2}$, where m is mass. We note that use is made of the conversion: $\sigma^2 = \frac{\hbar}{m}$. We note that such conversion requires further discussion (see Nelson, 1985). Hence, we obtain, as in Haven (2015)[3]:

$$\frac{-m\sigma^4}{2} \frac{\nabla^2 R'}{R'} = \frac{-m}{8}(b - c)^2. \tag{7}$$

We are now ready to consider our first application.

3. Application 1: Three Examples Showing Which Additional Information is Brought on by the Use of the Quantum Potential

The Newtonian motion with both the real and quantum potentials is: $m.a = -\nabla (V + Q)$ (see for instance Haven and Khrennikov, 2013, for more detail). If one consider the force, $-\nabla Q$, to be applied on Equation (7), one can see immediately that:

$$-\nabla Q = -\frac{\partial}{\partial x}\left[\frac{-m\sigma^4}{2} \frac{\nabla^2 R'}{R'}\right] = \frac{-\partial}{\partial x}\left[\frac{-m}{8}(b - c)^2\right] = 0. \tag{8}$$

The force derived from the real potential, $-\nabla V$, is $-b_+(x, t)$ or $-b_-(x, t)$. From an economics point of view, such force can be interpreted as an expected return. Hence, this force can thus incorporate preferences for risk.

In order to give an interpretation to the quantum potential, we need to re-consider Equation (5) but now for the case where $b_+(x, t)$ and/or $b_-(x, t)$ are not constant. Hence, let us consider the simple case where $b_+(x, t) = \mu x$, where we can set that μ is now the expected return. We note that making μ to be such an expected return follows in parallel to what is done in financial economics, where the drift term of the geometric Brownian motion is a product of the expected return and the position variable (i.e., the value of the stock price for instance). Let us assume, for easiness of purpose, that $b_-(x, t) = 0$. In this case, we are *not* in Newtonian mechanics since we are now explicitly imposing that $b_+(x, t) \neq b_-(x, t)$. We re-consider Equation (5) again:

$$\frac{\nabla^2 R'}{R'} = \frac{1}{\sigma^2} \frac{\partial}{\partial x}\left[\frac{1}{2}\left(\mu x\right)\right] + \frac{1}{4\sigma^4}\left[\mu^2 x^2\right]. \tag{9}$$

---

[2]This is Equation (26).
[3]This is Equation (36).

The quantum potential, Q, is written as: $\frac{-m\sigma^4}{2}\frac{\nabla^2 R'}{R'} = \frac{-m\sigma^2}{2}\frac{\partial}{\partial x}\left[\frac{1}{2}(\mu x)\right] - \frac{m}{8}\left[\mu^2 x^2\right]$. The force is then:

$$-\nabla Q = -\frac{\partial}{\partial x}\left[\frac{-m\sigma^2}{2}\frac{\partial}{\partial x}\left[\frac{1}{2}(\mu x)\right] - \frac{m}{8}\left[\mu^2 x^2\right]\right]. \qquad (10)$$

This yields then:

$$-\nabla Q = \frac{m}{4}\mu^2 x. \qquad (11)$$

If we assume that the force on the real potential, $-\nabla V = -b_+(x,t) = -\mu x$, then, one can also write that:

$$-\nabla V - \nabla Q = -\mu x + \frac{m}{4}\mu^2 x \qquad (12)$$

We observe from the above, that besides the information received, via the force on the real potential, i.e., the expected return times the position, additional information is now injected via the force on the quantum potential.

If we were to let $b_+(x,t) = \mu x^2$, then using Equation (5):

$$\frac{\nabla^2 R'}{R'} = \frac{1}{\sigma^2}\frac{\partial}{\partial x}\left[\frac{1}{2}(\mu x^2)\right] + \frac{1}{4\sigma^4}\left[\mu^2 x^4\right]. \qquad (13)$$

Using $\frac{-m\sigma^4}{2}\frac{\nabla^2 R'}{R'}$ then the force delivered by both the quantum and real potentials (assuming $-\nabla V = -b_+(x,t) = -\mu x^2$) is:

$$-\nabla V - \nabla Q = -\mu x^2 + \frac{m\mu^2}{2}x^3 + \frac{m\sigma^2}{2}\mu. \qquad (14)$$

If we would set $b_+(x,t) = \mu$, then [see Equation (8)], we would only be able to write that:

$$-\nabla V - \nabla Q = -\mu \qquad (15)$$

Let us compare those simple cases, Equations (12, 14 and 15). We can observe that additional terms are added to the force on the real potential. Remark however that we have assumed that the drift term in the pure Newtonian environment, is the same as the drift terms $b_+(x,t)$ and $b_-(x,t)$. If we translate the pure Newtonian environment, into the Nelson framework we obtain $\nabla R(x,t) = \frac{E\left[\frac{dx_+}{dt} - \frac{dx_-}{dt}\right]}{2\sigma^2} = 0$ and therefore, in that setting, $R(x,t)$ is constant. This would mean that the density function, in the Nelson framework $e^{2R(x,t)}$ would be constant. The quantum potential would also be zero. Hence, the equivalent information of the pure Newtonian setting into a Nelsonian setting would be senseless. However, the quantum potential still is comparable to the real potential, after all one can write: $-\nabla V - \nabla Q = m.a$! This is in some sense a dilemma.

Consider again Equations (12, 14 and 15), and let us re-write slightly, as follows:

- Under the Newtonian based theory if the expected return (i.e., force on real potential) is the expected return μ then the additional information (brought by the gradient of the quantum potential) is nil

- Under the Newtonian based theory if the expected return (i.e., force on real potential) is the expected return μx then the additional information (brought by the gradient of the quantum potential) is [set $m = 1$ (please see below for a comment on such setting)]: $\frac{1}{4}\mu^2 x$

- Under the Newtonian based theory if the expected return (i.e., force on real potential) is the expected return μx^2 then the additional information (brought by the gradient of the quantum potential) is [set $m = 1$ (please see below for a comment on such setting)]: $\frac{\sigma^2 \mu}{2} + \frac{\mu^2}{2}x^3$

Given the reasonableness of this force on the quantum potential to exist (i.e., volatility is not zero and the uncertainty given by $\left(\frac{x_{n+1}-x_n}{\epsilon}\right) \neq \left(\frac{x_n-x_{n-1}}{\epsilon}\right)$ to exist, it would then be reasonable to claim that if the expected drift is given as μx in a Newtonian world [where $\left(\frac{x_{n+1}-x_n}{\epsilon}\right) = \left(\frac{x_n-x_{n-1}}{\epsilon}\right)$ and $\sigma = 0$] then in a Nelson world this information would need to be augmented with: $\frac{1}{4}\mu^2 x$. Similarly, if the expected drift is given as μx^2 in a Newtonian world [where $\left(\frac{x_{n+1}-x_n}{\epsilon}\right) = \left(\frac{x_n-x_{n-1}}{\epsilon}\right)$ and $\sigma = 0$] then in a Nelson world this information would need to be augmented with: $\frac{\sigma^2 \mu}{2} + \frac{\mu^2}{2}x^3$.

Remark one very important issue which refers to the setting of $m = 1$ in the above discussion. The Nelson theory allows for a transition from pure Newtonian mechanics, into a stochastic environment [with the use of $R(x,t)$ as a scalar field] and from there, into a further transition into quantum mechanics [with the use of $R(x,t)$ now as an input into the wave function]. This transition from stochastics to quantum mechanics, also goes via the setting of $\sigma^2 = \frac{\hbar}{m}$.

For a given finite σ^2, in a quantum mechanical context when $\hbar = \sigma^2 m$, it would mean that m should be extremely!!! small indeed. The question becomes, whether the level of m has a continuum of values when transiting from the stochastic environment toward the quantum mechanics environment. If one considers the case $-\mu x + \frac{m}{4}\mu^2 x$ then when m is extremely small, the term $\frac{m}{4}\mu^2 x$ would need to be extremely small. The same can be said for: $-\mu x^2 + \frac{m\mu^2}{2}x^3 + \frac{m\sigma^2}{2}\mu$, where the terms which are added to μx^2 are then small too, because of small m: $\frac{m\mu^2}{2}x^3 + \frac{m\sigma^2}{2}\mu$. It would be a major achievement, if indeed we could find how m behaves when transiting from Newtonian $\rightarrow$ stochastics (with R as scalar field) $\rightarrow$ quantum mechanics (with R as an input to the wave function).

4. Application 2. An Example of How the Real and Quantum Potentials can be used in Lux's Noise Trader Infection Model

The "noise trader/infection" model was developed by Lux (1997) and we use it here to highlight the applications we can make, in a financial economics framework, of the potentials we have treated in our paper. From Equation (5), we can observe that σ^2 is essential to define the quantum potential. We will make the simple assumption, for the purposes of the model treated here,

that $E\left(\frac{dx_-}{dt}\right) = 0$. There is a non-zero uncertainty due to the fact that $E\left(\frac{dx_+}{dt}\right) \neq E\left(\frac{dx_-}{dt}\right)$.

4.1. Brief Set up of the Lux Model

The model has two main types of traders, fundamental and chartist (or also called noise) traders. The total number of chartist traders is $2N$. They are divided into two subgroups, n_+ (the number of noise traders who are positive about the development of the market) and n_- (the number of noise traders who are negative about the development of the market) and hence $n_+ + n_- = 2N$. An opinion index, which is, in the words of Lux (1997) (p. 8) "the distribution of attitudes among the 'population'" is also constructed and it is defined as: $x = \frac{n_+ - n_-}{2N}$. Remark that in the model, chartist traders may change from one subgroup to the other. The smallest difference in the opinion index is given by $\pm \frac{1}{N}$ and the asset price's minimal change is $\pm$ one cent. The central equation in which we are interested in for the purposes of our paper, is as follows (Equation 3.7b in Lux, 1997):

$$\frac{d\bar{p}}{dt} = \beta \left(\bar{x} T_c + T_f \left(p_f - \bar{p} \right) \right); \qquad (16)$$

where β is defined as (Lux, 1997, p. 8): "a parameter for the average speed of price adjustment in the presence of excess demand."; $\bar{x}$ is the average distribution of attitudes; T_c measures the trading volume of the chartist traders; T_f measures the trading volume of the fundamental traders; p_f is the perceived fundamental value of the asset and $\bar{p}$ is the expected price.

4.2. Embedding Lux's Model in the Quantum/Real Potential Environment

If we want to embed the above model in the quantum/real potential model presented here in this paper, then a departure of Lux's model could be as follows:

$$d\bar{p} = \beta \left(\bar{x} T_c + T_f \left(p_f - \bar{p} \right) \right) dt + dW; \qquad (17)$$

where dW is a Wiener process as defined before. Embedding this departure of Lux's model in the quantum/real potential model, we can then write that:

$$E\left[\frac{d\bar{p}}{dt}\right] = \bar{x}\beta T_c + \beta T_f p_f - \bar{p}\beta T_f + E(dW); \qquad (18)$$

which can then be re-written as, using $E(dW) = 0$:

$$E\left[\frac{d\bar{p}}{dt}\right] = \bar{x}\beta T_c + \beta T_f p_f - \bar{p}\beta T_f. \qquad (19)$$

Remark that the $\bar{x}$ parameter could be interpreted as being closely linked to some implicit preference functional, which in this model, is driven by chartists. We remark that Equation (3.7a; p. 13) in Lux's paper Lux (1997) provides for the time dependent evolution of the $\bar{x}$ parameter.

4.3. Consequence of the Absence of Expectation Operators

Remark that if we were to write Equation (16) (as it is thus written in Lux's model), as a result of having it embedded in our quantum/real potential model (thus now without expectation operator) then this would imply that $\frac{dW}{dt} = 0$. If we now go back to the importance of expectation operators in Nelson's theory we can say the following. Imagine we were to not use such an operator on $\frac{dW}{dt}$. For instance, can we then still write that $\frac{dx_+}{dt} = b_+(x, t)$, using the Brownian motion Equation (1)? The answer to this question would impose the requirement that no time reversibility can exist. The argument is quite straightforward. Let us follow the arguments of Merton (1990) (please see also Neftci, 2000, for a treatment of Merton's arguments which we follow here) who shows that $\frac{dW}{dt}$, as is well-known, can not be defined with ordinary derivatives. This problem is circumvented in the Nelson theory by using $E\left[\frac{dW}{dt}\right]$, where $E(dW) = 0$.

Assume we want to impose that $\frac{dW}{dt} = 0$ and we thus allow for the use of no expectation operators. We note again, we are well aware that $\frac{dW}{dt}$ does not exist. But let us do a quick thought experiment and assume it were to exist. What would be the consequences?

Define elapsed time as $h = t_k - t_{k-1}$ and let $n = \frac{T}{h^m}$ with $m > 1$, where T is total time. Following Neftci (2000), assume there exists a quantity A_2 so that: $\infty > A_2 > \sum_{k=1}^{\infty} E\left[\Delta W_k^2\right]$ and there exists a quantity A_3 so that $\frac{E\left[\Delta W_k^2\right]}{V_{\max}} > A_3$ with $A_3 \in]0, 1[$ where $V_{\max} = \max_k E\left[\Delta W_k^2\right]$. From the proof which allows for showing that $E\left[\Delta W_k^2\right] = \sigma_k^2 h$, the following relation is essential (see Neftci, 2000): $\frac{h}{T} \frac{A_2}{A_3} > E\left[\Delta W_k^2\right] > \frac{A_3 A_1}{T} h$, where $0 < A_1 < E\left[\Delta W_k^2\right]$. In order to come to show under what conditions $\frac{dW}{dt} = 0$ (thus without expectation operator) we want to impose, using $n = \frac{T}{h^m}$ ($m > 1$), that $\frac{h^m}{T} \frac{A_2}{A_3} > V_{\max}$. Clearly, if h is small (<1) then h^m ($m > 1$) will be smaller than h. Hence, it is reasonable to write that: $\frac{h^m}{T} \frac{A_2}{A_3} < \frac{h}{T} \frac{A_2}{A_3}$. Hence, if we still want $\frac{h^m}{T} \frac{A_2}{A_3} > V_{\max}$, we must impose that $m > \frac{\log V_{\max} + \log \frac{A_3}{A_2} T}{\log h}$. We know that $A_3 > 0$; $A_2 > A_1 > 0$. However, $h = t_k - t_{k-1}$ must clearly be positive! We can then write that: $\frac{h^m}{T} \frac{A_2}{A_3} > E\left[\Delta W_k^2\right] > \frac{A_3 A_1}{T} h^m$ and one can then define that: $E\left[\Delta W_k^2\right] = \sigma_k^2 h^m$. We can approximate: $\frac{dW}{dt} \simeq \lim_{h \to 0} \frac{\Delta W}{h} = \lim_{h \to 0} \frac{h^{\frac{m}{2}}}{h} = \lim_{h \to 0} h^{\frac{m-2}{2}}$ which for $m > 2$ will yield 0. Thus, we obtain that $\frac{dW}{dt} \simeq 0$ in non expectation operator form when: (i) $m > \frac{\log V_{\max} + \log \frac{A_3}{A_2} T}{\log h}$ and (ii) $m > 2$ and therefore we must impose that $\frac{\log V_{\max} + \log \frac{A_3}{A_2} T}{\log h} = 2$ and this condition would mean that the uncertainty concentrates in a very specific period of time, $V_{\max} = \frac{h^2 A_2}{A_3 T}$. It is clear from the condition $\frac{\log V_{\max} + \log \frac{A_3}{A_2} T}{\log h} = 2$ that there can not exist time reversibility since $h > 0$ (and $h \neq 1$) for the $\log h$ to be valid. This may on prima facie, "prove" that the expectation operators which have been imposed on $\frac{dW}{dt}$ in Nelson's theory are intrinsically connected to time reversibility.

A commutativity rule such as, $E\left[x_n\left(\frac{x_{n+1}-x_n}{\epsilon}\right)-\left(\frac{x_n-x_{n-1}}{\epsilon}\right)x_n\right] \equiv E(C)$, also needs to use such operators. Without the expectation operators, setting x_n fixed: $x_n.\left[\frac{dx_+}{dt}-E\left(\frac{dx_-}{dt}\right)\right]$, is possible. Time reversibility is obtained via the expectation operator. In the set up by Nelson, a function $u(x,t)$ is defined as: $u(x,t)=\frac{1}{2}(b_+(x,t)-b_-(x,t))$, and this function is zero in Newtonian mechanics. We also remark in Haven (2015)[4] that Nelson (1966) and Paul and Baschnagel (1999) define $u(x,t)=\frac{\sigma^2}{2}\nabla\ln f$, where f is a probability density function (for instance $f=e^{2R(x,t)}$).

From the above, we can not define $\frac{dx_-}{dt}$ (since this would require time reversibility) and hence, we can not define $u(x,t)=\frac{\sigma^2}{2}\nabla\ln f$. This is the case, because the relation is obtained under time reversal on the Fokker-Planck pde: $\frac{\partial f(x,t)}{\partial t}+\nabla\left(b_+(x,t)f(x,t)\right)+\frac{\sigma^2}{2}\Delta f(x,t)=0$, which is now impossible since $\log h$ is to exist. If $u(x,t)$ can not be defined then $\frac{\nabla^2 R'}{R'}=\Delta R+\left(\frac{u(x,t)}{\sigma^2}\right)^2$ is not definable.

4.4. What are the Real and Quantum Potential in Lux's Model?

We can write that using the real potential, V, $E\left[\frac{dx}{dt}\right]=\nabla V=b_\pm(x,t)$. In full analogy with this, we write Equation (19) now as:

$$\nabla V=\bar{x}\beta T_c+\beta T_f p_f-\bar{p}\beta T_f; \qquad (20)$$

which we write in shorthand format as: $\nabla V=\alpha-\bar{p}\beta T_f$. Recall Equation (5) which gave the expression for the "quantum potential":

$$\frac{\nabla^2 R'}{R'}=\frac{1}{\sigma^2}\frac{\partial}{\partial x}\left[\frac{1}{2}\left(b_+(x,t)-b_-(x,t)\right)\right]$$
$$+\frac{1}{4\sigma^4}\left[b_+(x,t)^2-2b_+(x,t)b_-(x,t)+b_-(x,t)^2\right]. \qquad (5)$$

Using Equation (20) in Equation (5), we obtain:

$$\frac{\nabla^2 R'}{R'}=\frac{1}{\sigma^2}\frac{\partial}{\partial\bar{p}}\left[\frac{1}{2}(\alpha-\bar{p}\beta T_f)\right]+\frac{1}{4\sigma^4}\left[(\alpha-\bar{p}\beta T_f)^2\right]; \qquad (21)$$

where we have to note that $\bar{x}$ is calculated as per Lux (1997), $\sum_x\sum_p x.L(x,p;t)$; where $L(x,p;t)$ is defined as the probability to occupy a state (x,p); i.e., where x is the distribution of attitudes and p is price and hence indirectly $\bar{x}$ is a function of p, via the probability L. Similarly, note that $\bar{p}=\sum_x\sum_p p.L(x,p;t)$. Similarly, T_c in Lux (1997) is defined as: $T_c\equiv 2Nt_c$, where $2N$ is the total noise trader population and t_c is the amount the chartist, individually buys or sells. We assume that t_c is not dependent on $\bar{p}$.

Simplifying Equation (21) is now straightforward and leads to:

$$\frac{\nabla^2 R'}{R'}=\frac{1}{\sigma^2}\frac{1}{2}(-\beta T_f)+\frac{1}{4\sigma^4}\left[(\alpha-\bar{p}\beta T_f)^2\right] \qquad (22)$$

---
[4]This is Equation (19).

Then multiplying Equation (22) with $\frac{-m\sigma^4}{2}$ and also taking $\frac{-d}{d\bar{p}}$ so as to get the force, we obtain:

$$-\nabla Q=-\frac{d}{d\bar{p}}\left[\left(\frac{-m\sigma^4}{2}\right)\frac{\nabla^2 R'}{R'}\right]=\frac{m}{4}\left(\alpha-\bar{p}\beta T_f\right)\left(-\beta T_f\right). \qquad (23)$$

Recall from Equation (20), that we now can write:

$$-\nabla V=\bar{p}\beta T_f-\alpha. \qquad (24)$$

Recall that the m factor in $-\nabla Q$ can indeed be very small if we move toward a quantum mechanical environment. If we thus write: $-\nabla V-\nabla Q$, we then obtain:

$$-\nabla V-\nabla Q=\bar{p}\beta T_f-\alpha+\frac{m}{4}\left(\alpha-\bar{p}\beta T_f\right)\left(-\beta T_f\right). \qquad (25)$$

This can be simplified to:

$$(\bar{p}\beta T_f-\alpha)\left[1+\beta T_f\left(\frac{m}{4}\right)\right]. \qquad (26)$$

Clearly, if $m\to 0$ and $\sigma^2\neq 0$ then $\beta T_f\left(\frac{m}{4}\right)$ is indeed very small. We can also write: $\frac{\nabla V}{\nabla R}=2\sigma^2$ and hence: $\nabla R=\frac{\nabla V}{2\sigma^2}$. We can then write that:

$$R(\bar{p})=\frac{1}{2\sigma^2}\int\alpha-\bar{p}\beta T_f d\bar{p}; \qquad (27)$$

which is worked out as: $R(\bar{p})=\frac{1}{2\sigma^2}(\alpha\bar{p})-\frac{1}{4\sigma^2}\beta T_f\bar{p}^2+C$. Recall that $s=\frac{1}{1+r}\int\exp(2R(\bar{p},t))d\bar{p}$, where s is the state price. In the context of the model we consider here this would yield:

$$s=\frac{1}{1+r}\int\exp\left(\left(\frac{1}{\sigma^2}(\alpha\bar{p})-\frac{1}{2\sigma^2}\beta T_f\bar{p}^2+C\right)\right)d\bar{p}. \qquad (28)$$

This means that the state price (an insurance price which is paid to guarantee a financial outcome when a particular state of nature occurs) is now dependent, using Lux's model which is embedded in our quantum/real potential model, on the (i) volatility; (ii) the expected price of an asset (which chartists and fundamental buyers buy); (iii) the parameter β for the average speed of price adjustment in the presence of excess demand; (iv) the term $\alpha=\bar{x}\beta T_c+\beta T_f p_f$. Remark that this density function used the amplitude function R from $\nabla R=\frac{\nabla V}{2\sigma^2}$. The quantum potential in its full form is absent from this relation, but parts of that potential (i.e., ∇R) are still figuring in the equation. In this formulation the preference factor (via $\bar{x}$) would be embedded in $\bar{p}$. The volatility parameter seems to be the "conduit" factor which links the changes in R with the changes in V, via $\nabla V=2\sigma^2\nabla R$. Thus, Equation (28) does only exist here if this "conduit" factor exists (i.e., if σ^2 is not zero).

5. What is the Relevance of Real and Quantum Potentials in Empirical Work?

This paper has not yet answered a crucial question, for which we have to thank one of the referees of this paper. Can the quantum and/or real potentials have empirical relevance? The answer is fortunately enough positive. Recent work by Belal Baaquie shows quite clearly how real potential functions can be estimated for traded commodities. Please consult Baaquie (2013). The same author also shows how the minimization of a real potential function, when that function is defined as being equal to the sum of supply and demand functions, yields a more general version of the equilibrium price which is well-known in basic economics. The quantum potential can also be estimated from real market data. In the paper by Tahmasebi et al. (2015) the quantum potential is estimated for the Standards and Poor (S&P) Index. More work is needed on how the path derived from the extended Newtonian motion (i.e., with thus two potentials), can be used to emulate price behavior over time.

6. Conclusion

We have attempted to show that preferences for risk seem to be captured by both types of potentials (i.e., the real and quantum potentials). When embedding the basics of the Lux model in the types of potential approach proposed in this paper, we seem to find that the quantum potential's influence depends on the mass parameter. This parameter varies depending on whether we are far removed (or not) from the quantum physical limit. The last section of the paper does indicate that the real and quantum potentials can be estimated within real financial data settings. But the question may remain, if the proposed analysis in this paper is of any value, how one can interpret the quantum potential in light of those real data interpretations? In the paper by Tahmasebi et al. (2015) it is shown quite clearly that a quantum potential with infinite walls occurs for the S&P Index for short time scales, and when the time scale grows those infinite walls disappear and the quantum potential for the S&P index for large time scales resembles the quantum potential for Gaussian white noise. Price variation is deemed to be very small for small time scales, but allowed to be larger for large time scales. This is intuitively acceptable.

From a Newtonian price path point of view, considering for instance Equation (25), our analysis in this paper seems to indicate, that the influence of the quantum potential (next to the real potential) on the price path, may vary. But to pinpoint, in an economics sense, what this parameter of variation really means is very difficult.

References

Baaquie, B. (2013). Statistical microeconomics. *Physica A* 392, 4400–4416.

Haven, E., and Khrennikov, A. (2013). *Quantum Social Science.* New York, NY: Cambridge University Press.

Haven, E. (2015). Potential functions and the characterization of economics-based information. *Found. Phys.* 45, 1394–1406. doi: 10.1007/s10701-015-9931-4

Khrennikov, A. Y. (1999). Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social and anomalous phenomena. *Found. Phys.* 29, 1065–1098.

Lux, T. (1997). Time variation of second moments from a noise trader infection model. *J. Econ. Dyn. Control* 22, 1–38.

Merton, R. (1990). *Continuous Time Finance.* Oxford: Blackwell Press.

Neftci, S. (2000). *An Introduction to the Mathematics of Financial Derivatives.* New York, NY: Academic Press.

Nelson, E. (1966). Derivation of the Schrödinger equation from newtonian mechanics. *Phys. Rev.* 150, 1079–1085.

Nelson, E. (1985). *Quantum Fluctuations.* Princeton, NJ: Princeton University Press.

Paul, W., and Baschnagel, J. (1999). *Stochastic Processes: From Physics to Finance.* Heidelberg: Springer Verlag.

Tahmasebi, F., Meskinimood, S., Namaki, A., Vasheghani Farahani, S., Jalalzadeh, S., and Jafar, G. R. (2015). Financial market images: a practical approach owing to the secret quantum potential. *Eur. Lett.* doi: 10.1209/0295-5075/109/30001

Quantum-like model of unconscious–conscious dynamics

Andrei Khrennikov *

Department of Mathematics, Mathematical Institute, Linnaeus University, Växjö, Sweden

We present a quantum-like model of sensation–perception dynamics (originated in Helmholtz theory of unconscious inference) based on the theory of quantum apparatuses and instruments. We illustrate our approach with the model of bistable perception of a particular ambiguous figure, the Schröder stair. This is a concrete model for unconscious and conscious processing of information and their interaction. The starting point of our quantum-like journey was the observation that perception dynamics is essentially contextual which implies impossibility of (straightforward) embedding of experimental statistical data in the classical (Kolmogorov, 1933) framework of probability theory. This motivates application of nonclassical probabilistic schemes. And the quantum formalism provides a variety of the well-approved and mathematically elegant probabilistic schemes to handle results of measurements. The theory of quantum apparatuses and instruments is the most general quantum scheme describing measurements and it is natural to explore it to model the sensation–perception dynamics. In particular, this theory provides the scheme of indirect quantum measurements which we apply to model unconscious inference leading to transition from sensations to perceptions.

Keywords: sensation, perception, quantum-like model, quantum apparatuses and instruments, bistable perception, unconscious inference

1. Introduction

In recent years the mathematical formalism of quantum mechanics was applied to a variety of problems outside of quantum physics: from molecular biology and genetics to cognition and decision making (see the monographs, Khrennikov, 2010b; Busemeyer and Bruza, 2012; Haven and Khrennikov, 2012) and the extended lists of references in them as well as in the papers (Aerts et al., 2014; Khrennikov et al., 2014).

The problem of mathematical modeling of bistable perception and, more generally, unconscious inference[1] is that it can be rather complex and that its nature is not understood well-enough to allow one to choose the optimal model. In spite of tremendous efforts during the last 200 years, this problem cannot be considered fully solved (cf. Newman et al., 1996; Laming, 1997). In this note we apply the theory of *quantum apparatuses and instruments* (Davies and Lewis, 1970; Busch et al., 1995; Ozawa, 1997) to quantum-like modeling of sensation–perception dynamics as the concrete example of *unconscious and conscious processing of information and their interaction.* Our model can be applied to general unconscious–conscious information processing. It generalized the quantum-like model developed in Khrennikov (2004). We also point out that this paper is the

---

[1]Unconscious inference (Conclusion) is a term of perceptual psychology invented by von Helmholtz (1866); Boring (1942), to describe an involuntary, pre-rational and reflex-like mechanism which is part of the formation of visual impressions.

first attempt to apply the theory of quantum apparatuses and instruments outside of physics, to cognition and psychology.

Special quantum structures were elaborated in order to mathematically represent most general measurement schemes and are applicable both in classical and quantum physics and, practically, in any domain of science. They generalize the pioneer quantum measurement representation by operators of the projection type, also known as von Neumann–Lüders measurements. In quantum physics, this new general framework is of vital importance since the projection type measurements do not completely cover real experimental situations (Davies and Lewis, 1970; Busch et al., 1995; Ozawa, 1997; Nielsen and Chuang, 2000). It seems that the same holds true in mathematical modeling in cognition and psychology (see Asano et al., 2010a,b; Khrennikov, 2010b; Asano et al., 2011, 2012; Khrennikov and Basieva, 2014; Khrennikov et al., 2014), although here the situation is not yet absolutely clear and, obviously, the underlying reason for using quantum instruments is different.

To motivate the use of the theory of quantum apparatuses and instruments, we shall compare it first to classical probabilistic methods and then to simpler quantum-like models of processing data from cognitive science and psychology based on the von Neumann–Lüders measurements. A detailed discussion on violation of laws of classical probability theory by statistical data collected in cognitive science and psychology can be found in Khrennikov, 2010b and SS. We can, for example, point to the order effect (Khrennikov, 2010b; Wang and Busemeyer, 2013) and the disjunction effect (Khrennikov, 2010b; Busemeyer and Bruza, 2012). In the probabilistic terms these are just various exhibitions of violation of the formula of total probability. In general, during recent years quantum probability and decision making were successfully applied to describe a variety of problems, paradoxes, and probability judgment fallacies, such as Allais paradox (humans violate Von Neumann–Morgenstern expected utility axioms), Ellsberg paradox (humans violate Aumann–Savage subjective utility axioms) (see e.g., Haven et al., 2009; Asano et al., 2010a,b, 2011, 2012; Busemeyer et al., 2011; Pothos and Busemeyer, 2013; Wang and Busemeyer, 2013; Aerts et al., 2014; Khrennikov and Basieva, 2014). Psychologists and economists explore the new way inspired by one simple fact from physics: *quantum probability can work in situations where classical probability does not.* Why? Answers may differ (see Khrennikov, 2010b). We point to contextuality of data as one of the main sources of its non-classicality (Khrennikov, 2010b; Dzhafarov and Kujala, 2012a,b, 2013).

As was pointed out, at the beginning of quantum theory physicists attempted to represent quantum measurements they were dealing with by projectors. The same attitude could be observed in applications of the quantum formalism outside of physics. Granted, some statistical psychological effects can be nicely described with the help of the von Neumann–Lüders measurements (see e.g., Haven et al., 2009; Busemeyer et al., 2011; Busemeyer and Bruza, 2012; Pothos and Busemeyer, 2013; Wang and Busemeyer, 2013; Aerts et al., 2014). However, more detailed analysis showed (Asano et al., 2010a,b; Khrennikov, 2010b; Asano et al., 2011, 2012; Khrennikov and Basieva, 2014; Khrennikov et al., 2014) that, in general, data from

cognitive psychology cannot be embedded into the projection-measurement scheme. Therefore, it is natural to follow the development of quantum physics and proceed within a general theory of measurements.

In this paper we do this by illustrating the general theory of quantum instruments with one concrete example: *bistable perception* of the concrete ambiguous figure, *the Schröder stair.* Why do we use a quantum-like model? Here the argument is more complicated than in the case of the order and disjunction effects and other probability fallacies mentioned above. The deviation from classical probability theory is expressed not as a violation of the *formula of total probability*, but as a violation of one of the *Bell-type inequalities*, namely, the Garg–Leggett inequality (Asano et al., 2014). We point out that the Bell-type inequalities play an important role in modern quantum physics. If such an inequality is violated, then the data cannot fit a classical probability space. As was shown in our previous study (Asano et al., 2014), the data collected in a series of experiments performed at Tokyo University of Science (see Asano et al., 2014) for details, violate the Garg–Leggett inequality (statistically significantly)[2].

The first step toward creation of a quantum-like model of bistable perception was done by Atmanspacher and Filk (2012, 2013). We studied this problem in Asano et al. (2014), where we demonstrated a violation of the Garg–Leggett inequality for experimental probabilistic data collected for rotating image of Schröder stair (the experiment was performed at Tokyo University of Science), in Accardi et al. (in press) we presented a quantum-like adaptive dynamical model for bistable perception. The latter is based on a more general formalism than the theory of quantum instruments—on the theory of adaptive quantum systems. In the present paper, the traditional approach to quantum measurement theory is used for modeling sensation–perception transition and unconscious inference.

Finally, we point out that violation of laws of classical probability theory is a statistical exhibition of violation of laws of classical Boolean logic. Thus, in logical terms the quantum-like modeling of cognition is modeling of a nonclassical reasoning, decision making, and problem solving. In particular, in our model unconscious inference, generation of a perception from a sensation, is not based on the rules of classical logics. We also remark that the so called quantum logic corresponding to the quantum formalism is just one special type of nonclassical logic. In principle, there are no reasons to assume that human (mental) cognition, even if it has a non-Boolean structure, can be modeled completely with the aid of quantum logic and quantum probability. Still more general models might be explored, see (Khrennikov and Basieva, 2014) for a discussion.

---

[2]We remark that the formula of total probability and the Bell-type inequalities can be treated as just two special statistical tests of non-classicality of the data (see Conte et al., 2008; Bruza et al., 2010; Khrennikov, 2010b; Asano et al., 2014; Dzhafarov and Kujala, 2014) for discussion. This is the "minimal interpretation." In quantum physics the standard interpretation of these inequalities is related to whether we can proceed with a realistic and local model. The Garg–Leggett inequality is a rather special type of Bell's inequalities, since it is about time correlations for a single system and the original Bell's inequality is about spatial correlations for pairs of systems.

2. Advantageousness of Quantum Instrumental Modeling in Cognitive Psychology

We emphasize that, as well as quantum physics (Plotnitsky, 2006, 2009), cognitive and social sciences also can be treated as theories of measurements. A great deal of effort has been put into the development of measurement formalisms, cf. with, e.g., the time-honored *Stimulus–Organism–Response (S–O–R) scheme* for explaining cognitive behavior (Woodworth, 1921). Just like the situation in quantum physics, cognitive and social scientists cannot approach the mental world directly; they work with results of observations. Both quantum physics and cognitive and social sciences are fundamentally based on operational formalisms for observations.

The basic notions of the operational formalism for the quantum measurement theory are *quantum apparatus and instrument* (Davies and Lewis, 1970; Busch et al., 1995; Ozawa, 1997). Quantum apparatuses are mathematical structures representing at a high level of abstraction physical apparatuses used for measurements. They encode the probabilities of the results of observations as well as the back-actions of the measurements on the states of physical systems. Such back-actions are mathematically represented with the aid of another important mathematical structure, a quantum instrument. Our aim is to explore the theory of quantum apparatuses and instruments and especially its part devoted to *indirect measurements* in cognitive and social sciences.

The scheme of *indirect measurements* is very useful for applications, both in quantum physics and humanities. In this scheme, besides the "principle system" S, a probe system S' is considered. A measurement on S is composed of the unitary interaction with S' and a subsequent measurement on the latter.

In our cognitive modeling S *represents unconscious information processing and S' conscious.* In the concrete example of Helmholtz unconscious inference, S represents processing of sensation (its unconscious nature was emphasized already by Helmholtz) and S' represents processing of perception - conscious representation of sensation.

This approach provides a possibility to extend the class of quantum measurements which originally were only von Neumann–Lüders measurements of the projection type. Such an extension serves not only the natural seeking of generality. Generalized quantum measurements have some new features. Here we shall concentrate only on those of them relevant to our project on quantum-like cognition.

For us, one of the main problems of exploring solely projective (direct) measurements is their fundamentally invasive nature: *as the feedback of a measurement, the quantum state is "aggressively modified"—it is projected onto the subspace corresponding to the result of this measurement.* In any event, this feature is not so natural for the dynamics of sensation and perception states. Of course, each "perception–creation" modifies the states of sensation and perception, *but these modifications are not of the collapse type,* as they should be in the case of projections.

Important for our applications is that *a variety of different quantum instruments (describing back-reaction transformations resulting from measurements) can correspond to one and the same observable* on the principle system S. That is, measurements having the same statistical results may lead to very different state transformations (due to very different types of interaction between the principle and probe systems). In quantum mechanics (as Ozawa emphasized Ozawa, 1997), the same observable can be measured by different apparatuses having different state-transforming quantum instruments. This is a very important characteristic of the theory of generalized quantum measurements. It is also very useful for cognitive modeling, since it reflects the individuality of measurement apparatuses/instruments which are used by cognitive systems (e.g., human beings) to generate the same perception.

We point out that the scheme of indirect measurements accounts for state dynamics in the process of measurement, which is not just a "yes"/"no" collapse as in the original von Neumann–Lüders approach. The possibility to mathematically describe the mental state dynamics in the process of perception–creation by means of the quantum formalism is very attractive. A study in this direction was already presented in the work of Pothos and Busemeyer (2013), although without appealing to the operational approach to quantum mechanics. In the series of works of Asano et al. (2010a,b, 2011, 2012), the process of decision making was described by a novel scheme of measurements generalizing the standard theory of quantum apparatuses and instruments (Asano et al., 2010a,b, 2011, 2012).

Now we list once again the main advantageous properties of the quantum instrument/apparatus modeling in cognitive psychology:

1. A possibility to model the feedback reaction of a "mental measurement" (including self-measurements such as decision making and problem judgment) without collapse-like projections of mental states (belief states).
2. The same (self-)measurement output can correspond to a variety of mental state processing.
3. This is the only way to consistently model indirect measurements in which the output of one psychological function of the brain is (self-) measured through the output of another psychological function.

3. Quantum States

We start with a brief introduction to the quantum basics and define pure and mixed quantum states. The state space of a quantum system is complex Hilbert space. Denote it by H. This is a complex linear space endowed with a scalar product, a positive-definite non-degenerate Hermitian form. Denote the latter by $\langle \cdot | \cdot \rangle$. It generates the norm on H: $\|\psi\| = \sqrt{\langle \psi | \psi \rangle}$.

A reader who does not feel comfortable in the abstract framework of functional analysis can simply proceed with the Hilbert space $H = \mathbf{C}^n$, where $\mathbf{C}$ is the set of complex numbers, and the scalar product $\langle u | v \rangle = \sum_i u_i \bar{v}_i$, $u = (u_1, ..., u_n)$, $v = (v_1, ..., v_n)$. Instead of linear operators, one can consider matrices.

Pure quantum states are represented by normalized vectors, $\psi \in H : \|\psi\| = 1$. Two colinear vectors, $\psi' = \lambda\psi, \lambda \in \mathbf{C}, |\lambda| = 1$, represent the same pure state. Each pure state can also be represented as the projection operator P_ψ which projects H onto the one dimensional subspace based on ψ. For a vector $\phi \in H$, $P_\psi\phi = \langle\phi|\psi\rangle \ \psi$. Any projector is a Hermitian and positive-definite operator[3]. We also remark that the trace of the one dimensional projector P_ψ equals to 1: $\mathrm{Tr}\ P_\psi = 1$. (We recall that, for a linear operator A, its trace can be defined as the sum of diagonal elements of its matrix in any orthonormal basis: $\mathrm{Tr}\ A = \sum_i a_{ii}$.) We summarize these properties of an operator (matrix) $\rho = P_\psi$ representing a pure state. It is

1. Hermitian,
2. positive-definite,
3. trace one,
4. idempotent: $\rho^2 = \rho$.

A linear operator is an orthogonal projector if and only if it satisfies (1) and (4); in particular, (2) is a consequence of (4). The properties (1–4) are characteristic for one dimensional orthogonal projectors—pure states [for a projector, (3) implies that it is one dimensional], i.e., any operator satisfying (1–4) represents a pure state.

The next step in the development of quantum mechanics was the extension of the class of quantum states, from pure states represented by one dimensional projectors to states represented by linear operators (matrices) having the properties (1–3). Such operators (matrices) are called *density operators* (density matrices). (This nontrivial step of extension of the class of quantum states was based on the efforts of Landau and von Neumann). One typically distinguish pure states, as represented by one dimensional projectors, and mixed states, those density operators which cannot be represented by one dimensional projectors. The terminology "mixed" has the following origin: any density operator can be represented as a "mixture" of pure states (ψ_i):

$$\rho = \sum_i p_i P_{\psi_i}, \ p_i \in [0, 1], \sum_i p_i = 1. \tag{1}$$

The state is pure if and only if such a mixture is trivial: all p_i, besides one, equal to zero. However, by operating with the terminology "mixed state" one has to take into account that the representation in the form Equation (1) is not unique. The same mixed state can be interpreted as mixtures of different collections of pure states.

Any operator ρ satisfying (1–3) is diagonalizable (even in the infinite-dimensional Hilbert space), i.e., in some orthonormal basis it is represented as a diagonal matrix, $\rho = \mathrm{diag}(p_j)$, where $p_j \in [0, 1], \sum_j p_j = 1$. Thus, it can be represented in the form Equation (1) with mutually orthogonal one dimensional

---

[3]We recall that a linear operator A in H is called Hermitian if it coincides with its adjoint operator, $A = A^\star$. If an orthonormal basis in H is fixed, (e_i), and A is represented by its matrix, $A = (a_{ij})$, where $a_{ij} = \langle Ae_i|e_j\rangle$, then it is Hermitian if and only if $\bar{a}_{ij} = a_{ji}$. A linear operator is positive-definite if, for any $\phi \in H$, $\langle A\phi|\phi\rangle \geq 0$. It is equivalent to positive definiteness of its matrix. We remark that, for a Hermitian operator, all its eigenvalues are real.

projectors. The property (4) can be used to check whether a state is pure or not. We point out that pure states are merely mathematical abstractions; in real experimental situations it is possible to prepare only mixed states; one defines the degree of purity as $\mathrm{Tr}[\rho^2 - \rho]$. Experimenters are satisfied by getting this quantity less than some small ϵ.

4. Atomic Instruments/Apparatuses

The notions of instrument and apparatus are based on very simple and natural consideration. Consider systems of any origin (physical, biological, social, financial). Suppose that the states of such systems can be represented by points of some set X. These are statistical states, i.e., by knowing the state of a system one can determine the values of observables only with some probabilities. Then, for each state $x \in X$ and observable A and its concrete value a_i, there is defined a map

$$p_i = f_{A,a_i}(x) \tag{2}$$

giving the probability of the result $A = a_i$ for systems in the state $x \in X$. Here $f_{A,a_i} : X \to [0, 1]$. Then its is natural to assume that the measurement modifies the state x, i.e., there is is defined another map

$$x_i = g_{A,a_i}(x), \tag{3}$$

here $g_{A,a_i} : X \to X$. This scheme is applicable both in classical and quantum physics as well as in psychology—Stimulus–Organism–Response (S–O–R) scheme for explaining behavior (Woodworth, 1921) of humans and other cognitive systems.

For the fixed observable A, the system of the state transformation maps (g_{A,a_i}) corresponding to all possible values (a_i) of A is called an *instrument* and the collection of maps $(f_{A,a_i}; g_{A,a_i})$ is called an *apparatus.* Of course, this scheme is too general and, to get something fruitful, one has to select the state space X having a special structure and special classes of f- and g-maps. Quantum theory is characterized by selection of the state space starting with a complex Hilbert space. This choice leads to theory of *quantum instruments and apparatuses.*

The general theory of quantum measurements is mathematically advanced, Section 9. Therefore, it is useful to illustrate it by a simple example. We consider the simplest class of quantum instruments extending the class of von Neumann–Lüders instruments of the projection type. These are *atomic instruments.*

Suppose that the range of values of a measurement, spectrum of an observable, is discrete $O = \{a_1, ..., a_n\}$. The main point of theory of instruments is that each measurement resulting in a concrete value a_i generates the feedback action to the original state ρ of a quantum system, i.e., ρ is transformed into a new state ρ_{a_i}, see Equation (3):

$$\rho \to \rho_{a_i}. \tag{4}$$

We start with the standard von Neumann–Lüders measurements. which gives us an important class of quantum instruments/apparatuses (especially from the

historical viewpoint) . These measurements are mathematically represented by Hermitian operators,

$$A = \sum_i a_i P_{a_i}, \tag{5}$$

where P_{a_i} is the projector onto the eigensubspace corresponding to the eigenvalue a_i. For pure states, the transformation (Equation 4) is based on the projection P_{a_i}:

$$\psi \to P_{a_i} \psi, \tag{6}$$

this map is linear and it is convenient to work with it. However, if $P_{a_i} \neq I$, where I is the unit operator, then $\|P_{a_i}\psi\| < 1$, so the output of Equation (6) is not a state. To get a state, it has to be normalized by its norm:

$$\psi \to \frac{P_{a_i}\psi}{\|P_{a_i}\psi\|}. \tag{7}$$

This is a map from the space of pure states into the space of pure states, but it is nonlinear. This type of the feedback reaction to the result of measurement was postulated by von Neumann. It is well-known as the *projection postulate* of quantum mechanics (the state reduction postulate or the state collapse postulate, see (Khrennikov and Basieva, 2014) for a psychologist-friendly discussion on these postulates and their role in quantum physics and cognitive psychology and psychophysics) [4].

Now, for a pure state ψ, one can consider its representation by the density operator $\rho = P_\psi$. In such terms, the state transform (Equation 6) can be written as

$$\rho \to P_{a_i} \rho P_{a_i}. \tag{8}$$

This is the simplest example of a transformation which in quantum measurement theory is called a *quantum operation*. It can be extended to the linear map from the space of linear operators (matrices) to itself—by the same formula (Equation 8). For a finite spectral set O, the collection of quantum operations (Equation 8) , $a_i \in O$, gives the simplest example of a *quantum instrument*.

We are again interested in a map from the space of density operators (matrices) to itself, see Equation (4). Thus, we again have to make normalization:

$$\rho \to \rho_{a_i} = \frac{P_{a_i} \rho P_{a_i}}{\mathrm{Tr}\, P_{a_i} \rho P_{a_i}}. \tag{9}$$

---

[4]It is less known (in fact, practically unknown) that von Neumann sharply distinguished the case of observables with non-degenerate spectra, i.e., all (P_{a_i}) in the spectral decomposition of A, see Equation (5), are one dimensional projectors, and degenerate spectra, i.e., some of (P_{a_i}) are projectors onto multi-dimensional subspaces. In the first case he postulated aforementioned state-collapse (Equation 7), but in the second case he pointed out that the measurement feedback can generate state transformations different from one given by Equation (7); in particular, the output of the initial pure state can be a mixed state. Later Lüders extended the von Neumann projection postulate even to projectors with degenerate spectra, i.e., in fact, he reduced the class of possible state transformations (quantum operations). This simplification was convenient in theoretical studies and the projection postulate was widely treated as applicable generally, i.e., even to observables with degenerate spectra. The name of Lüders was washed out from the majority of foundational works and nowadays the projection postulate is typically known as the von Neumann projection postulate (see Khrennikov, 2008) for more details.

It is nonlinear and physicists work with quantum operations (forming instruments), by making normalization by trace only at the final step of calculations which can involve a chain of measurements.

However, we are primarily interested not in the measurement feedback to the initial quantum state ρ, but in the probabilities to get the results $a_i \in O$. Denote them $p(a_i|\rho)$. Here they are given by Born's rule. If the initial state is pure $\rho = P_\psi$, then

$$p(a_i|\psi) = \langle P_{a_i}\psi|\psi\rangle = \|P_{a_i}\psi\|^2. \tag{10}$$

It is easy to see that

$$p(a_i|\psi) = \mathrm{Tr}\, \mathrm{P}_{a_i} \mathrm{P}_\psi. \tag{11}$$

This formula can be easily generalized, e.g., via Equation (1), to an arbitrary initial state ρ:

$$p(a_i|\rho) = \mathrm{Tr}\, \mathrm{P}_{a_i} \rho. \tag{12}$$

A *quantum apparatus* is the combination of feedback state-transformations, i.e., a quantum instrument, and detection probabilities.

In the von Neumann–Lüders approach the quantum instrument is uniquely determined by an observable, the Hermitian operator A. The latter is the basis of the construction. However, even in this approach we could start directly with an instrument determined by a family of mutually orthogonal projectors (P_{a_i}), i.e.,

$$\sum_i P_{a_i} = I, \tag{13}$$

where $P_{a_i} \perp P_{a_j}$, $i \neq j$, and then define the observable A simply as this family (P_{a_i}). In quantum information the values a_i have merely the meaning of *labels* for the results of measurement. For future generalization, we remark that the normalization condition (Equation 13) can be written as

$$\sum_i P_{a_i}^\star P_{a_i} = I, \tag{14}$$

because, for any orthogonal projector P, $P^\star = P$ and $P^2 = P$.

Now we move to general atomic instruments and apparatuses. Here quantum operations have the form:

$$\rho \to Q_{a_i} \rho Q_{a_i}, \tag{15}$$

where, for each value a_i, Q_{a_i} is a linear operator which is a contraction (i.e., its norm is bounded by 1). These operators are constrained by the normalization condition, cf. (Equation 14):

$$\sum_i Q_{a_i}^\star Q_{a_i} = I, \tag{16}$$

These operations determine an atomic quantum instrument. Each quantum operation induces the corresponding state transformation:

$$\rho \to \rho_{a_i} = \frac{Q_{a_i} \rho Q_{a_i}}{\mathrm{Tr}\, Q_{a_i} \rho Q_{a_i}}. \tag{17}$$

In particular, pure states are transformed into pure states (similar to the von Neumann–Lüders measurements):

$$\psi \rightarrow \frac{Q_{a_i}\psi}{\|Q_{a_i}\psi\|}. \tag{18}$$

Probabilities of the results of measurements are given by the following generalization of Equation (12):

$$p(a_i|\rho) = \mathrm{Tr}\, M_{a_i}\rho, \tag{19}$$

where

$$M_{a_i} = Q_{a_i}^{\star} Q_{a_i}. \tag{20}$$

(We remark that if Q_{a_i} is a projector, then $Q_{a_i}^{\star} = Q_{a_i}$ and $Q_{a_i}^2 = Q_{a_i}$. Thus, in this case (Equation 19) matches with (Equation 12). In this way we obtain the corresponding quantum instrument.

The class of atomic instruments and apparatuses is the most direct generalization of the von Neumann–Lüders class. In particular, in general quantum instruments do not transfer pure states into pure states, see Appendix.

5. Bistable Perception of Schröder Stair

The experiment is about perception of on the ambiguous figure, the Schröder stair, see **Figure 1**. Here we reproduce data from paper (Asano et al., 2014), where the reader can find a more detailed presentation.

A total of 151 subjects participated in the test performed at Tokyo University of Science. They were divided into three groups ($n_A = 55$, $n_B = 48$, $n_C = 48$). To the subjects of all three groups, we showed 11 pictures of the Schröder stair which was leaning at different angles. Subjects answered L = "I can see that left side is front," or R = "I can see that right side is front" for each picture. Thus, we have a random variable for perception, $X_\theta = L, R$. We denote the experimental probability that a subject answers "Left side is front" by $p(X_\theta = L)$.

For the first group (A), order of showing pictures is randomly selected for each subject. For the second group (B), angle θ changed from 0 to 90 as if the picture was rotating clockwise.

Inversely, for the third group (C), the angle θ was changed from 90 to 0. As a result, we obtained perception trends with respect of angles, see **Figure 2**. These graphs demonstrate contextuality of data, its dependence on experimental contexts, (A)–(C), (see Asano et al., 2014) for numerical estimation of the degree of contextuality as violation of the Garg–Leggett inequality. As was discussed in Introduction, contextual statistical data can be modeled by using the quantum formalism.

6. Mental Apparatuses

We shall proceed with finite dimensional state spaces by making remarks on the corresponding modifications in the infinite dimensional case. The symbol $D(H)$ denotes the space of density operators in the complex Hilbert space H; $L(H)$ the space of all linear operators in H (bounded operators in the infinite dimensional case).

The space $L(H)$ can itself be endowed with the structure of the linear space. We also have to consider linear operators from $L(H)$ into itself; such maps, $T : L(H) \rightarrow L(H)$ are called *superoperators*. We shall use this notion only in Section 9. Thus, for a moment, the reader can proceed without it.

Moreover, on the space $L(H)$ it is possible to introduce the structure of Hilbert space with the scalar product

$$\langle A|B\rangle = \mathrm{Tr}\, A^{\star}B.$$

Therefore, for each superoperator $T : L(H) \rightarrow L(H)$, there is defined its adjoint (super)operator $T^{\star} : L(H) \rightarrow L(H)$, $\langle T(A)|B\rangle = \langle A|T^{\star}(B)\rangle$, $A, B \in L(H)$.

For reader's convenience we remind the notion of POVM.

Definition. *A positive operator valued measure (POVM) is a family of positive operators $\{M_j\}$ such that $\sum_{j=1}^{m} M_j = I$, where I is the unit operator.*

Consider a cognitive system, to be concrete consider a human individual, call her Keiko. She confronts some recognition-problem, i.e., in our problem of bistable perception of Schröder stair she has to make the choice between two perception $A = L, R$. In the quantum(-like) model the space of her mental states is represented by complex Hilbert space H (pure states are
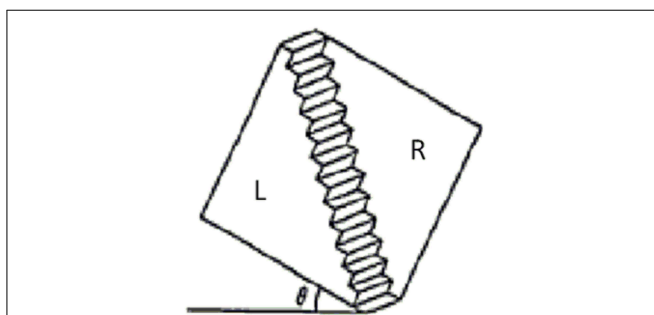

FIGURE 1 | Schröder Stair is an ambiguous figure which may have two different interpretations, "left part (L) is front and right part (R) is back," and its converse. Humans percept either of them, and the tendency of the perception depends on the roatating angle θ.
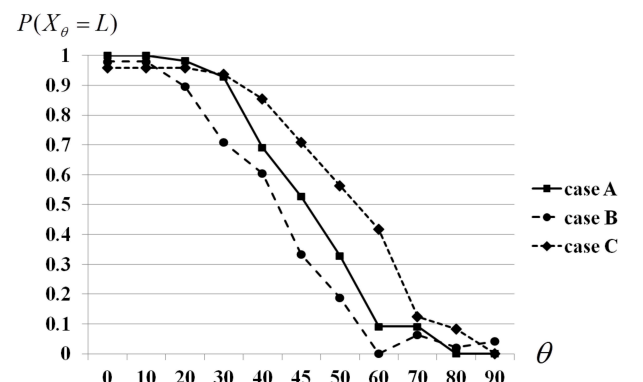

FIGURE 2 | Optical illusion is affected by memory bias: subject's perception is shifted in response to rotation direction of the figure.

represented by normalized vectors and mixed states by density operators).

In the model under construction H is tensor-factorized into two components, namely, $H = H \otimes K$, where H is the space of *sensation-states* and K is the space of *perception-states*. The states of the latter are open for *conscious introspection*, but the states of the former are in general not approachable consciously. We recall that we model Helmholtz unconscious inference.

In general suppose that Keiko confronts some concrete recognition problem A with possible perceptions labeled as $a_i, i = 1, 2, ..., m$. We denote the set of possible values of A by the symbol O, i.e., $O = \{a_1, .., a_m\}$. By interacting with a figure (in our concrete case the figure is ambiguous) she generates the the sensation-state ρ (e.g., a pure state, i.e., $\rho = |\psi\rangle\langle\psi|, \psi \in H, \|\psi\| = 1$). The process of generation of ρ can be mathematically represented as a unitary transformation in the space H. Denote the pre-recognition state of sensation by ρ_0. Then

$$\rho = U\rho_0 U^\star,$$

where the unitary operator $U : H \to H$ depends on the figure; in our concrete case $U = U_{\text{Schr}}$.

To come to the concrete perception, Keiko uses a "mental apparatus," denoted as A, which produces the results (perceptions) a_i randomly with the probabilities $p(a_i|\rho)$, the output probabilities[5]. An apparatus represents not only perceptions and the corresponding probabilities, but also the results of the evolution of the initial sensation-state ρ as induced by the back-reaction to the concrete perception a_i. This is a sort of the state reduction, "sensation-state collapse" as the result of creation of the concrete perception a_i. Thus, the sensation state ρ which Keiko created from her visual image is transformed into the output state ρ_{a_i}.

However, as we shall see, in general this sensation-state update can be sufficiently peaceful, so our model differs crucially from the orthodox quantum models of cognition (Busemeyer and Bruza, 2012) based on the projection-type state update. Thus, each mental apparatus A corresponding to the recognition-problem A is mathematically represented by

- probabilities for concrete perceptions $p(a_i|\rho)$;
- transformations of the initial sensation-state corresponding to the concrete results of perception,

$$\rho \to \rho_{a_i}. \tag{21}$$

The rigorous mathematical description of such state transformations leads to the notion of *a quantum instrument*, see Section 9.

6.1. Mixing Law

In the quantum operational formalism it is assumed that these probabilities, $p(a_i|\rho)$, satisfy the *mixing law*. We remark that, for any pair of states (density operators) ρ_1, ρ_2 and any pair of probability weights $q_1, q_2 \geq 0, q_1 + q_2 = 1$, the convex

---

[5]We are going toward creation of a cognitive analog of the quantum operational model of measurements with the aid of physical apparatuses.

combination $\rho = q_1\rho_1 + q_2\rho_2$ is again a state (density operator). In accordance with the mixing law any apparatus produces probabilities such that

$$p(a_i|q_1\rho_1 + q_2\rho_2) = q_1 p(a_i|\rho_1) + q_2 p(a_i|\rho_2). \tag{22}$$

In our model of bistable perception the mixing law can be formulated as follows:

A probabilistic mixture of sensations produces the mixture of probabilities for perception outputs.

In physics this is a very natural assumption. However, in modeling of cognitive phenomena, in particular, unconscious inference, an additional analysis of its validity has to be performed. We have no possibility to do this in this note, so we postpone such analysis to one of coming publications. Now we mimic quantum physics explicitly and proceed under the assumption (Equation 22).

6.2. Composition of the Apparatuses

It is natural to assume that after resolving the recognition-problem A a person is ready to look at another image B and proceed to its perception. In general perception of B depends on the preceding perception of A. Such a sequence of perceptions represented as a new mental apparatus, the composition of the apparatuses A and $B : BA$. Its outputs are ordered pairs of perceptions (a_i, b_j). It is postulated that the corresponding output probabilities and states are determined as

$$p((a_i, b_j)|\rho) = p(b_j|\rho_{a_i})p(a_i|\rho); \tag{23}$$

$$\rho_{(a_i, b_j)} = (\rho_{a_i})_{b_j}. \tag{24}$$

The law (Equation 23) can be considered as the quantum generalization of the Bayes rule. The law (Equation 24) is the natural composition law.

In our experiment with rotation of the Schröder stair, we are interested in a sequence of instruments A_θ corresponding to some sample of angles $C = \{\theta_1, ..., \theta_m\}$. Here C determines the context of the experiment. Our data from Section 5 can be represented as the superposition of quantum apparatuses: $A_C = A_{\theta_m}...A_{\theta_1}$. Here A_C is the quantum apparatus representing the context C. In our experimental study we considered not only deterministic contexts corresponding to clockwise and counter-clockwise rotations, but even the random context determined by the uniform probability distribution.

7. Perception through Unitary Interaction Between the Sensation and Perception-states

The above operational description of "perception–production" was formulated solely in terms of sensation-states. However, a sensation-state is a complex informational state which is in general unapproachable for conscious introspective. The operational representation of observables in the space of sensation-states is not straightforward and in general it cannot be formulated in terms of mutually exclusive perceptions. For

example, in our experiment Keiko's perceptions can be binary encoded: $A = L, R$. However, her sensation of the Schröder stair is a complex information state depending on a variety of parameters (in particular, we are interested in dependence on the rotation angle). *The subspaces corresponding to sensations leading to the L-perception and R-perception are in general not orthogonal.* This non-orthogonality of sensation subspaces for different perceptions is the fundamental feature of bistable perception, recognition of ambiguous figures.

Therefore, it is more fruitful to define the perception-observable directly by using an additional state space, the space of the perception-states K. *In the perception space a perception-observable can be defined as the standard von Neumann–Lüders projection observable.*

Example 1. Consider the simplest case: recognition of the fixed figure A, with dichotomous output, i.e., there are two possible outcomes of "perception-measurement," e.g., $L = 0$ and $R = 1$ for the Schröder stair. This observable can be represented by the pair of projectors (P_0, P_1) onto the subspaces K_0 and K_1 of the perception space K. Since the perceptions $a_0 = 0$ and $a_1 = 1$ are mutually exclusive, and sharply exclusive, the subspaces K_0 and K_1 are orthogonal. Hence, the projectors P_0 and P_1 can be selected as orthogonal. The perception-observable A can be represented as the conventional von-Neumann-Lüders observable $\hat{A} = a_0 P_0 + a_1 P_1 (= P_1)$. However, we emphasize that this representation is valid only in the perception-state space K. It is often (but not always!) possible to proceed with one dimensional projectors, i.e., to represent possible perceptions just by the basis vectors in the two dimensional perception-state space, $(|0\rangle, |1\rangle)$. Here each perception-state can be represented as superposition

$$\phi = c_0|0\rangle + c_1|1\rangle, \ |c_0|^2 + |c_1|^2 = 1. \tag{25}$$

Measurement of A leads to probabilities of perceptions given by squared coefficients, $p_0 = |c_0|^2, p_1 = |c_1|^2$.

In the case of the finite-dimensional perception-state, a perception-observable A can be represented as

$$A = \sum_i a_i P_i, \tag{26}$$

where (P_i) is the family of mutually orthogonal projectors in the space of perception-states K and (a_i) are real numbers encoding possible answers (perceptions).

Now we shall explore the cognitive analog of the standard scheme of *quantum indirect measurements.*

In our cognitive framework "indirectness" means that the sensation-states are in general unapproachable for conscious introspection. Therefore, it is impossible to perform the direct measurement on the sensation-state ρ (in particular, on a pure state $\rho = |\psi\rangle\langle\psi|$). Moreover, in the sensation-state the alternatives, say 0/1, encoded in a perception-observer A are not represented exclusively, they can have overlap. (Mathematically the overlap is expressed as non-orthogonality of sensation-subspaces corresponding to various perceptions.)

In the quantum measurement framework, this situation is described as follows: in the sensation space an observable A is represented as an unsharp observable of the POVM-type. Roughly speaking in the H-representation the A-zero contains partially the A-one and vice versa. The latter is simply a consequence of interpretation of POVM observables as unsharp observables.

Remark 1. To map the quantum physics scheme (Ozawa, 1997) of indirect measurements onto the quantum(-like) cognition scheme, one has to associate the state of the principle physical system S with the sensation-state and the state of the probe physical system S' with the perception-state. We point out that in the cognitive framework we do not consider analogs of physical systems. In principle, one can consider the sensation-system S as a part of the neuronal system representing sensations and the perception system S' as another part of the neuronal system representing possible perceptions. The latter can be specified: different measurements can be associated with different neuronal networks responsible for the corresponding perceptions. However, in principle we need not associate sensation and perception states with the concrete physical neuronal networks. In the case of cognition usage of isolated physical systems as carriers of the corresponding information states might be ambiguous. The interconnectivity of neuronal networks is very high. Therefore, the picture of distributed computational system is more adequate. (Of course, even in physics the notion of an isolated system is just an idealization of the real situation). Therefore, it is useful to proceed in the *purely information approach* by operating solely with states, without coupling them to bio-physical systems. This is, in fact, the quantum information approach, where systems play the secondary role, and one operates with states; especially for the information interpretation of quantum mechanics (Zeilinger, 2010).

In the simplest model we can assume that at the beginning of the process of perception-creation the sensation and perception-states, ρ and σ, are not entangled[6]. Thus, mathematically, in accordance with the quantum formalism, the integral sensation–perception-state, the complete mental state corresponding to the problem under consideration, can be represented as the tensor product

$$R = \rho \otimes \sigma.$$

In the process of perception-creation the sensation and perception-states (cf. Remark 1) "interacts" and the evolution of the sensation–perception-state R is mathematically represented by a unitary operator[7] $U : H \to H$:

$$R \to R_{\text{out}} \equiv URU^\star. \tag{27}$$

In the space of sensation–perception-states H the perception-observer A is represented by the operator $I \otimes A$. Thus, the probabilities of perceptions are given by

$$p_{a_i}^{A \otimes I} = \text{Tr } R_{\text{out}}(I \otimes P_i) = \text{Tr } URU^\star(I \otimes P_i), \tag{28}$$

---

[6] One can say that they are independent. But one can use this terminology carefully, since the notion of quantum independence is more complicated than the classical one and it is characterized by diversity of approaches.

[7] As was mentioned, in the works of Asano et al. (2010a,b, 2011, 2012) and Accardi et al. (in press) even non-unitary evolutions were in charge.

where the projectors (P_i) form the spectral decomposition of the Hermitian observable A in K, see Equation (26).

Since only the perception-state belonging K is a subject of conscious introspective, at the conscious level the perception process can be represented solely in the state space K. The post-interaction perception-state σ_{out} can be (mathematically) extracted from the integral state R_{out} with the aid of the operation of the partial trace:

$$\sigma_{out} = Tr_H R_{out}. \qquad (29)$$

Then perceptions can be represented as the results of the A-measurement (measurement of the projection-type) in the perception space; measurement on the output state σ_{out}. The probabilities of the concrete perceptions (a_i) are given by the standard Born rule:

$$p_{a_i}^A = Tr_K \sigma_{out} P_i = Tr_K(Tr_H R_{out}) P_i = Tr R_{out}(I \otimes P_i) = p_{a_i}^{A \otimes I}. \qquad (30)$$

Thus, Equations (28) and (30) match each other.

If the concrete result $A = a_i$ was observed, then the state of perception σ is transformed into

$$\sigma_{i;out} = \frac{Tr_H R_{out}(I \otimes P_i)}{Tr R_{out}(I \otimes P_i)}. \qquad (31)$$

What does happen in the sensation space?

The expression (Equation 28) for the probability of the perception a_i can be represented as

$$p(a_i|\rho) = p_{a_i}^{A \otimes I} = Tr R_{out}(I \otimes P_i) = Tr\rho \otimes \sigma U^\star(I \otimes P_i)U$$
$$= Tr_H \rho M_{a_i}, \qquad (32)$$

where

$$M_{a_i} = Tr_K(I \otimes \sigma)U^\star(I \otimes P_i)U. \qquad (33)$$

The operator $M_i; H \rightarrow H$ can also be represented in the following useful form (a consequence of the cyclic property of the trace operation):

$$M_{a_i} = Tr_K U^\star(I \otimes P_i)U(I \otimes \sigma) \qquad (34)$$

We remark that (Equation 33) implies:

$$\sum_i M_{a_i} = Tr_K(I \otimes \sigma)U^\star(I \otimes \sum_i P_i)U = Tr_K I \otimes \sigma = (Tr_K \sigma)I.$$

We also remark that each operator M_{a_i} is positively defined and Hermitian.

Thus, in the sensation space the perception-observable of the projection-type A (acting in K) with the spectral family (P_i) is represented as POVM $M = (M_i)$. We remark that in general the operators M_i are not projectors. Such measurement cannot separate sharply sensations leading to perceptions (a_i) for different i.

The operational formalism also gives the "post-perception sensation-state," i.e., the state of sensation created as the feedback to the consciously recognized perception a_i,

$$\rho_{a_i} = \frac{Tr_K R_{out}(I \otimes P_i)}{Tr R_{out}(I \otimes P_i)}. \qquad (35)$$

The output sensation-state depends not only on the initial sensation-state ρ, but also on the initial perception-state σ, interaction between believes and possible perceptions given by U and the question-observable A acting in K.

8. The Indirect Measurement Scheme for Rotation Contexts for Perception of Schröder Stair

As at the very end of Section 6.2, we consider contextual measurements for the Schröder stair: a sequence of perceptions corresponding to some sample of angles $C = \{\theta_1, ..., \theta_m\}$. Here C determines the context of the experiment. We apply the scheme of indirect measurements. We can assume that the perception space K is two dimensional with the orthogonal basis $|L\rangle$, $|R\rangle$ representing the "left-faced" and "right-faced" preceptions of the stair. Thus, projectors P_i, $i = L, R$, are one dimensional.

We start with the initial sensation state ρ_0. By the visual image rotated at the angle θ_1 this state is transformed to

$$\rho_{\theta_1} = U_{Sch;\theta_1} \rho_0 U_{Sch;\theta_1}^\star, \qquad (36)$$

where $U_{Sch;\theta_1}$ represents the unitary dynamics induced by this image. Then the perception of the image is modeled starting with

$$R_{\theta_1} = \rho_{\theta_1} \otimes \sigma_0, \qquad (37)$$

where σ_0 represents the state of perception preceding interaction with the state of sensation. It is natural to assume that $\sigma_0 = |\phi_0\rangle\langle\phi_0|$, where

$$\phi_0 = (|L\rangle + |R\rangle)/\sqrt{2} \qquad (38)$$

is the neutral composition of the states "left-faced" and "right-faced." It represents the deepest state of uncertainty. Suppose (for simplicity) that independently of the angle the interaction of sensation and perception states is given by the same unitary operator U. Then Keiko's perception of the Schröder stair observed at the angle θ_1 with the fixed result $i_1 = L$ or R leads to the new states of sensation and perception:

$$\sigma_{i_1;\theta_1} = \frac{Tr_H U R_{\theta_1} U^\star(I \otimes P_{i_1})}{Tr U R_{\theta_1} U^\star(I \otimes P_{i_1})}, \quad \rho_{i_1;\theta_1} = \frac{Tr_K U R_{\theta_1} U^\star(I \otimes P_{i_1})}{Tr U R_{\theta_1} U^\star(I \otimes P_{i_1})}. \qquad (39)$$

The probability of creation of the perception i can be calculated as

$$p_{i_1;\theta_1} = Tr_H \rho_{\theta_1} M_{i_1;\theta_1}. \qquad (40)$$

Here POVM's component $M_{i_1;\theta_1}$, $i_1 = L, R$, has the form:

$$M_{i_1;\theta_1} = Tr_K U^\star(I \otimes P_{i_1})U(I \otimes \sigma_0). \qquad (41)$$

For the next measurement corresponding to rotation of Schröder's stair for the angle θ_2, Keiko selects $\rho_{i_1;\theta_1}$ and $\sigma_{i_2;\theta_1}$ as the initial states. This means that creation of the fixed perception i_1 leads to disentanglement of her mental state into the product of two states, the state of sensation and perception. Then

$$\rho_{\theta_2} = U_{\text{Sch};\theta_2} \rho_{i_1;\theta_1} U^\star_{\text{Sch};\theta_2}, \tag{42}$$

where $U_{\text{Sch};\theta_1}$ represents the unitary dynamics induced by the θ_2-image. Then

$$R_{\theta_2} = \rho_{\theta_2} \otimes \sigma_{i_1;\theta_1}, \tag{43}$$

Then Keiko's perception of the Schröder stair observed at the angle θ_2 with the fixed result $j = L$ or R leads to the new states of sensation and perception:

$$\sigma_{i_2;\theta_2} = \frac{\text{Tr}_H U R_{\theta_2} U^\star (I \otimes P_{i_2})}{\text{Tr} U R_{\theta_2} U^\star (I \otimes P_{i_2})}, \;\; \rho_{i_2;\theta_2} = \frac{\text{Tr}_K U R_{\theta_2} U^\star (I \otimes P_{i_2})}{\text{Tr} U R_{\theta_2} U^\star (I \otimes P_{i_2})}. \tag{44}$$

The probability of creation of the perception i_2 can be calculated as

$$p_{i_2;\theta_2} = \text{Tr}_H \rho_{\theta_2} M_{i_2;\theta_2}. \tag{45}$$

Starting with $\rho_{i_2;\theta_2}, \sigma_{i_2;\theta_2}$, Keiko generates the perception of the θ_3-rotated stair and so on. After the last test, Keiko's states of sensation and perception $\rho_{i_n;\theta_n}, \sigma_{i_n;\theta_n}$ depend on the sequence of angles C and the sequence of her perceptions $(i_1, i_2, ..., i_n)$. The same is valid for the probability $p_{i_n;\theta_n}$. If the experiment is performed for two different contexts $C = \{\theta_1, ..., \theta_m\}$ and $C' = \{\theta'_1, ..., \theta'_m\}$. Then in general it is impossible to embed the probabilities of perceptions in a single Kolmogorov probability space. Therefore, the use of quantum theory of measurement and "quantum probabilities" can be fruitful. Our approach provides the possibility to model probabilities of perceptions depending on a context, a sequence of angles.

9. Representing Perception by Quantum Instruments

The considered model of perception as the result of unitary interaction between the sensation-state and the perception-state describes an important class of transformations of the sensation-state, see Equation (35). We now turn to the general case which was considered in Section 6, see Equation (21). Set

$$E(a_i)\rho = p(a_i|\rho)\rho_{a_i} \tag{46}$$

and, for a subset Γ of O, where $O = \{a_1, ..., a_m\}$ is the set of all possible perceptions, we set

$$E(\Gamma)\rho = \sum_{a_i \in \Gamma} E(a_i)\rho = \sum_{a_i \in \Gamma} p(a_i|\rho)\rho_{a_i}. \tag{47}$$

We point to the basic feature of this map:

$$\text{Tr} E(O)\rho = \sum_{a_i \in O} p(a_i|\rho)\text{Tr}\rho_{a_i} = 1. \tag{48}$$

For each concrete perception a_i, $E(a_i)$ maps density operators to linear operators (in the infinite dimensional case, these are trace-class operators, but we proceed in the finite dimensional case, where all operators have finite traces).

The mixing law implies that, for any $\Gamma \subset O$,

$$E(\Gamma)(q_1\rho_1 + q_2\rho_2) = q_1 E(\Gamma)\rho_1 + q_2 E(\Gamma)\rho_2. \tag{49}$$

As was shown by Ozawa (1997), under the assumption on the existence of composition of the apparatuses any such a map $E(\Gamma)$: $D(H) \rightarrow L(H)$ can be extended to a linear map (superoperator)

$$E(\Gamma) : L(H) \rightarrow L(H) \tag{50}$$

such that:

- each $E(\Gamma)$ is positive, i.e., it transfers the set of positively defined operators into itself;
- $E(O) = \sum_i E(a_i)$ is trace preserving:

$$\text{Tr} E(O)\rho = \text{Tr}\rho. \tag{51}$$

The latter property is a consequence of Equation (48)[8].

Thus, the two very natural and simple assumptions, the mixing law for probabilities and the existence of composite apparatuses, have the fundamental mathematical consequence, the representation of the evolution of the state by a superoperator (Equation 50).

In quantum physics such maps are known as state transformers (Busch et al., 1995) or DL (Davis–Levis, Davies and Lewis, 1970) quantum operations[9].

Thus, each perception induces the back-reaction which can be formally represented as a state transformer. In these terms

$$\rho_{a_i} = \frac{E(a_i)\rho}{\text{Tr} E(a_i)\rho} \tag{52}$$

We remark that the map $\Gamma \rightarrow L(L(H))$, from subsets of the set of possible perceptions O into the space of superoperators, is additive:

$$E(\Gamma_1 \cup \Gamma_2) = E(\Gamma_1) + E(\Gamma_2), \;\; \Gamma_1 \cap \Gamma_2 = \emptyset. \tag{53}$$

This is a measure with values in the space $L(L(H))$. Such measures are called (DL) instruments (Davies and Lewis, 1970). To specify the domain of applications in our case, we shall call them *perception instruments*.

The class of such instruments is essentially wider than the class of instruments based on the unitary interaction between sensation and perception components of the mental state, see

---

[8] If one wants to extend $E(\Gamma)$ from the set of density operators to the set of all linear operators (in the infinite dimensional case it has to be the set of finite-trace operators) by linearity then it has to be set $E(\Gamma)\mu = E(\Gamma)\text{Tr}\mu(\mu/\text{Tr}\mu) = \text{Tr}\mu\, E(\Gamma)(\mu/\text{Tr}\mu)$ and, in particular, $E(O)\mu = \text{Tr}\mu\, E(O)(\mu/\text{Tr}\mu) = \text{Tr}\mu$.

[9] DL-notion of the quantum operation is more general than the notion used nowadays. The latter is based on complete positivity, instead of simply positivity as the DL-notion, see Appendix for the corresponding definition and a discussions on whether the reasons used in physics to restrict the class of state transformers can be automatically used in cognitive science.

Equation (35). The evident generalization of the scheme of Section 7 is to consider nonunitary interactions between the components of the mental state; another assumption which can be evidently violated in modeling of cognition is that the initial sensation and perception states are not entangled ("independent") (see Asano et al., 2010a,b, 2011, 2012) for generalizations of the aforementioned scheme.

We start with a discussion on possible *nonunitarity of interaction between the sensation and perception states.* In quantum physics the assumption of unitarity of interaction between the principle system S and the probe system S' (representing a part of the measurement apparatus interacting with S) is justified, because the compound system $S + \tilde{S}'$ can be considered (with a high degree of approximation) as an isolated quantum system and its evolution can be described (at least approximately) by the Schrödinger equation. And the latter induces the unitary evolution of a state.

In cognition the situation is totally different. The main scene of cognition is not the physical space-time, but the brain. It is characterized by huge interconnectivity and parallelism of information processing. Therefore, it is more natural to consider the sensation and perception states corresponding to different visual inputs as interacting, especially at the level of the sensation-states. Thus, the perception-creation model based on the assumption of isolation of different perception-creation processes from each other seems to be too idealized, although it can be used in many applications, where the concentration on one fixed problem may diminish the influence of other perception-creation processes.

In physics, the assumption that the initial state of the system $S + \tilde{S}'$ is factorized is also justified, since the exclusion of the influence of the state of the measurement device to the state of a system S prepared for measurement (and vice versa) is the experimental routine. In cognition the situation is more complicated. One cannot exclude that in some situations *the initial sensation and perception state are entangled.*

The representation of probabilities with the aid of POVMs is not a feature of only the unitary interaction representation of apparatuses, see Equation (32). In general, any DL-instrument generates such a representation. Take an instrument E, where, for each $a_i \in O, E(a_i) : L(H) \rightarrow L(H)$ is a superoperator. Then we can define the adjoint operator $E^\star(a_i) : L(H) \rightarrow L(H)$. Set $M_{a_i} = E^\star(a_i)I$, where $I : H \rightarrow H$ is the unit operator. Then, since $p_{a_i} = \text{Tr} E(a_i)\rho = \text{Tr I}; E(a_i)\rho = \langle I|E(a_i)\rho\rangle == \langle E^\star(a_i)I|\rho\rangle = \text{Tr}(E^\star(a_i)I)\rho = \text{Tr} M_{a_i}\rho$. By using the properties of an instrument it is easy to show that M_{a_i} is POVM. Thus, each mental apparatus can be represented by a POVM. We interpret this POVM as the mathematical representation of "unconscious" inference. Such "unconscious measurements" are not sharp, they cannot separate completely different perceptions a_i which are mutually exclusive at the conscious level. Mathematically, we have that the subspaces $H_{a_i} = M_i H$ need not be orthogonal. Sensation states corresponding to the perceptions a_i and a_j, say $\psi_i \in H_{a_i}$ and $\psi_j \in H_{a_j}$, in general have nonzero overlap $\langle \psi_i | \psi_j \rangle \neq 0$.

10. Concluding Remarks

This paper is an attempt to present the theory of generalized quantum measurements based on quantum apparatuses and instruments in a humanities-friendly way. This is a difficult task, since this theory is based on advanced mathematical apparatus. We hope that the reader can at least follow our introductory presentation in Sections 3, 4. Although we applied quantum apparatuses and instruments to the concrete problem of cognition, modeling bistable perception and, more generally, Helmholtz unconscious inference, this approach can be used to model general unconscious–conscious information processing. We hope that in future other interesting examples will be presented with the aid of this formalism (cf. Khrennikov, 2010a, 2014).

References

Accardi, L., Khrennikov, A., Ohya, M., Tanaka, Y., and Yamato, I. (in press). Application of non-Kolmogorovian probability and quantum adaptive dynamics to unconscious inference in visual perception process. *Open Syst. Inform. Dyn.* doi: 10.1007/978-94-017-9819-8

Aerts, D., Sozzo, S., and Tapia, J. (2014). Identifying quantum structures in the ellsberg paradox. *Int. J. Theor. Phys.* 53, 3666–3682. doi: 10.1007/s10773-014-2086-9

Aerts, D., Broeakert, J., Czachor, M., Kuna, M., Sinervo, B., and Sozzo, S. (2014). Quantum structure in competing lizard communities. *Ecol. Model.* 281, 38–51. doi: 10.1016/j.ecolmodel.2014.02.009

Asano, M., Ohya, M., and Khrennikov, A., (2010a). Quantum-like model for perception making process in two players game. *Found. Phys.* 41, 538–548. doi: 10.1007/s10701-010-9454-y

Asano, M., Ohya, M., Tanaka, Y., Khrennikov, A., and Basieva, I. (2010b). On application of Gorini-Kossakowski-Sudarshan-Lindblad equation in cognitive psychology. *Open Syst. Inf. Dyn.* 17, 1–15. doi: 10.1142/S1230161211000042

Asano, M., Ohya, M., Tanaka, Y., Khrennikov, A., and Basieva, I. (2011). Dynamics of entropy in quantum-like model of perception making. *J. Theor. Biol.* 281, 56–64. doi: 10.1016/j.jtbi.2011.04.022

Asano, M., Basieva, I., Khrennikov, A., Ohya, M., Tanaka, Y., and Yamato, I. (2012). Quantum-like model of diauxie in *Escherichia coli*: operational description of precultivation effect. *J. Theor. Biol.* 314, 130–137. doi: 10.1016/j.jtbi.2012.08.022

Asano, M., Khrennikov, A., Ohya, M., Tanaka, Y., and Yamato, I. (2014). Violation of contextual generalization of the Leggett-Garg inequality for recognition of ambiguous figures. *Phys. Scr.* T163, 014006. doi: 10.1088/0031-8949/2014/t163/014006

Atmanspacher, H., and Filk, T. (2012). "Temporal nonlocality in bistable perception," in *Quantum Theory: Reconsiderations of Foundations - 6, Special Section: Quantum-like Decision Making: From Biology to Behavioral Economics, AIP Conf. Proc. Vol. 1508,* eds A. Khrennikov, H. Atmanspacher, A. Migdall, and S. Polyakov (Melville, NY: American Institute of Physics), 79–88.

Atmanspacher, H., and Filk, T. (2013). The Necker-Zeno model for bistable perception. *Top. Cogn. Sci.* 5, 800–817. doi: 10.1111/tops.12044

Boring, E. G. (1942). *Sensation and Perception in the History of Experimental Psychology.* New York, NY: Appleton-Century Co.

Bruza, P., Kitto, K., Ramm, B., Sitbon, L., Blomberg, S. and Song, D.(2010). Quantum-like non-separability of concept combinations, emergent associates and abduction. *Log. J. IGPL* 20, 455–457. doi: 10.1093/jigpal/jzq049

Busch, P., Grabowski, M., and Lahti, P. (1995). *Operational Quantum Physics.* Berlin: Springer Verlag.

Busemeyer, J. R., and Bruza, P. D. (2012). *Quantum Models of Cognition and Perception*. Cambridge: Cambridge Press.

Busemeyer, J. R., Pothos, E. M., Franco, R., and Trueblood, J. (2011). A quantum theoretical explanation for probability judgment errors. *Psychol. Rev.* 118, 193–218. doi: 10.1037/a0022542

Conte, E., Khrennikov, A., Todarello, O., Federici, A., and Zbilut, J. P. (2008). A preliminary experimental verification on the possibility of Bell inequality violation in mental states. *Neuroquantology* 6, 214–221. doi: 10.14704/nq.2008.6.3.178

Davies, E., and Lewis, J. (1970). An operational approach to quantum probability. *Comm. Math. Phys.* 17, 239–260. doi: 10.1007/BF01647093

Dzhafarov, E. N., and Kujala, J. V. (2012a). Selectivity in probabilistic causality: where psychology runs into quantum physics. *J. Math. Psychol.* 56, 54–63. doi: 10.1016/j.jmp.2011.12.003

Dzhafarov, E. N., and Kujala, J. V. (2012b). Quantum entanglement and the issue of selective influences in psychology: an overview. *Lec. Notes Comput. Sci.* 7620, 184–195. doi: 10.1007/978-3-642-35659-9/17

Dzhafarov, E. N., and Kujala, J. V. (2013). All-possible-couplings approach to measuring probabilistic context. *PLoS ONE* 8:e61712. doi: 10.1371/journal.pone.0061712

Dzhafarov, E. N., and Kujala, J. V. (2014). On selective influences, marginal selectivity, and Bell/CHSH inequalities. *Top. Cogn. Sci.* 6, 121–128. doi: 10.1111/tops.12060

Haven, E., and Khrennikov, A. (2012). *Quantum Social Science*. Cambridge: Cambridge Press.

Haven, E., and Khrennikov, A. (2009). Quantum mechanics and violation of the sure-thing principle: the use of probability interference and other concepts. *J. Math. Psychol.* 53, 378–388. doi: 10.1016/j.jmp.2009.01.007

Khrennikov, A., and Basieva, I. (2014). Quantum model for psychological measurements: from the projection postulate to interference of mental observables represented as positive operator valued measures. *Neuroquantology* 12, 324–336. doi: 10.14704/nq.2014.12.3.750

Khrennikov, A., Basieva, I., Dzhafarov, E. N., and Busemeyer, J. R. (2014). Quantum models for psychological measurements: an unsolved problem. *PLoS ONE* 9:e110909. doi: 10.1371/journal.pone.0110909

Khrennikov, A. (2004). "Information dynamics in cognitive, psychological, social, and anomalous phenomena," in *Fundamental Theories of Physics* (Dordreht: Kluwer).

Khrennikov, A. (2008). The role of von Neumann and Lüders postulates in the Einstein, Podolsky, and Rosen considerations: comparing measurements with degenerate and nondegenerate spectra. *J. Math. Phys.* 49, 052102. doi: 10.1063/1.2903753

Khrennikov, A. (2010b). *Ubiquitous Quantum Structure: From Psychology to Finances*. Berlin; Heidelberg; New York: Springer.

Khrennikov, A. Y. (2010a). Modelling of psychological behavior on the basis of ultrametric mental space: encoding of categories by balls. *P-Adic Num. Ultram. Anal. Appl.* 2, 1–20. doi: 10.1134/S2070046610010012

Khrennikov, A. Y. (2014). Cognitive processes of the brain: an ultrametric model of information dynamics in unconsciousness. *P-Adic Num. Ultram. Anal. Appl.* 6, 293–302. doi: 10.1134/S2070046614040049

Laming, D. R. J. (1997). *The Measurement of Sensation No. 30*. Oxford: Oxford University Press.

Newman, L. S., Moskowitz, G. B., and Uleman, J. S. (1996). People as flexible interpreters: evidence and issues from spontaneous trait inference. *Adv. Exp. Soc. Psychol.* 28, 211–279.

Nielsen, M. A., and Chuang, I. L. (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

Ozawa, M. (1997). An operational approach to quantum state reduction. *Ann. Phys.* 259, 121–137. doi: 10.1006/aphy.1997.5706

Plotnitsky, A. (2006). *Reading Bohr: Physics and Philosophy*. Dordrecht: Springer.

Plotnitsky, A. (2009). *Epistemology and Probability: Bohr, Heisenberg, Schrödinger, and the Nature of Quantum-Theoretical Thinking*. Heidelberg; Berlin; New York: Springer.

Pothos, E. M., and Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* 36, 255–274. doi: 10.1017/S0140525X12001525

Shaji, A., and Sudarshan, E. C. G. (2005). Who's afraid of not completely positive maps? *Phys. Lett. A* 341, 48–54. doi: 10.1016/j.physleta.2005.04.029

von Helmholtz, H. (1866). *Treatise on Physiological Optics*. Transl. by Optical Society of America in English. New York, NY: Optical Society of America.

Wang, Z., and Busemeyer, J. R. (2013). A quantum question order model supported by empirical tests of an a priori and precise prediction. *Top. Cogn. Sci.* 5, 689–710. doi: 10.1111/tops.12040

Woodworth, R. S. (1921). *Psychology: A Study of Mental Life*. New York, NY: H. Holt.

Zeilinger, A. (2010). *Dance of the Photons: From Einstein to Quantum Teleportation*. New-York, NY: Farrar, Straus and Giroux.

Appendix

Do we Need Complete Positivity?

Nowadays theory of the DL-instruments is considered old-fashioned; the class of such instruments is considered to be too general: it contains mathematical artifacts which have no relation to real physical measurements and state transformations as back-reactions to these measurements. The modern theory of instruments is based on *the extendability postulate* (e.g., Busch et al., 1995; Ozawa, 1997; Nielsen and Chuang, 2000):

For any apparatus A_S corresponding to measurement of observable A on a system S and any system $\tilde{S}$ noninteracting with S there exists an apparatus $A_{S+\tilde{S}}$ representing measurement on the compound system $S + \tilde{S}$ such that

- $p(a_i|\rho \otimes r) = p(a_i|\rho)$;
- $(\rho \otimes r)_{a_i} = \rho_{a_i} \otimes r$

for any state ρ of S and any state r of $\tilde{S}$.

In physics this postulate is quite natural: if, besides the quantum system S which is the object of measurement, there is (somewhere in the universe) another system $\tilde{S}$ which is not entangled with S, i.e., their joint pre-measurement state has the form $\rho \otimes r$, then the measurement on S with the result a_i can be considered as measurement on $S + \tilde{S}$ as well with the same result a_i. It is clear that the back-reaction cannot change the state of $\tilde{S}$. Surprisingly this very trivial assumption has tremendous mathematical implications.

Since we proceed only in the finite dimensional case, the corresponding mathematical considerations are simplified. Consider an instrument E_S representing the state update as the result of the back-reaction from measurement on S. For each Γ, this is a linear map from $L(H) \rightarrow L(H)$, where H is the state space of S. Let W be the state space of the system $\tilde{S}$. Then the state space of the compound system $S+\tilde{S}$ is given by the tensor product $H \otimes W$. We remark that the space of linear operators in this state space can be represented as $L(H \otimes W) = L(H) \otimes L(W)$. Then the superoperator $E_S(\Gamma) : L(H) \rightarrow L(H)$ can be trivially extended to the superoperator $E_S(\Gamma) \otimes I : L(H \otimes W) \rightarrow L(H \otimes W)$. It is easy to prove that the state transformer corresponding to the apparatus for measurements on $S + \tilde{S}$ has to have this form $E_{S+\tilde{S}}(a_i) = E_S(a_i) \otimes I$. Hence, this operator also has to be positively defined. We remark that if the state space W has the dimension k, then the space of linear operators $L(W)$ can be represented as the space of $k \times k$ matrices which is further denoted as $\mathbf{C}^{k \times k}$.

Formally, a superoperator $T : L(H) \rightarrow L(H)$ is called *completely positive* if it is positive and each its trivial extension $T \otimes I : L(H) \otimes \mathbf{C}^{k \times k} \rightarrow L(H) \otimes \mathbf{C}^{k \times k}$ is also positive. There are natural examples of positive maps which are not completely positive (Nielsen and Chuang, 2000).

A CP quantum operation is a DL quantum operation which is additionally completely positive; a CP instrument is based on CP quantum operations representing back-reactions to measurement. As was pointed out, in modern literature only CP

quantum operations and instruments are in the use, so they are called simply quantum operations and instruments.

The main mathematical feature of (CP) quantum operations is that the class of such operations can be described in a simple way, namely, with the aid of the Kraus representation (Busch et al., 1995; Ozawa, 1997; Nielsen and Chuang, 2000):

$$T\rho = \sum_j V_j^\star \rho V_j, \tag{A1}$$

where (V_j) are some operators acting in H. Hence, for a (CP) instrument, we have: for each $a_i \in O$, there exist operators (V_{a_ij}) such that

$$E(a_i)\rho = \sum_j V_{a_ij}^\star \rho V_{a_ij}. \tag{A2}$$

Thus,

$$\rho_{a_i} = \frac{\sum_j V_{a_ij}^\star \rho V_{a_ij}}{\sum_j V_{a_ij}^\star \rho V_{a_ij}}, \tag{A3}$$

where the trace one condition (Equation 48) implies that

$$\sum_i \sum_j V_{a_ij}^\star V_{a_ij} = I. \tag{A4}$$

The corresponding POVMs M_{a_i} can be represented as

$$M_{a_i} = \sum_j V_{a_ij}^\star V_{a_ij}. \tag{A5}$$

This is a really elegant mathematical representation. However, it might be that this mathematical elegance, and not a real physical situation, has contributed to widespread use of CP in quantum information theory (cf. Shaji and Sudarshan, 2005).

Is the use of the extendability postulate justified in the operational approach to cognition?

Seemingly, not (although further analysis is required). Any concrete perception takes place at the conscious level, and it is based on interaction with the sensation of a visual image. The state of this sensation corresponds to the state of the system S in the above considerations. To be able to consider the state of another sensation, the analog of the state of the system $\tilde{S}$, the brain has to *activate* this sensation. Thus, we cannot simply consider all possible sensations as existing in some kind of the mental universe simultaneously. Hence, in general, sensations generated by different visual stimuli cannot be treated as existing simultaneously.

It is more natural to develop the theory of perception instruments as the theory of DL instruments and not CP instruments. In particular, although the Kraus representation can be used as a powerful analytic tool, we need not to overestimate its applicability for modeling of cognition.

Quantum-Like Bayesian Networks for Modeling Decision Making

Catarina Moreira * *and Andreas Wichert*

Department of Computer Science, Instituto Superior Técnico, University of Lisbon, INESC-ID, Lisbon, Portugal

In this work, we explore an alternative quantum structure to perform quantum probabilistic inferences to accommodate the paradoxical findings of the Sure Thing Principle. We propose a Quantum-Like Bayesian Network, which consists in replacing classical probabilities by quantum probability amplitudes. However, since this approach suffers from the problem of exponential growth of quantum parameters, we also propose a similarity heuristic that automatically fits quantum parameters through vector similarities. This makes the proposed model general and predictive in contrast to the current state of the art models, which cannot be generalized for more complex decision scenarios and that only provide an explanatory nature for the observed paradoxes. In the end, the model that we propose consists in a nonparametric method for estimating inference effects from a statistical point of view. It is a statistical model that is simpler than the previous quantum dynamic and quantum-like models proposed in the literature. We tested the proposed network with several empirical data from the literature, mainly from the Prisoner's Dilemma game and the Two Stage Gambling game. The results obtained show that the proposed quantum Bayesian Network is a general method that can accommodate violations of the laws of classical probability theory and make accurate predictions regarding human decision-making in these scenarios.

Keywords: Bayesian networks, decision making, quantum probability, quantum cognition, sure thing principle

1. INTRODUCTION

The present work proposes a new model to make predictions in paradoxical situations where the Sure Thing Principle is being violated. The Sure Thing Principle (Savage, 1954) is a fundamental principle in economics and probability theory and states that if one prefers action A over B under state of the world X, and if one also prefers A over B under the complementary state of the world, $\neg\, X$, then one should always prefer action A over B even when the state of the world is unspecified. Several experiments have shown that people violate this principle in decisions under uncertainty, leading to paradoxical results and violations of the classical law of total probability (Tversky and Kahnenman, 1974; Tversky and Kahneman, 1983; Tversky and Shafir, 1992; Aerts et al., 2004; Birnbaum, 2008).

1.1. Motivation

More recently, cognitive scientists have turned to quantum probability theory in order to accommodate these paradoxical findings. Although many models have been proposed in the literature, most of them cannot be considered predictive. Most of these models require a set of quantum parameters to be fitted and, so far, the only way these models have to fit the parameters is to use the final outcome of the experiment to set the parameters in order to explain that outcome.

Moreover, these models cannot scale to more complex decision scenarios, because the number of parameters is exponentially large (Khrennikov, 2003a, 2004, 2006) or because of computational constraints in the computation of very large unitary operators (Busemeyer et al., 2006b, 2009; Pothos and Busemeyer, 2009).

1.2. Contributions

For these reasons, in this work, we propose a network structure framework that can easily scale to more complex decision scenarios. In other words, we propose a quantum-like Bayesian Network formalism, which consists in replacing classical probabilities by quantum probability amplitudes. However, since this approach also suffers from the problem of exponential growth of quantum parameters that need to be fit, we also propose a similarity heuristic (Shah and Oppenheimer, 2008) that automatically computes this exponential number of quantum parameters through vector similarities. A Bayesian Network can be understood as an acyclic directed graph, in which each node represents a random variable and each edge represents a direct causal influence from the source node to the target node (conditional dependence).

In this article, we will address the problem of violations to the Sure Thing Principle by examining two major problems in which these violations were verified: the Prisoner's Dilemma game and the Two Stage Gambling game. These violations were initially reported by Tversky and Shafir (1992) and later simulated in several works in the literature that also reported similar results (Li and Taplin, 2002; Busemeyer et al., 2006a; Hristova and Grinberg, 2008). We will show how the current classical models fail to explain the paradoxical findings implied in the violations of the Sure Thing Principle and we will make a more deep discussion about the drawbacks of the most representative quantum-like models in the literature.

1.3. Research Questions

With the present work, we intend to address the following research questions. An answer to these questions is given in Section 8.

1. Why do we need another quantum-like model to explain violations to the Sure Thing Principle?
2. What is the advantage of the proposed approach? How can it make a difference toward the current well-established quantum models that have been proposed in the literature?

2. VIOLATIONS OF THE SURE THING PRINCIPLE

In this section, we present two experiments from the literature, in which it was observed violations to the Sure Thing Principle and consequently to the laws of classical probability theory and logic. The two experiments are the Prisoner's Dilemma game and the Two Stage Gambling game.

2.1. The Prisoner's Dilemma Game

The Prisoner's Dilemma game corresponds to an example of the violation of the Sure Thing Principle. In this game, there are two prisoners who are in separate solitary confinements with no means of speaking to or exchanging messages with the other. The police offer each prisoner an agreement: each prisoner is given the opportunity either to betray the other (*Defect*), by testifying that the other committed the crime, or to *Cooperate* with the other by remaining silent.

In order to test the veracity of the Sure Thing Principle under the Prisoner's Dilemma game, an experiment was made in which three conditions were tested:

- Participants were informed that the other participant chose to *Defect*.
- Participants were informed that the other participant chose to *cooperate*.
- Participants had no information about the other participant's decision.

Table 1 summarizes the results of several works of the literature, which have performed this experiment. Note that all entries of **Table 1** show a violation of the law of total probability. According to the total law of probability, it is expected that:

$$Pr(P_2 = Defect \mid P_1 = Defect) \geq Pr(P_2 = Defect)$$
$$\geq Pr(P_2 = Defect \mid P_1 = Cooperate)$$

Note that, $Pr(P_2 = Defect \mid P_1 = Defect)$ corresponds to the probability of the second player choosing the *Defect* action given that he knows that the first player chose to *Defect*. In **Table 1** this corresponds to the entry *Known to Defect*. In the same way, $Pr(P_2 = Defect \mid P_1 = Cooperate)$ corresponds to the entry *Known to Collaborate*. The observed probability during the experiments concerned with player 2 choosing to *Defect*, $Pr(P2 = Defect)$, corresponds to the entry *unknown* of **Table 1**, since there is no evidence about the first player's actions. Finally, the entry *Classical Probability* corresponds to the classical probability $Pr(P_2 = Defect)$, which is computed through the law of total probability:

TABLE 1 | Works of the literature reporting the probability of a player choosing to *Defect* under several conditions for the Prisoner's Dilemma Game: when the action of the second player is known to be *Defect* (Known to Defect), when the action of the second player is known to be *Cooperate* (Known to Collaborate), and when the action of the second player is not known (Unknown).

Literature	Known to Defect	Known to Collaborate	Unknown	Classical probability
Shafir and Tversky, 1992	0.9700	0.8400	0.6300	0.9050
Croson, 1999[a]	0.6700	0.3200	0.3000	0.4950
Li and Taplin, 2002[b]	0.8200	0.7700	0.7200	0.7950
Busemeyer et al., 2006a	0.9100	0.8400	0.6600	0.8750
Hristova and Grinberg, 2008	0.9700	0.9300	0.8800	0.9500
Average	0.8700	0.7400	0.6400	0.8050

a corresponds to the average of the results reported in the first two payoff matrices of the work of Croson (1999).
b corresponds to the average of all seven experiments reported in the work of Li and Taplin (2002).

Literature	Known to Win	Known to Lose	Unknown	Classical probability
Tversky and Shafir, 1992	0.69	0.58	0.37	0.6350
Kuhberger et al., 2001	0.72	0.47	0.48	0.5950
Lambdin and Burdsal, 2007	0.63	0.45	0.41	0.5400
Average	0.68	0.50	0.42	0.5900

$$Pr(P_2 = Defect) = Pr(P_1 = Defect) \cdot Pr(P_2 = Defect|$$
$$P_1 = Defect) +$$
$$+ Pr(P_1 = Cooperate) \cdot Pr(P_2 = Defect|P_1$$
$$= Cooperate)$$

2.2. The Two Stage Gambling Game

The Two Stage Gambling game is another game that shows violations of the Sure Thing Principle. In this game, participants were asked at each stage to make the decision of whether or not to play a gamble that has an equal chance of winning $200 or losing $100. Three conditions were verified:

1. Participants were informed if they had won the first gamble;
2. Participants were informed if they had lost the first gamble;
3. Participants did not know the outcome of the first gamble;

The overall results revealed that participants who knew that they won the first gamble, decided to play again. Participants who knew that they lost the first gamble, also decided to play again. Through Savage's Sure Thing Principle, it was expected that the participants would choose to play again, even if they did not know the outcome of the first gamble. However, the results obtained revealed something different. If the participants did not know the outcome of the first gamble, then many of them decided not to play the second one.

We conclude this section by clarifying why we will only validated the proposed quantum-like Bayesian Network in small decision problems (such as the Prisoner's Dilema and the Two Stage Gambling Game), since we are defending a general quantum-like structure that is able to deal with complex decision scenarios. We used small decision scenarios, because we cannot find literature showing violations to the Sure Thing Principle for more complex decision scenarios. Actually, after performing some research, we believe that the violations of the Sure Thing Principle tend to diminish with the complexity of the decision scenario. Imagine for instance a three stage gambling game. It will be very hard to find significant data that shows a player wishing to play the last gamble, given that he has lost the two previous gambles. **Table 2** shows the results obtained in several works of the literature.

3. VIOLATION OF THE SURE THING PRINCIPLE: CLASSICAL APPROACHES

There are many classical approaches that could be used to try to accommodate violations to the Sure Thing Principle. Two of these main models are the Classical Markov Models and the Classical Bayesian Networks. In this section, we will describe how these two models work and we will explain why they cannot be used to simulate violations to the Sure Thing Principle.

3.1. Classical Markov Model

A Markov Model can be generally defined as a stochastic probabilistic undirected graphical model that satisfies the Markov property. This means that the process evolves (and tries to perform a prediction) based only on the present state. The current state is independent of any past or future states. These probabilistic models are very useful to model systems that change states according to a transition matrix that specifies some probability distribution or some transition rules that depend solely on the current state.

The initial state is given by a vector, which contains the probabilities of each event occurring. This vector requires that the sum of these probabilities is one.

$$P_I = \begin{bmatrix} a_0 & a_1 & \dots & a_n \end{bmatrix} \cdot \frac{1}{\sum_i a_i}$$

The state transition is represented by a differential equation, which consists in the multiplication of this initial probability state P_I by a transition function $T(t)$. This function is represented by a matrix containing positive real numbers and with the constraint that each row must sum to one (normalization axiom). In other words, this matrix represents the new probability distribution across all possible outcomes through some time period t (Pothos and Busemeyer, 2009).

$$\frac{d}{dt} T(t) = K \cdot T(t) \quad (1)$$

The intensity matrix K corresponds to the problem's settings. For instance, for the Prisoner's Dilemma Game, it represents the payoffs of each player, in the Two Stage Gambling Game, it represents the rewards/losses that the player can have in each gamble. A solution to this equation is given by Equation 2, which allows one to construct a transition matrix for any time point from the fixed intensity matrix. In other words, the intensity matrix performs a transformation in the probabilities of the current state in order to favor a certain action in the decision problem.

$$T(t) = e^{K.t} \quad (2)$$

In the end, we can compute the solution for the probability distribution over time by multiplying the transition matrix by the initial probability state.

$$P_F(t) = e^{K.t} \cdot P_I(0) \quad (3)$$

In Equation 3, we do not need to perform any normalization in the end, because the operation in Equation 1 together with the intensity matrix K assure that the values computed are already probability values.

Since, in the end, the Markov Model has to obey to the rules of probability theory and set theory, even if we parameterize the intensity matrix K, we would find that there are no values that could explain the violations of the Sure Thing Principle without violating the laws of classical probability theory. Some studies have been proposed in the literature demonstrating that the classical Markov Model cannot accommodate violations to the Sure Thing Principle (Busemeyer et al., 2009; Pothos and Busemeyer, 2009).

3.2. Classical Bayesian Networks

A classical Bayesian Network can be defined by a directed acyclic graph structure in which each node represents a different random variable from a specific domain and each edge represents a direct influence from the source node to the target node. The graph represents independence relationships between variables and each node is associated with a conditional probability table, which specifies a distribution over the values of a node given each possible joint assignment of values of its parents. This idea of a node, depending directly from its parent nodes, is the core of Bayesian Networks. Once the values of the parents are known, no information relating directly or indirectly to its parents or other ancestors can influence the beliefs about it (Koller and Friedman, 2009).

A Bayesian Network can be understood as the representation of a full joint probability distribution through conditional independence statements. This way, a Bayesian Network can be used to answer any query about the domain by combining (adding) all relevant entries from the joint probability.

The full joint distribution (Russel and Norvig, 2010) of a Bayesian Network, where X is the list of variables, that is, the set of nodes of the Bayesian Network and is given by:

$$Pr(X_1, \ldots, X_n) = \prod_{i=1}^{n} Pr(X_i | Parents(X_i)) \qquad (4)$$

The formula for computing classical exact inferences on Bayesian Networks is based on the full joint distribution (Equation 4). Let e be the list of observed variables (nodes) and let Y be the remaining unobserved variables (nodes) in the network. For some query X, the inference is given by Equation 5. Note that, $Pr(X, e, y)$ corresponds to the full joint probability distribution.

$$Pr(X|e) = \alpha \left[\sum_{y \in Y} Pr(X, e, y) \right] \qquad (5)$$

$$\text{Where} \quad \alpha = \frac{1}{\sum_{x \in X} Pr_c(X = x, e)}$$

The summation is over all possible y, i.e., all possible combinations of values of the unobserved variables y. The

α parameter, corresponds to the normalization factor for the distribution $Pr(X|e)$ (Russel and Norvig, 2010). This normalization factor comes from some assumptions that are made in Bayes rule.

One might think that if we parameterize the Bayesian Network, it could be possible to explain the paradoxical findings of the Sure Thing Principle. This line of thought is legitimate, however one must take into account that in the end, the probabilistic inferences computed through the Bayesian Network must obey set theory and to the law of total probability. This means that, even if we parameterize the network, we could not find any closed form optimization that would accommodate violations to the Sure Thing Principle.

4. VIOLATION OF THE SURE THING PRINCIPLE: QUANTUM-LIKE APPROACHES

In this section, we introduce the most import quantum decision models that have been proposed in the literature that can accommodate the violations to the Sure Thing Principle. The models that we describe in this section are the following: the Quantum Dynamical Model (Section 4.1), the Quantum-Like Approach (Section 4.2), and the Quantum Prospect Decision Theory (Section 4.3).

4.1. The Quantum Dynamical Model

The Quantum Dynamical Model was originally proposed by Busemeyer (Busemeyer et al., 2009; Pothos and Busemeyer, 2009) and consists on a general framework that corresponds to a quantum version of a classical dynamical Markov model. The Quantum Dynamical Model takes into account time evolution. Quantum interference effects are also taken into account though a superposition of paths.

The initial belief state corresponds to a quantum state representing a superposition of the participant's beliefs in the form of a vector. The term ψ corresponds to a quantum probability amplitude.

$$P_I = \begin{bmatrix} \psi_0 & \psi_1 & \ldots & \psi_n \end{bmatrix} \cdot \frac{1}{\sum_i |\psi_i|^2} \qquad (6)$$

Next, we need to create a unitary matrix. In quantum mechanics, a unitary matrix restricts the allowed evolution of quantum systems, ensuring that the sum of probabilities of all possible outcomes of any event is always 1. This means that the matrix must be orthonormal (the rows are mutually orthogonal unit vectors, as are the columns). In the Quantum Dynamical Model, this matrix encodes all state transitions that a person can experience while choosing a decision. The unitary matrix U is computed by a differential equation called Schrödinger's equation.

$$\frac{\delta}{\delta t} U(t) = -i \cdot H \cdot U(t) \qquad (7)$$

The parameter t corresponds to the time evolution. Under the Dynamical Quantum Model, this parameter is set to $\pi/2$,

corresponding to the average time that a participant takes to make a decision (approximately 2 s) (Pothos and Busemeyer, 2009). The matrix H is called the Hamiltonian matrix, which must be Hermitian in order to generate a unitary matrix.

$$U(t) = \exp(-i \cdot H \cdot t) \tag{8}$$

By multiplying the unitary matrix with the initial superposition belief state, one can compute the transition of the participants' beliefs at each time. The final vector Q_F represents the amplitude distribution across states after deliberation.

In the end, we can compute the solution for the probability distribution over time by multiplying the transition matrix by the initial probability state.

$$Q_F = U \cdot Q_i = e^{-i \cdot H \cdot t} \cdot Q_I(0) \tag{9}$$

In Equation 9, we do not need to perform any normalization in the end, because the operation in Equation 8 together with the intensity matrix H assure that the values computed are in accordance with the normalization axiom.

4.2. The Quantum-Like Approach

The Quantum-Like Approach has its roots in contextual probabilities. This model was proposed by Khrennikov and corresponds to a general contextual probability space from which the classical and quantum probability models can be derived (Khrennikov, 2009b, 2010).

In the Quantum-Like Approach, the context relates to the circumstances that form the setting for an event in terms of which it can be fully understood, clarifying the meaning of the event. More specifically, it is a complex of conditions under which a measurement is performed. For instance, in domains outside of physics, such as cognitive science, one can have mental contexts. In social sciences, we can have a social context. And the same idea is applied to many other domains, such as economics, politics, game theory, biology, etc. (Khrennikov, 1999, 2001, 2003b, 2005a,b).

The Quantum-Like Approach corresponds to a contextual probabilistic model given by $M = (\mathcal{C}, \mathcal{O}, \pi(\mathcal{O}, \mathcal{C}))$. Where $\mathcal{C}$ is a set of contexts, $\mathcal{O}$ is the set of observables and $\pi(\mathcal{O}, \mathcal{C})$ corresponds to a probability distribution of some observables belonging to a specific context. Associated with a context, there are a set of observables. In quantum mechanics, an observable corresponds to a self-adjoint operator on a complex Hilbert Space. Under the Quantum-Like Approach, these observables correspond to the set of possible events with their respective values.

Let's assume, for a context $C \in \mathcal{C}$, that there are two dichotomous observables $a, b \in \mathcal{O}$, and each of these observables can take some values $\alpha \in a$ and $\beta \in b$, respectively.

The Quantum-Like Approach can be built from the general structure of the quantum law of total probability. The quantum law of total probability is very similar to the classical law of total probability, except that it uses complex amplitudes instead of real probability values. In order to obtain a probability value, the magnitude of the quantum amplitude must be

squared Busemeyer and Bruza (2012). This will generate an additional term called the *interference term*. This term does not exist in classical probability and enables the representation of interferences between quantum states.

$$Pr(b = \beta) = Classical\_Probability(b = \beta) + Interference\_Term \tag{10}$$

Under this representation, we can replace *Classical_Probability* by the classical law of total probability, and also replace the quantum *Interference_Term* by a measure of supplementary, represented by $\delta(\beta|a, C)$.

If we perform the normalization of the probability measure of supplementary $\delta(\beta|a, C)$ by the square root of the product of all probabilities, we obtain:

$$\lambda_\theta = \frac{\delta(\beta|a, C)}{2\sqrt{\prod_{\alpha \in a} Pr(a = \alpha|C)Pr(b = \beta|a = \alpha, C)}} \tag{11}$$

From Equation 11, the general probability formula of the Quantum-Like Approach can be derived. For two variables, is given by:

$$Pr(b = \beta|C) = \sum_{\alpha \in a} Pr(a = \alpha|C)Pr(b = \beta|a = \alpha, C)$$
$$+2\lambda_\theta \sqrt{\prod_{\alpha \in a} Pr(a = \alpha|C)Pr(b = \beta|a = \alpha, C)} \tag{12}$$

If we look closely to Equation 12, we will see that the first summation of the formula corresponds to the classical law of total probability. The second term of the formula (the one that contains the λ_θ parameter), does not exist in the classical model and it is called the interference term.

In a quantum context, since the supplementary term $\delta(\beta|a, C)$ is being normalized in a quantum fashion, then we automatically know that the indicator term λ_θ will always have to be smaller than 1 in order to obtain quantum probabilities, $\lambda_\theta \leq 1$. So, under trigonometric contexts, the Quantum-Like Approach for quantum probabilities becomes:

$$\lambda_\theta = \cos(\theta) \rightarrow Pr(\beta|C) = \sum_{\alpha \in a} Pr(\alpha|C)Pr(\beta|\alpha, C)$$
$$+2\sqrt{\prod_{\alpha \in a} Pr(\alpha|C)Pr(\beta|\alpha, C)} \cos(\theta) \tag{13}$$

Equation 13 can be simplified in the following way:

$$Pr(\beta|C) = \left| \sqrt{Pr(\alpha_1|C)Pr(\beta|\alpha_1, C)} \right.$$
$$\left. + e^{i\theta_{\beta|\alpha, C}} \sqrt{Pr(\alpha_2|C)Pr(\beta|\alpha_2, C)} \right|^2 \tag{14}$$

Equation 14 corresponds to the representation of the quantum law of total probability. In this equation, the angle $\theta_{\beta|\alpha, C}$ corresponds to the phase of a random variable and incorporates the phase of both $A = \alpha_1$ and $A = \alpha_2$ in the following way: $\theta_{\beta|\alpha, C} = \theta_{\beta|\alpha_1} - \theta_{\beta|\alpha_2}$.

One should note that, the Quantum-Like Approach can be extended to more complex decision scenarios, that is, with more than two random variables. However, this will lead to the very difficult task of tuning an exponential number of quantum θ parameters. Peter Nyman noticed this problem when he generalized the Quantum-Like Approach for three dichotomous variables (Nyman, 2010, 2011b; Nyman and Basieva, 2011a,b).

4.2.1. The Hyperbolic Interference

Although the Quantum-Like Approach provides great possibilities comparing with the classical one, it seems that it cannot cover completely data from psychology and that a quantum formalism was not enough to explain some paradoxical findings (see Khrennikov et al., 2014), so hyperbolic spaces were proposed (Khrennikov, 2005c; Nyman, 2011a,b).

From Equation 12, if $Pr(b = \beta) - \sum_{\alpha \in a} Pr(a = \alpha|C)Pr(b = \beta|a = \alpha, C)$ is different from zero, then some interference effects occur. In order to determine which type of interference happened, one tests the Quantum-Like Approach for quantum probabilities. This can be determined by normalizing the supplementary measure in a quantum fashion, just like presented in Equation 11.

If the probability $Pr(b = \beta)$ was not computed in a trigonometric space (that is, it is not quantum), then, it is straightforward that the quantum normalization applied in Equation 11 will give a value bigger than 1. Since we are not in the context of quantum probabilities, the quantum normalization factor will fail to normalize the interference term, and will produce a number bigger than the normalization factor. Under these circumstances, the Quantum-Like Approach incorporates the generalization of *hyperbolic probabilities*, arguing that the context in which these probabilities were computed was in a Hyperbolic context (Khrennikov, 2009a, 2010; Nyman, 2011a).

Under Hyperbolic contexts, the Quantum-Like Approach contextual probability formula becomes:

$$\lambda_\theta = \cosh(\theta) \quad \rightarrow \quad Pr(\beta|C) = \sum_{\alpha \in a} Pr(\alpha|C)Pr(\beta|\alpha, C)$$
$$\pm 2\sqrt{\prod_{\alpha \in a} Pr(\alpha|C)Pr(\beta|\alpha, C)} \cosh(\theta) \quad (15)$$

In summary, according to the values computed by the indicator function λ_θ, the Växjö Model enables the computation of probabilities in the following contexts:

- If $|\lambda_\theta| = 0$, then there is no interference and the Växjö Model collapses to classical probability theory.
- If $|\lambda_\theta| \leq 1$, then we fall into the realm of quantum mechanics and the context becomes a Hilbert space. The indicator function is then replaced by the trigonometric function $\cos(\theta)$.
- If $|\lambda_\theta| > 1$, then we fall into the realm of hyperbolic numbers and the context becomes a hyperbolic space. The indicator function is then replaced by the hyperbolic function $\cosh(\theta)$.

4.3. The Quantum Prospect Decision Theory

The Quantum Prospect Decision Theory was developed by Yukalov and Sornette (2008, 2011) and developed throughout many other works (Yukalov and Sornette, 2009a,b, 2010a,b). The foundations of this theory are very similar to the previously presented Quantum-Like Approach.

In the Quantum-Like Approach, we start with two dichotomous observables. In the Quantum Prospect Decision Theory, these observables are referred to *intensions*. An intension can be defined by an intended action and a set of intended actions is defined by a *prospect*.

Each prospect can contain a set of action modes, which are concrete representations of an intension. Making a comparison with the Quantum-Like Approach, a prospect can be seen as a random variable and the set of action modes are the assignments that each random variable can have. For instance, the intension *to play* can have two representations: *play action A* or *play action B*.

Following the work of Yukalov and Sornette (2011), two intensions A and B with the respective representations: $A = x$ where $x \in a_1, a_2$ and $B = y$, where $y \in b_1, b_2$. The corresponding state of mind is given by:

$$| \psi_s(t) \rangle = \sum_{i,j} c_{i,j}(t) | A_i B_j \rangle \quad (16)$$

Equation 16 represents a linear combination of the prospect basis states. From a psychological perspective, the state of mind is a fixed vector characterizing a particular decision maker with his/her beliefs, habits, principles, etc. That is, it describes each decision maker as a unique subject.

The prospect states corresponding to the intensions A and B are given by Equation 17. The ψ symbol corresponds to quantum amplitudes associated with the prospect state. Under the Quantum Prospect Decision Theory, these amplitudes represent the weights of the intended actions, while a person is still deliberating about them.

$$|\pi_{A=a_1}\rangle = \psi_{11}|A = a_1 B = b_1\rangle + \psi_{12}|A = a_1 B = b_2\rangle$$
$$|\pi_{A=a_2}\rangle = \psi_{21}|A = a_2 B = b_1\rangle + \psi_{22}|A = a_2 B = b_2\rangle \ (17)$$

The probabilities of the prospects can be obtained by computing the squared magnitude of the prospect states (just like in the Quantum-Like Approach and the Quantum Dynamical Model). Consequently, the final probabilities are given by:

$$Pr(\pi_{A=a_1}) = Pr(A = a_1, B = b_1) + Pr(A = a_1, B = b_2)$$
$$+ Interference_{A=a_1}$$
$$Pr(\pi_{A=a_2}) = Pr(A = a_2, B = b_1) + Pr(A = a_2, B = b_2) \quad (18)$$
$$+ Interference_{A=a_2}$$

Where the interference term in defined by:

$$Interference_{A=a_1} = 2 \cdot \varphi(\pi_{A=a_1})\sqrt{Pr(A = a_1, B = b_1)} \cdot$$
$$\sqrt{Pr(A = a_1, B = b_2)}$$
$$Interference_{A=a_2} = 2 \cdot \varphi(\pi_{A=a_2})\sqrt{Pr(A = a_2, B = b_1)} \cdot \quad (19)$$
$$\sqrt{Pr(A = a_2, B = b_2)}$$

In Equation 19, the symbol φ corresponds to the uncertainty factor and is given by:

$$\varphi(\pi_{A=a_1}) = \cos\left(arg\left(\psi_{11} \cdot \psi_{12}\right)\right)$$
$$\varphi(\pi_{A=a_2}) = \cos\left(arg\left(\psi_{21} \cdot \psi_{22}\right)\right) \quad (20)$$

The interference term corresponds to the effects that emerge during the process of deliberation, that is, while a person is making a decision. These interference effects result from conflicting interests, ambiguity, emotions, etc. (Yukalov and Sornette, 2011).

One can notice that the Quantum Prospect Decision Theory is very similar to the Quantum-Like Approach proposed by Khrennikov (2009c). Both theories end up with the same quantum probability formula. However, the Quantum Prospect Decision Theory provides some heuristics in how to choose the uncertainty factors. This information will be addressed in the next section.

4.3.1. Choosing the Uncertainty Factor

In order to accommodate the violations of the Sure Thing Principle, the uncertainty factor must be set in such a way that it will enable accurate predictions. Two methods were proposed by Yukalov and Sornette (2011) to estimate the uncertainty factor: the Interference Alternation method and the Interference Quarter Law.

- **Interference Alternation**—Under normalized conditions, the probabilities of the prospects $p\left(\pi_j\right)$ must sum to 1. This normalization only occurs if one characterizes the interference term as an *alternation*, such that the interference effects disappear while summing the probability of the prospects. The interference alternation property is in accordance with the findings of Epstein (1999): the destructive interference effects can be associated with uncertainty aversion. This leads to a less probable action under uncertainty conditions. In contrast, the probabilities of other actions that contain less uncertainty are enhanced through constructive quantum interference effects. This uncertainty aversion happens quite frequently in situations where the Sure Thing Principle is violated. This implies that one of the probabilities of the prospects must be enhanced, whereas the other must be decreased.

$$sign\left[\varphi(\pi_{A=a_1})\right] = -sign\left[\varphi(\pi_{A=a_2})\right] \quad \text{where}$$
$$\left|\varphi(\pi_{A=a_i})\right| \in [0, 1] \quad (21)$$

- **Interference Quarter Law**—the interference terms generated by quantum probabilistic inferences, have a free quantum parameter, which is the uncertainty factor. The Interference Quarter Law corresponds to a quantitative estimation of this parameter. The modulus of the interference term q can be quantitatively estimated by computing the expectation value of the probability distribution of a random variable ξ in the interval $[0, 1]$.

$$q \equiv \int_0^1 \xi \cdot pr\left(\xi\right) d\xi = \frac{1}{4} \quad (22)$$

The probability distribution $p(\xi)$ is given by Equation 22 and can be computed by making the average of two probability distributions.

$$Pr\left(\xi\right) = \frac{1}{2}\left[pr_1\left(\xi\right) + pr_2\left(\xi\right)\right] = \delta\left(\xi\right) + \frac{1}{2}\Theta\left(1 - \xi\right) \quad (23)$$

4.4. Quantum-Like Bayesian Networks in the Literature

There are two main works in the literature that have contributed to the development and understanding of Quantum Bayesian Networks. One belongs to Tucci (1995) and the other to Leifer and Poulin (2008).

In the work of Tucci (1995), it is argued that any classical Bayesian Network can be extended to a quantum one by replacing real probabilities with quantum complex amplitudes. This means that the factorization should be performed in the same way as in a classical Bayesian Network.

One big problem with Tucci's work was the lack of methods to set the phase parameters. The author states that, one could have infinite Quantum Bayesian Networks representing the same classical Bayesian Network depending on the values that one chooses to set the parameter. This requires that one knows *a priori* which parameters would lead to the desired solution for each node queried in the network (which we never know). So, for these experiments, Tucci's model (Tucci, 1995) cannot predict the results observed, since one does not have any information about the quantum parameters.

In the work of Leifer and Poulin (2008), the authors argue that, in order to develop a quantum Bayesian Network, a quantum version is required of probability distributions, quantum marginal probabilities and quantum conditional probabilities (**Table 3**). The authors made a preliminary study on these concepts. Generally speaking, a quantum probability distribution corresponds to a density matrix contained in a Hilbert space, with the constraint that the trace of this matrix must sum to 1. In quantum probability theory, a full joint distribution is given by a *density matrix* ρ. This matrix provides the probability distribution of all states that a Bayesian Network can have. The marginalization operation corresponds to a quantum partial trace (Nielsen and Chuang, 2000; Rieffel and Polak, 2011). In the end, these models from the literature fail to provide any advantage relatively to the classical models, because they cannot take into account interference effects between random variables. So, they provide no advantages in modeling decision-making problems that try to predict decisions that violate the laws of total probability.

TABLE 3 | Relation between classical and quantum probabilities used in the work of Leifer and Poulin (2008).

	Classical probability	Quantum probability
State	$Pr(A)$	$\left|e^{j\theta}\psi_A\right|^2$
Joint probability distribution	$Pr(A, B)$	ρ_{AB}
Marginal probability distribution	$Pr(B) = \sum_A Pr(A, B)$	$\rho_B = Tr_A(\rho_{AB})$
Conditional state	$Pr(B|A)$	$\rho_{B|A}$
	$\sum_{b \in B} Pr(b|A) = 1$	$Tr(\rho_{B|A}) = I_A$

5. PROBLEMS WITH CURRENT CLASSICAL AND QUANTUM-LIKE APPROACHES

In this section, we summarize the three main models that were presented in the previous sections (**Table 4**) and point out the advantages and disadvantages of each one of them.

The Quantum-Like Approach is a very simple framework that enables the computation of quantum probabilities by performing the direct mapping between classical real probabilities and quantum probability amplitudes through Born's rule (Zurek, 2005, 2011). Although this model can be extended for N random variables and also go beyond quantum probabilities by incorporating hyperbolic spaces, this model cannot be called predictive, since there are no mechanisms to estimate the quantum θ parameters. One is required to know a priori the outcome of the decision scenario in order to fit the quantum parameters. So, this model has an explanatory nature in what concerns accommodating the paradoxical findings derived from violations of the Sure Thing Principle.

The Quantum Dynamical Model provides an elegant framework that can estimate decisions though time evolution. However, it also suffers from a major disadvantage related to Hamiltonian matrices. Creating a manual Hamiltonian is a very hard problem. It is required that all possible interactions of the decision problem are known and this specification must be made in such a way that the matrix is double stochastic. For more complex decision scenarios, this process is intractable. Furthermore, the Hamiltonian matrix grows exponentially with the complexity of the decision problem and the computation of a Unitary operator from such matrices is a very complex process. Most of the times, approximations are used, because of the complexity of the calculations involved in the matrix exponentiation operation.

The Quantum Prospect Decision Theory is a model very similar to the Quantum-Like Approach, but it is not extended to the hyperbolic spaces. The main advantage of the Quantum Prospect Decision Theory toward the other known quantum models is its predictive nature. The Quantum-Like Approach and the Quantum Dynamical model are more explanatory models. That is, they require that the outcome of an experiment is known in order to fit the parameters of the model and explain the paradoxical findings. The Quantum Prospect Decision Theory, on the other hand, contains an heuristic (the interference quarter law) that enables the estimation of the quantum parameters, turning the model predictive. However, the interference quarter law is a static heuristic. This means that, independently of the decision scenario and independently of the complexity of the decision, this interference term remains constant for every problem.

All of the above models exhibit different growth rates in parameters. For instance, the Dynamical Model parameterizes actions plus an additional parameter to model cognitive dissonance effects. So the number of parameters would be static if we consider the N-Person Prisoner's Dilemma Game. That is, instead of having only 2 players, this would be extended to N players. In the case of the Quantum-Like Approach, we would have 2^N parameters for the N-Person Prisoner's Dilemma Game. The number 2 comes from the fact that each player has two actions (either *Defect* or *Cooperate*). The same applies to the Quantum-Like Bayesian Networks and to the Quantum Prospect Theory Model. If we extend these models for N random variables, the number of parameters grows at a rate of $N_{actions}^{N_{person}}$, but these parameters will be automatically set using the Law of Quantum Interference, in the case of the Quantum Prospect Theory. The same is applied to the proposed Quantum-Like Bayesian Network, but instead of a static heuristic, we automatically set these parameters using a dynamic heuristic.

At this point, the reader might be thinking that the Quantum Dynamical Model provides great advantages toward the existing models, since the number of parameters required corresponds to the players actions with an additional cognitive dissonance parameter. Although this line of thought is correct, one should also take into account how the model unfolds. Although the numbers of parameters do not grow exponentially large as in the Quantum-Like Approach, the size of the Hamiltonian does. In fact, it grows exponentially large with the following size: $N_{actions}^{N_{players}} \times N_{actions}^{N_{players}}$, where $N_{actions}$ represents the number of actions of the players and $N_{players}$ corresponds to the number of players. The computation of a unitary operator from such matrices is a very complex process. Most of the times, approximations are used, because of the complexity of the calculations involved in the matrix exponentiation operation. **Table 5** summarizes the parameter growth rate of each approach.

TABLE 4 | Summary of the most relevant quantum decision models of the literature.

Model	State representation	Quantum interference	Predictive	Comments
Quantum dynamical model	Superposition of subject's beliefs	Shröedinger's equation	No . requires manual fit	Enables time evolution
Quantum-Like approach	Contextual probabilities (observables/random variables)	Measure of supplementarity	No. requires manual fit	Can deal with hyperbolic spaces
Quantum prospect decision theory	Contextual probabilities (prospects/random variables)	Interference quarter law	Yes. Uses a static heuristic	It is predictive uses a heuristic
Quantum-Like Bayesian networks	Contextual probabilities (observables/random variables)	None	No. manual fit	Can easily scale to more complex scenarios

	Approach	Parameter growth	Comments
Busemeyer et al., 2009; Pothos and Busemeyer, 2009	Quantum dynamical model	$N_{actions}$	Hamiltonian size exponential: $N_{actions}^{Nperson}$
Khrennikov, 2009c	Quantum-Like approach	$N_{actions}^{Nperson}$	Number of parameters Grows exponentially large
Yukalov and Sornette, 2011	Quantum prospect decision theory	$N_{actions}^{Nperson}$	Static heuristic Interference quarter law

For these reasons, in this work, we propose a network structure framework that can easily scale to more complex decision scenarios. In other words, we propose a quantum-like Bayesian Network formalism, which consists in replacing classical probabilities by quantum probability amplitudes. However, since this approach also suffers from the problem of exponential growth of quantum parameters that need to be fit, we also propose a similarity heuristic that automatically computes this exponential number of quantum parameters (Shah and Oppenheimer, 2008).
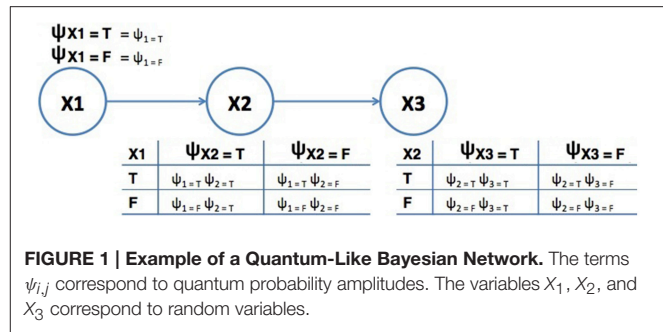
6. A QUANTUM-LIKE BAYESIAN NETWORK FOR DECISION AND COGNITION

The reason why we chose Bayesian Networks is because it provides a link between probability theory and graph theory. And a fundamental property of graph theory is its modularity: one can build a complex system by combining smaller and simpler parts. It is easier for a person to combine pieces of evidence and to reason about them, instead of calculating all possible events and their respective beliefs (Griffiths et al., 2008). In the same way, Bayesian Networks represent the decision problem in small modules that can be combined to perform inferences. Only the probabilities, which are actually needed to perform the inferences, are computed.

A Quantum-Like Bayesian Network can be defined in the same way as a classical Bayesian Network with the difference that real probability numbers are replaced by quantum probability amplitudes (Tucci, 1995). **Figure 1** shows an example of the proposed Quantum-Like Bayesian Network, containing quantum probability amplitudes, $\psi_{i,j}$, instead of real probability values.

In this sense, the quantum counterpart of the full joint probability distribution corresponds to the application of Born's rule to Equation 4. This results in the quantum like version of the full joint probability distribution:

$$Pr(X_1, \ldots, X_n) = \left| \prod_{i=1}^{n} \psi(X_i | Parents(X_i)) \right|^2 \quad (24)$$

FIGURE 1 | Example of a Quantum-Like Bayesian Network. The terms $\psi_{i,j}$ correspond to quantum probability amplitudes. The variables X_1, X_2, and X_3 correspond to random variables.

In order to perform exact inference in Bayesian Networks, the probability amplitude of each assignment of the network is propagated and influences the probabilities of the remaining nodes. That is, *every* assignment of *every* node of the Bayesian Network propagate throughout the network until they reach the node representing the query variable. Note that, by taking multiple assignments and paths at the same time, these trails influence each other producing interference effects.

The quantum counterpart of the Bayesian exact inference formula corresponds to the application of Born's rule to the classical marginal probability distribution equation (Equation 5).

$$Pr(X|e) = \alpha \left| \sum_y \prod_{x=1}^{N} \psi(X_x | Parents(X_x), e, y) \right|^2 \quad (25)$$

Expanding Equation 25, it will lead to the quantum marginalization formula with interference effects (Moreira and Wichert, 2014):

$$Pr(X|e) = \alpha \sum_{i=1}^{|Y|} \left| \prod_{x}^{N} \psi(X_x | Parents(X_x), e, y = i) \right|^2 + 2 \cdot Interference \quad (26)$$

$$Interference = \sum_{i=1}^{|Y|-1} \sum_{j=i+1}^{|Y|} \left| \prod_{x}^{N} \psi(X_x | Parents(X_x), e, y = i) \right| \cdot$$
$$\left| \prod_{x}^{N} \psi(X_x | Parents(X_x), e, y = j) \right| \cdot \cos(\theta_i - \theta_j)$$

In the Quantum Dynamical Model, since it uses unitary operators, the double symmetric property of these operators does not require the normalization of the computed values. In the proposed approach, on the other hand, since we do not have the constraints of double stochastic operators, we need to normalize the final scores that are computed in order to achieve a probability value. In classical Bayesian inference, normalization of the inference scores is also necessary due to assumptions made in Bayes rule. The normalization factor corresponds to α in Equation 26.

Note that, in Equation 26, if one sets $(\theta_i - \theta_j)$ to $\pi/2$, then $\cos(\theta_i - \theta_j) = 0$, which means that the quantum Bayesian

TABLE 6 | Table representation of a quantum full joint probability distribution.

X_1	...	X_N	$\psi\,(X_1, ..., X_N)$
T	...	T	$\psi_1 \cdot e^{j\theta_1}$
T	...	F	$\psi_2 \cdot e^{j\theta_2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
F	...	F	$\psi_M \cdot e^{j\theta_M}$

Network collapses to its classical counterpart. That is, the proposed Quantum-Like Bayesian Network can behave in a classical way, if one sets the interference term to zero. Setting the angles to right angles means that all cosine similarities are 0 or 1. This transforms a continuous-valued system to a Boolean-valued system. Moreover, in Equation 26, if the Bayesian Network has N binary random variables, we will end up with 2^N free quantum θ parameters.

The proposed Bayesian Network leaves an open research question regarding the quantum θ parameters: *how can one compute such parameters in order to obtain realistic inferences?* By realistic, we mean the probability that an event that was observed in an experiment. These probabilities are impossible to compute using exact Bayesian inference in experiments where the Sure Thing Principle is being violated. In the next section, we answer this question by proposing a similarity heuristic that is able to compute the quantum θ parameters through vector similarities between beliefs/actions in superposition.

6.1. Representation of Beliefs/Actions

The superposition quantum vector, comprising all possible events, is given by the quantum full joint probability distribution already presented in Equation 24. The full joint probability distribution can be illustrated in table form just like it is presented in **Table 6**.

The quantum probability inference formula is composed of two parts: one representing the classical probability and the other representing the quantum interference term. The interference term performs a summation over several combinations of the entries of the full joint probability distribution in groups of two variables: $\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} |\psi_i|\,|\psi_j| \cos\left(\theta_i - \theta_j\right)$. For each pair of variables, we will represent them as a 2-dimensional vector: one component represents the probability of ψ_i and the other corresponds to ψ_j. Moreover, the different probabilities represented in the full joint probability distribution table can be seen as the different beliefs/actions that one might have available before making a decision.

$$\mathbf{a}(X = T) = \begin{bmatrix} \left|\psi_i \cdot e^{i\theta_i}\right|^2 \\ \left|\psi_i \cdot e^{i\theta_j}\right|^2 \end{bmatrix} \quad \mathbf{b}(X = F) = \begin{bmatrix} \left|\psi_i \cdot e^{i\theta_i}\right|^2 \\ \left|\psi_j \cdot e^{i\theta_j}\right|^2 \end{bmatrix} \quad (27)$$

We always have two vectors, because the proposed Quantum-Like Bayesian network only supports binary random variables, that is, the query that it is performed to the network corresponds to a *yes* or *no* answer. In other words, one vector corresponds to
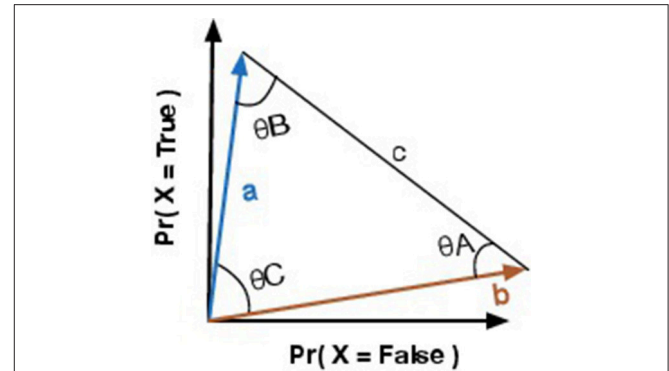

FIGURE 2 | Vector representation of two events representing a certain state.

the probability of the query random variable returning a positive answer, and the other corresponds to the probability of the query random variable returning a negative one. In a geometric space, these vectors are represented as in **Figure 2**. From these two vectors, similarity measures like the angles between the vectors or the distances between them can be computed. These similarity measures will be addressed in more detail in Section 6.2.
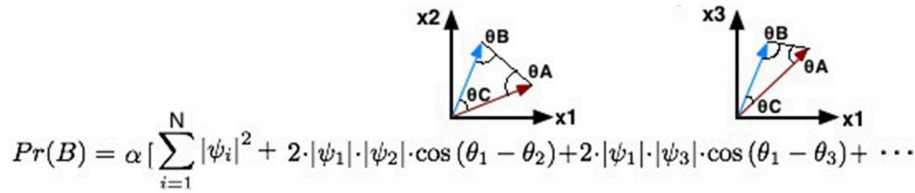
One could ask why these feature vectors are represented by probabilities. In our model, the goal is to find a quantum parameter that can be used to compute quantum probability inferences. The only information that one has are the probability distributions of a given scenario, which are encoded in the Bayesian Network.

In quantum mechanics, quantum states are always represented by unit length vectors. Since the proposed model is inspired by quantum formalisms, one might be wondering why the vectors are not unit length as well. There are two reasons for this choice. First, this representation of beliefs/actions as probabilities in feature vectors is not new, and it is a common practice in the literature (Osherson, 1995). Second, since our model is represented by a Bayesian Network and the vectors extracted directly from the network (through the representation of the full joint probability distribution), we do not need to have unit length vectors. Instead, this normalization will be performed during the inference process through the computation of the normalization factor α.

In the end, the quantum interference term is computed by computing different vector representations for each pair of variables that are being computed (**Figure 3**). These vectors are extremely important to compute, since they will enable the calculation of different quantum θ parameters.

6.2. Acquisition of Additional Information

It is important to note that, over the current literature, quantum parameters must be assigned manually in order to obtain a prediction. So, for different experiments, we will have disparate quantum parameters. For this reason, it is very hard to create a universal heuristic that can assign quantum parameters for different applications. In this work, we propose a heuristic that

$$Pr(B) = \alpha \left[\sum_{i=1}^{N} |\psi_i|^2 + 2 \cdot |\psi_1| \cdot |\psi_2| \cdot \cos(\theta_1 - \theta_2) + 2 \cdot |\psi_1| \cdot |\psi_3| \cdot \cos(\theta_1 - \theta_3) + \cdots \right]$$

FIGURE 3 | Illustration of the different 2-dimensional vectors that will be generated for each step of iteration during the computation of the quantum interference term.

is able to perform accurate predictions for the several different experiments reported in the literature related to the Prisoner's Dilemma Game and the Two Stage Gambling Game.

The goal of this similarity heuristic is to determine an angle between the vectors **a** and **b** (Equation 27) that can be used as the θ parameter in Equation 26. Moreover, by computing the Euclidean distance between vectors **a** and **b**, one can obtain vector **c**. Equation 28 shows how to obtain the norm of vector **c** through vectors **a** and **b** (**Figure 2**). Additional information is gained by comparing the similarity between the two vectors. This new information allows one to infer hidden properties of a participant's beliefs/actions from visible ones. This vector representation is similar to the approach proposed in the work of Pothos et al. (2013), where the authors represent a person's beliefs/actions in an n-dimensional vector space and the similarity between the vectors is measured by a projection operator, which corresponds to the computation of the squared length of the projected vector. This is similar to our approach, since we are also computing the length between the vectors **a** and **b**.

$$||\mathbf{c}|| =$$
$$||\mathbf{a} - \mathbf{b}|| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \tag{28}$$

Since we are interested in the angles that these vectors make between each other, we used trigonometric laws, such as the law of cosines, to determine these angles. The law of cosines is given by Equations 29–31, where θA corresponds to the angle between vectors **b** and **c**. θB corresponds to the angle between vectors **a** and **c**. And θC corresponds to the angle between vectors **a** and **b**. Since we know the coordinates of vectors **a** and **b**, one can also compute angle θC through the similarity between two vectors using the cosine similarity measure: $\cos(\theta C) = \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{a}|| \cdot ||\mathbf{b}||}$. However, since we only know the length of vector **c**, we need to compare the similarity of the vectors through the law of cosines.

$$||\mathbf{a}||^2 = ||\mathbf{b}||^2 + ||\mathbf{c}||^2 - 2 \cdot ||\mathbf{b}|| \cdot ||\mathbf{c}|| \cdot \cos(\theta A) \Leftrightarrow \theta A$$
$$= \cos^{-1}\left(\frac{||\mathbf{b}||^2 + ||\mathbf{c}||^2 - ||\mathbf{a}||^2}{2 \cdot ||\mathbf{b}|| \cdot ||\mathbf{c}||}\right) \tag{29}$$

$$||\mathbf{b}||^2 = ||\mathbf{a}||^2 + ||\mathbf{c}||^2 - 2 \cdot ||\mathbf{a}|| \cdot ||\mathbf{c}|| \cdot \cos(\theta B) \Leftrightarrow \theta B$$
$$= \cos^{-1}\left(\frac{||\mathbf{a}||^2 + ||\mathbf{c}||^2 - ||\mathbf{b}||^2}{2 \cdot ||\mathbf{a}|| \cdot ||\mathbf{c}||}\right) \tag{30}$$

$$||\mathbf{c}||^2 = ||\mathbf{a}||^2 + ||\mathbf{b}||^2 - 2 \cdot ||\mathbf{a}|| \cdot ||\mathbf{b}|| \cdot \cos(\theta C) \Leftrightarrow \theta C$$
$$= \cos^{-1}\left(\frac{||\mathbf{a}||^2 + ||\mathbf{b}||^2 - ||\mathbf{c}||^2}{2 \cdot ||\mathbf{a}|| \cdot ||\mathbf{b}||}\right) \tag{31}$$

6.3. Definition of the Similarity Heuristic

Violations to the Sure Thing principle imply a decrease in the final probability values when compared to the classical theory. This suggests that, somehow, we need to force the quantum parameters to have a destructive interference effect. This can be obtained by setting the quantum parameter to π (which is the angle that provides the smallest cosine value). The additional information that we incorporated in **Figure 2**, namely the Euclidean distance between vectors and their similarities, is translated into a triangle. This shape has a well-known property that all their inner angles must sum to $180°$ or π radians. Moreover, we would like to have a destructive interference effect that takes into account the similarity of the original vectors. Equation 32, shows how one can obtain this relationship.

$$\theta_A + \theta_B + \theta_C = \pi \Leftrightarrow \pi - \theta_C = \theta_A + \theta_B \tag{32}$$

$$\Leftrightarrow \pi - \frac{\theta_C}{2} = \frac{\theta_A + \theta_B + \pi}{2}$$

When, the similarity of the vectors is very small, that is θ_C is very small, then we can add a third relationship:

$$\theta_A + \theta_B + \theta_C = \pi \Leftrightarrow \pi = \theta_A + \theta_B$$

In this sense, we can formulate the general formula of the proposed similarity heuristic :

$$h(a, b) = \begin{cases} \pi & \text{if } \phi < 0 \\ \pi - \theta_C/2 & \text{if } \phi > 0.2 \\ \pi - \theta_C & \text{otherwise} \end{cases} \tag{33}$$

We also came up with a similarity measure ϕ, which is given by the ratio between all the angles that the vectors make between them. In order words, it represents the similarity between the additional information found by manipulating the original vectors and is given by Equation 34.

$$\phi = \frac{\theta_C}{\theta_A} - \frac{\theta_B}{\theta_A} \tag{34}$$

The thresholds shown in the proposed similarity heuristic were taken by observing the data from several experiments violating

the Sure Thing Principle. These include several experiments in the literature of the Prisoner's Dilemma Game and the Two Stage Gambling Game. Yukalov and Sornette (2011) also did something similar. They analyzed the experiments violating the Sure Thing Principle and came up with a static interference term (the Interference Quarter Law) that allows them to apply their model without knowing exactly a priori the outcome of some specific experiment. The proposed model works under similar conditions. We analyzed several experiments from the literature from different games and mapped the trends of the data into a dynamic heuristic. So, in the end, the proposed model works under some rules that enables a dynamic behavior (after all each experiment is unique, so there should be the freedom of different quantum interferences) and also enables the application of the model without specific a priori knowledge from a specific experiment.

In quantum mechanics, the θ parameter corresponds to the phase of a wave. When representing a quantum state in a Hilbert space, this phase is given by the inner product between two quantum states (Busemeyer and Bruza, 2012). The proposed similarity heuristic is motivated by the same idea. For two vectors representing a person's belief/action, we find which angle (or in this case, a combination of angles) that can lead to the observed probabilities for the Prisoner's Dilemma and for the Two Stage Gambling game.

6.4. Summary of the Proposed Model

The proposed model is built based on observed data to perform quantum probabilistic inferences. We are using a similarity heuristic, which relies in the data of the Bayesian Network to indicate the parameters that will allow us to perform quantum probabilistic inferences. One should keep in mind that this function is a heuristic: it generally provides good results in many situations (in this case, the Two Stage Gambling game, and the Prisoner's Dilemma), but at the cost of occasionally not giving us very accurate results (Shah and Oppenheimer, 2008).

In sum, the proposed model works as follows:

- Definition of a quantum-like Bayesian Network containing cause/effect relationships of a given scenario. Each node of the Bayesian Network corresponds to a binary random variable and is associated to a conditional probability table. These tables represent conditional probability distributions, which can be converted to quantum amplitudes through Born's rule.
- When performing a query to the quantum-like Bayesian Network, a set of quantum parameters will emerge, because of the application of Equation 26. These parameters can be determined with the similarity heuristic that takes into account similarities between vectors.
- The proposed similarity heuristic takes into account two 2-dimensional vectors. Each vector corresponds to one assignment of the query variable (for instance, the probability of the query being true or the probability of the query being false).
- The two features of each vector correspond to each entry of the full joint probability distribution of the Bayesian Network that has the same assignment of the query variable. For instance,

all entries of the distribution that have the assignment of the query variable set to *true*.
- After knowing the similarities that the vectors share between them, we can apply the proposed similarity heuristic given in Equation 33 to obtain a θ parameter that enables the computation of the final probability value of the query.

One might be thinking that we use two of the three data points directly in the model (known *Defect* and known *Collaborate*). Then, they use one free parameter to account for the remaining data point (the probability of Defection in the unknown condition). However, this is not what we state with this work. As already mentioned, this work is a nonparametric method for estimating inference effects from a statistical point of view. It is a statistical model that is simpler than the previous quantum dynamic and quantum-like models proposed in the literature. Again, this work is not about simulation methods of fitting. We are simply providing a Bayesian Network structure that enables a simple representation of more complex decision-making scenarios, and the incorporation of a similarity heuristic (which results from algebraic manipulations) in order to assign values to quantum parameters in such a way that provides accurate predictions (that is, it can represent the data accurately).

In the next sections, we will present a full example of how the proposed Quantum-Like Bayesian Network can be applied (Section 6.5). We will also present experimental results of the proposed model applied to several works of the literature concerned with the Prisoner's Dilemma game (Section 7.1) and the Two Stage Gambling game (Section 7.2).

6.5. Example of Application of the Proposed Model

In this section, we will demonstrate how the proposed Bayesian Network can be applied to the average results presented in **Table 1** for the Prisoner's Dilemma game. The proposed Quantum-Like Bayesian Network can be summarized in the following steps:

- **Step 1: Create a Bayesian Network Representation of the Problem:** In the Prisoner's dilemma game, if nothing is told to the participants, then there is a 50% chance of the first participant choosing to *Defect* or *Cooperate*. The decision of the first participant is then followed by the decision of the second participant. A Bayesian Network representation of this problem is illustrated in **Figure 4**.
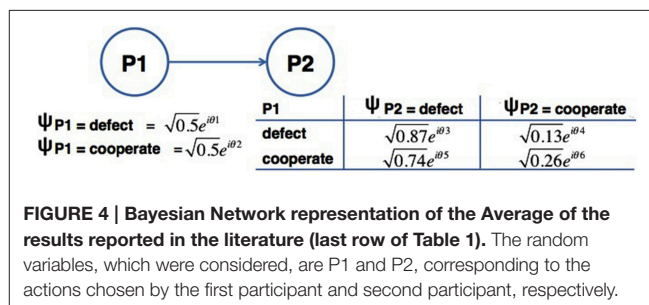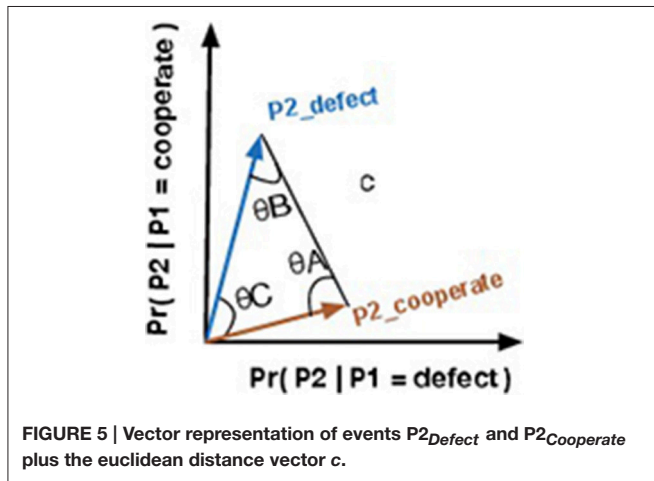

P1	$\Psi_{P2 = defect}$	$\Psi_{P2 = cooperate}$
defect	$\sqrt{0.87}e^{i\theta 3}$	$\sqrt{0.13}e^{i\theta 4}$
cooperate	$\sqrt{0.74}e^{i\theta 5}$	$\sqrt{0.26}e^{i\theta 6}$

$\Psi_{P1 = defect} = \sqrt{0.5}e^{i\theta 1}$
$\Psi_{P1 = cooperate} = \sqrt{0.5}e^{i\theta 2}$

FIGURE 4 | Bayesian Network representation of the Average of the results reported in the literature (last row of Table 1). The random variables, which were considered, are P1 and P2, corresponding to the actions chosen by the first participant and second participant, respectively.

TABLE 7 | Quantum full joint probability distribution representation of the Bayesian Network in Figure 4.

P1	P2	Pr(P1, P2)
Defect	Defect	$\sqrt{0.5} \cdot e^{j \cdot \theta_1} \times \sqrt{0.87} \cdot e^{j \cdot \theta_3} = 0.6595 \cdot e^{j \cdot (\theta_1 + \theta_3)}$
		$= 0.6595 \cdot e^{j \cdot \theta_A}$
Defect	Cooperate	$\sqrt{0.5} \cdot e^{j \cdot \theta_1} \times \sqrt{0.13} \cdot e^{j \cdot \theta_4} = 0.2550 \cdot e^{j \cdot (\theta_1 + \theta_4)}$
		$= 0.2550 \cdot e^{j \cdot \theta_B}$
Cooperate	Defect	$\sqrt{0.5} \cdot e^{j \cdot \theta_2} \times \sqrt{0.74} \cdot e^{j \cdot \theta_5} = 0.6083 \cdot e^{j \cdot (\theta_2 + \theta_5)}$
		$= 0.6083 \cdot e^{j \cdot \theta_C}$
Cooperate	Cooperate	$\sqrt{0.5} \cdot e^{j \cdot \theta_2} \times \sqrt{0.26} \cdot e^{j \cdot \theta_6} = 0.3606 \cdot e^{j \cdot (\theta_2 + \theta_6)}$
		$= 0.3606 \cdot e^{j \cdot \theta_D}$

FIGURE 5 | Vector representation of events $P2_{Defect}$ and $P2_{Cooperate}$ plus the euclidean distance vector c.

- **Step 2: Compute the Vectors associated to each action.** Since we want to determine the $Pr(P2 = Defect)$, this probability will be given by the quantum full joint probability distribution, which is represented in **Table 7**.

$$\mathbf{P2}_{Defect} = \begin{bmatrix} |0.6595 \cdot e^{i \cdot \theta_A}|^2 \\ |0.6083 \cdot e^{i \cdot \theta_C}|^2 \end{bmatrix} = \begin{bmatrix} 0.435 \\ 0.370 \end{bmatrix}$$

$$\mathbf{P2}_{Cooperate} = \begin{bmatrix} |0.2550 \cdot e^{i \cdot \theta_B}|^2 \\ |0.3606 \cdot e^{i \cdot \theta_D}|^2 \end{bmatrix} = \begin{bmatrix} 0.065 \\ 0.130 \end{bmatrix} \quad (35)$$

This way, one can build feature vectors using classical probabilities. For instance, the probability of $Pr(P2 = Defect)$ is given by a 2-dimensional feature vector with entries: $Pr(P1 = Defect) \cdot Pr(P2 = Defect|P1 = Defect)$ and $Pr(P1 = Cooperate) \cdot Pr(P2 = Defect|P1 = Cooperate)$. The feature vector corresponding to the action $Cooperate$ can be achieved in the same way (Equation 35).

- **Step 3: Determine the quantum parameters using the proposed similarity heuristic:** Since we only have two random variables, we only need to compute one θ parameter. This parameter can be obtained by directly by first computing the Euclidean distance between $\mathbf{P2}_{Defect}$ and $\mathbf{P2}_{Cooperate}$, and by computing the inner angles of the resulting triangle (**Figure 5**).

$$||\mathbf{c}|| = ||\mathbf{P2}_{Defect} - \mathbf{P2}_{Cooperate}||$$
$$= \sqrt{(0.435 - 0.065)^2 + (0.37 - 0.13)^2} = 0.4410 \ (36)$$

The norm of vectors $P2_{Defect}$ and $P2_{Cooperate}$ is given by:

$$||\mathbf{P2}_{Defect}|| = \sqrt{0.435^2 + 0.370^2} = 0.5711$$
$$||\mathbf{P2}_{Cooperate}|| = \sqrt{0.065^2 + 0.130^2} = 0.1453 \quad (37)$$

The inner angles of the triangle formed by vectors $\mathbf{P2}_{Defect}$ and $\mathbf{P2}_{Cooperate}$ and $\mathbf{c}$ can be computed from the law of Cosines presented in Equations 38–40.

$$A = \cos^{-1}\left(\frac{||\mathbf{P2}_{Cooperate}||^2 - ||\mathbf{P2}_{Defect}||^2 + c^2}{2 \cdot \mathbf{c} \cdot ||\mathbf{P2}_{Cooperate}||} \right) = 2.6102$$
$$(38)$$

$$B = \cos^{-1}\left(\frac{||\mathbf{P2}_{Defect}||^2 - ||\mathbf{P2}_{Cooperate}||^2 + c^2}{2 \cdot \mathbf{c} \cdot ||\mathbf{P2}_{Defect}||} \right) = 0.1294$$
$$(39)$$

$$C = \cos^{-1}\left(\frac{||\mathbf{P2}_{Defect}||^2 + ||\mathbf{P2}_{Cooperate}||^2 - c^2}{2 \cdot ||\mathbf{P2}_{Defect}|| \cdot ||\mathbf{P2}_{Cooperate}||} \right) = 0.4023$$
$$(40)$$

Given that $\frac{\theta C}{\theta A} - \frac{\theta B}{\theta A} = 0.1046$, then the final quantum θ parameter can be computed by using the third condition of Equation 33

$$\theta = \pi - \theta C = \pi - 0.4023 = 2.7393 \quad (41)$$

- **Step 4: Perform the Probabilistic Inference.** In order to compute $Pr(P2 = Defect)$ we also need to compute the opposite probability, that is, $Pr(P2 = Cooperate)$. Equation 42 represents quantum amplitudes through the symbol ψ. The sub indexes D and C correspond to the actions $Defect$ and $Cooperate$, respectively.

$$Pr(P2 = Defect) = \alpha \left[|\psi_{P2=D|P1=D}|^2 + |\psi_{P2=D|P1=D}|^2 \right.$$
$$\left. + 2 \cdot |\psi_{P2=D|P1=D}| \cdot |\psi_{P2=D|P1=C}| \cdot \cos(\theta) \right] \ (42)$$

$$Pr(P2 = Defect) = \alpha [0.5 \times 0.87 + 0.5 \times 0.74$$
$$+ 2 \times \sqrt{0.5 \times 0.87} \times \sqrt{0.5 \times 0.74} \cos(2.7393)]$$
$$(43)$$

Computing the probability of $Pr(P2 = Cooperate)$ in the same way, we obtain:

$$Pr(P2 = Defect) = \alpha \cdot 0.0667$$
$$Pr(Cooperate) = \alpha \cdot 0.0258 \quad (44)$$

- Step 5: Compute Normalization Factor and Final Probabilities.

$$\alpha = \frac{1}{0.0667 + 0.0258} = \frac{1}{0.0925} = 10.8108 \quad (45)$$

The final probabilities are given by Equation 45. Note that in **Table 1**, the observed probability of a player choosing to *Defect* was 0.64. The proposed Bayesian Network estimated this probability to be approximately 0.72, which corresponds to a fit error percentage of 12.63%.

$$Pr(P2 = Defect) = 0.7208 \quad Pr(P2 = Cooperate) = 0.2792 \tag{46}$$

7. EXPERIMENTAL RESULTS

Violations to the Sure Thing Principle are hard to verify in complex decision-making problems. For this reason, there is not much data available in the literature for validation purposes. So, in this work, we will validate our model for several different experiments made to detect violations of the Sure Thing Principle in the Prisoner's Dilemma Game (Section 7.1) and for the Two Stage Gambling game (Section 7.2).

7.1. Quantum Bayesian Network Applied to the Prisoner's Dilemma Game

In this section, we apply our model to predict the results obtained for the Prisoner's Dilemma game for several works in the literature.

It is common (and good) practice in cognitive science to compare the results of one's model to the results of leading comparable models. The fit error percentages that we present in the following sections would be much easier to interpret if there could be other models to compare with. However, we cannot perform this comparison directly, because the current models of the literature only work for isolated experiments, just like it was shown for the Quantum Dynamical Model (Section 4.1) and the Quantum-Like Approach (Section 4.2). That is, each time there is a new experiment, the parameters of their respective models would need to be tuned manually in order to perform correct predictions. We propose a general and scalable framework that is able to perform predictions in several different setting with small amounts of fit errors.

In this sense, we modeled each result reported in **Table 1** with the proposed Bayesian Network and using the proposed similarity heuristic. We obtained the results that are presented in **Figure 6**.

For a more detailed analysis of **Figure 6**, **Table 8** shows the quantum θ parameters that were computed for each experiment and the quantum parameter that would be expected to achieve a 0% fit error. The fit error is a percentage value and was computing in the following way: $(1 - \frac{computed\_probability}{observed\_probability}) * 100$. In **Table 8**, the term *computedprobability* corresponds to the column *Pr(Defect) predicted* and the term *observed_probability* corresponds to the column *observed_probability*.

In **Table 8**, one can see that the proposed similarity heuristic was able to perform good approximations to the data. The dynamical heuristic enabled to perform different estimations of quantum interference effects for different decision problems. However, since it is an heuristic, it can sometimes lead to overestimations, which was the case in the work of Busemeyer et al. (2006a). These overestimations occur due to the sensitivity of the quantum parameters. That is, a small change in a quantum parameter will lead to a completely different probability value. This will be discussed in more depth in Section 7.1.2.

As one might have noticed, the work of Croson (1999) was not taken into account in the analysis of these results. We decided to analyse these results in the next section, because they contained properties that were different from the remaining works. In Croson (1999), the participants were never told about the actions of the other player. The author asked for the participants to first try to guess what action the other player chose and then make a decision. In another setting, participants were just asked to make a decision.

7.1.1. The Special Case of Croson's (1999) Experiments

In work of Croson (1999), we used the results reported for the first two payoff matrices tested in their work and performed the average of the results. When trying to compute the optimum quantum θ parameter that would lead to the computation of the probability with a 0% fit error, we could not find any. There was no possible parameter that could be obtained from the two feature vectors representing the probability of choosing either a *Defect* action or a *Cooperate* action.
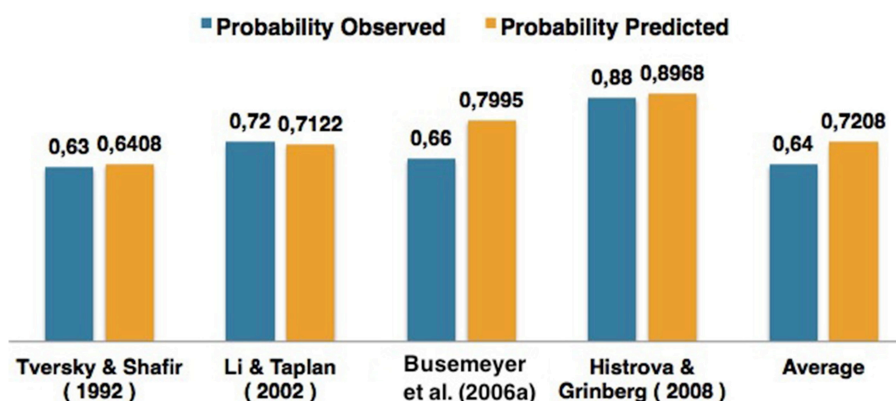

FIGURE 6 | Comparison of the results obtained for different works of the literature concerned with the Prisoner's Dilemma game.

TABLE 8 | Analysis of the quantum θ parameters computed for each work of the literature using the proposed similarity function.

Literature	Expected θ	Computed θ	Pr(Defect)	Pr(Defect) predicted	Fit error
Shafir and Tversky, 1992	2.8151	2.8057	0.6300	0.6408	1.71
Li and Taplin, 2002[b]	3.3033	3.0121	0.7200	0.7122	1.09
Busemeyer et al., 2006a	2.9738	3.3628	0.6600	0.7995	21.13
Hristova and Grinberg, 2008	2.8255	2.7400	0.8800	0.8968	3.01
Average	2.8718	2.7393	0.6400	0.7208	12.63

Expected θ corresponds to the quantum parameter that leads to the observed probability value in the experiment. Computed θ corresponds to the quantum parameter computed with the proposed heuristic.
b corresponds to the average of all seven experiments reported.

TABLE 9 | Results for the two games reported in the work of Croson (1999) for the Prisoner's Dilemma Game for several conditions: when the action of the second player was guessed to be *Defect* (Guessed to Defect), when the action of the second player was guessed to be *Cooperate* (Guessed to Collaborate), and when the action of the second player was not known (Unknown).

Croson, 1999	Guessed to Defect	Guessed to Cooperate	Unknown	Unknown predicted	Violation of STP
Game 1	0.1700	0.6800	0.2250	0.5877	No
Game 2	0.4700	0.6500	0.3750	0.4390	Yes
Average	0.6700	0.32	0.3000	0.5053	Yes
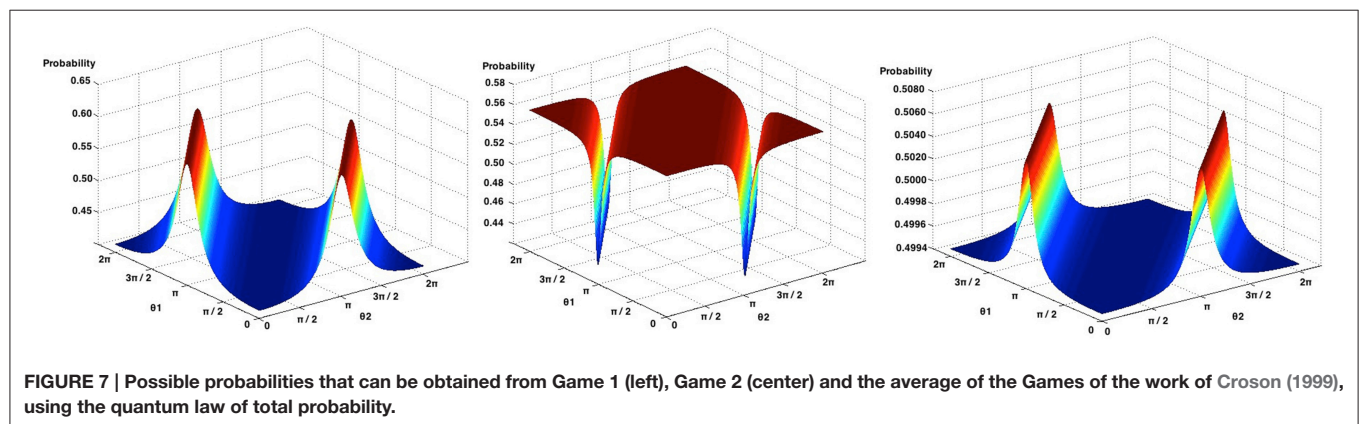
FIGURE 7 | Possible probabilities that can be obtained from Game 1 (left), Game 2 (center) and the average of the Games of the work of Croson (1999), using the quantum law of total probability.

As a first thought, we noticed that the average of the results could be the cause of such impossibility, because they were not the true probabilities of the events reported. So, we decided to analyse the outcome of each experiment of the work of Croson (1999) individually. **Table 9** specifies those results.

We again analyzed the individual results of **Table 9**, and again, we could not find any quantum θ parameter that would lead to the computation of probabilities with a 0%. On the contrary, the minimum fit errors found were 64.89, 83.25, and 17.06% for Game 1, Game 2 and the Average of these games, respectively. **Figure 7** present all possible probabilities that can be computed using the quantum law of total amplitude.

Analysing Game 1 (**Figure 7**, left), the probability that leads to the smallest fit error is obtained when both θ parameters are set to zero, with a probability of 0.4123. The observed probability reported in this experiment corresponds to 0.2250, leading to a computed fit error of 64.69%.

For Game 2 (**Figure 7**, center), when $\theta_1 = 0$ and $\theta_2 = \pi$, we obtain the probability that leads to the smallest fit error, which is 0.4390, with a fit error of 83.25 %.

When computing the average of both games (**Figure 7**, right), the quantum θ parameters found were $\theta_1 = 0$ and $\theta_2 = 0$. This leads to a probability of 0.4947, corresponding to a fit error of 17.06%.

7.1.2. Analysing Li and Taplin (2002) Experiments

Table 10 specifies the results collected by Li and Taplin (2002), which corresponded to the average of the results obtained in seven different experiments for the Prisoner's Dilemma game. In this section we analyse each of these seven experiments, by trying to predict their outcome using the proposed Bayesian Network.

The results reported in the experiments conducted by Li and Taplin (2002) are presented in **Table 10**. Note that Games 3, 6 and 7 are not violating the Sure Thing Principle, because: $Pr(Defect) \geq Pr(Unknown) \leq Pr(Cooperate)$ or $Pr(Cooperate) \geq Pr(Unknown) \leq Pr(Defect)$. Additionally, the results reported for the unknown condition in Games 3, 6 and 7 are very close to the classical probability theory. The goal of the study performed by Li and Taplin was to question

TABLE 10 | Experimental results reported in work of Li and Taplin (2002) for the Prisoner's Dilemma game for several conditions: when the action of the second player is known to be *Defect* (Known to Defect), when the action of the second player is known to be *Cooperate* (Known to Collaborate), and when the action of the second player was not known (Unknown).

Li and Taplin, 2002	Known Defect	Known Cooperate	Unknown	Classical probability	Violation of STP
Game 1	0.7333	0.6670	0.6000	0.7000	Yes
Game 2	0.8000	0.7667	0.6300	0.7833	Yes
Game 3	**0.9000**	**0.8667**	**0.8667**	**0.8834**	**No**
Game 4	0.8333	0.8000	0.7000	0.8167	Yes
Game 5	0.8333	0.7333	0.7000	0.7833	Yes
Game 6	**0.7667**	**0.8333**	**0.8000**	**0.8000**	**No**
Game 7	**0.8667**	**0.7333**	**0.7667**	**0.8000**	**No**
Average	0.8200	0.7700	0.7200	0.7950	Yes

The column Violations of STP corresponds to determining if the collected results are violating the Sure Thing Principle. The values in bold represent the experiments that are not violating the Sure Thing Principle.
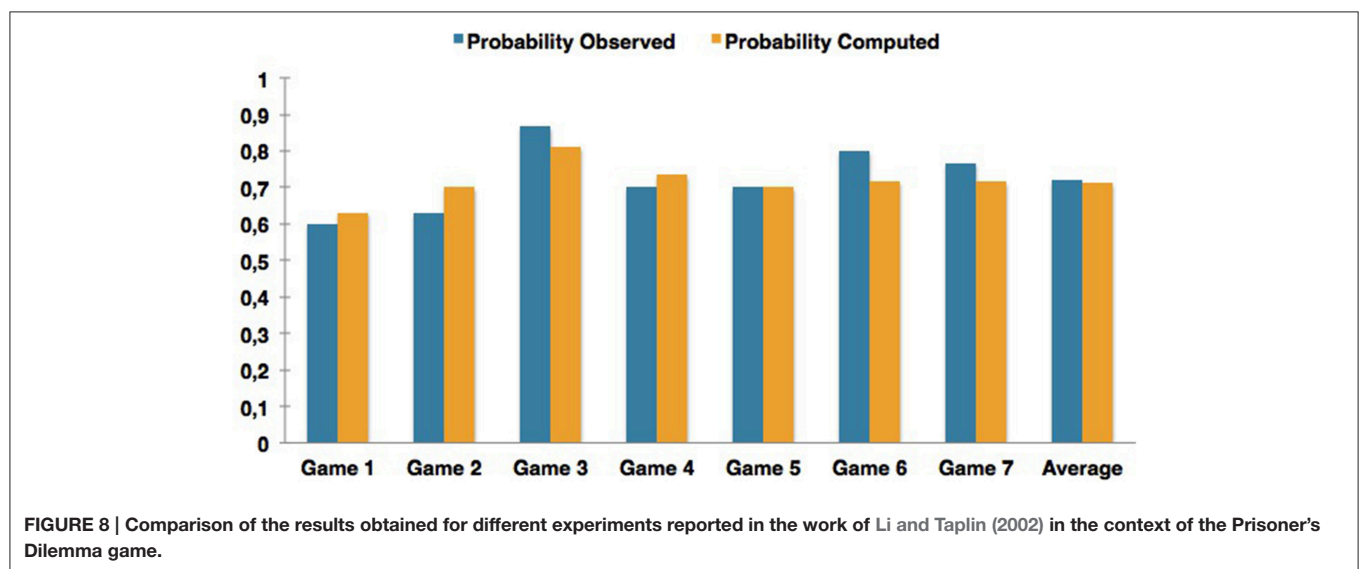
FIGURE 8 | Comparison of the results obtained for different experiments reported in the work of Li and Taplin (2002) in the context of the Prisoner's Dilemma game.

if there was really violations of the Sure Thing Principle under the Prisoner's Dilemma game. According to **Table 10** three of the seven experiments did not show a violation, and reported results very similar to the classical probability theory.

By applying the proposed quantum-like Bayesian Network each game in **Table 10**, we obtained the results illustrated in **Figure 8**.

The experiments that achieved the highest fit error rates correspond to Games 2 and 6. Game 6 corresponds to a situation where the Sure Thing Principle was not being violated. This leads to the conclusion that the proposed Bayesian Network can also predict classical probabilities, but with some fit errors.

Table 11 shows the quantum parameters that were computed and compares them with the parameters that would be expected in order to obtain the smallest fit error percentage. One thing worth mentioning in the computation of these quantum parameters is their sensitivity. Consider the row of **Table 11** addressing the results of Game 2. The difference between expected quantum parameter with the one that was computed using the similarity heuristic corresponds to a difference of just 0.0322. However, this small difference introduced a fit error of almost 11.28% in the computation of the final probabilities.

Figure 9 illustrates the relation between the quantum θ parameter and the final probabilities that can be obtained in Li's Game 2, Game 6 and the work of Busemeyer et al. (2006a).

Small changes in the θ parameters can lead to a completely different probability outcomes. This has some relation with *deterministic chaos*, in which small differences in initial conditions yield widely diverging outcomes in a system. This chaos suggests how difficult the task of predicting human decisions is and how random it can be (Sterman, 1989).

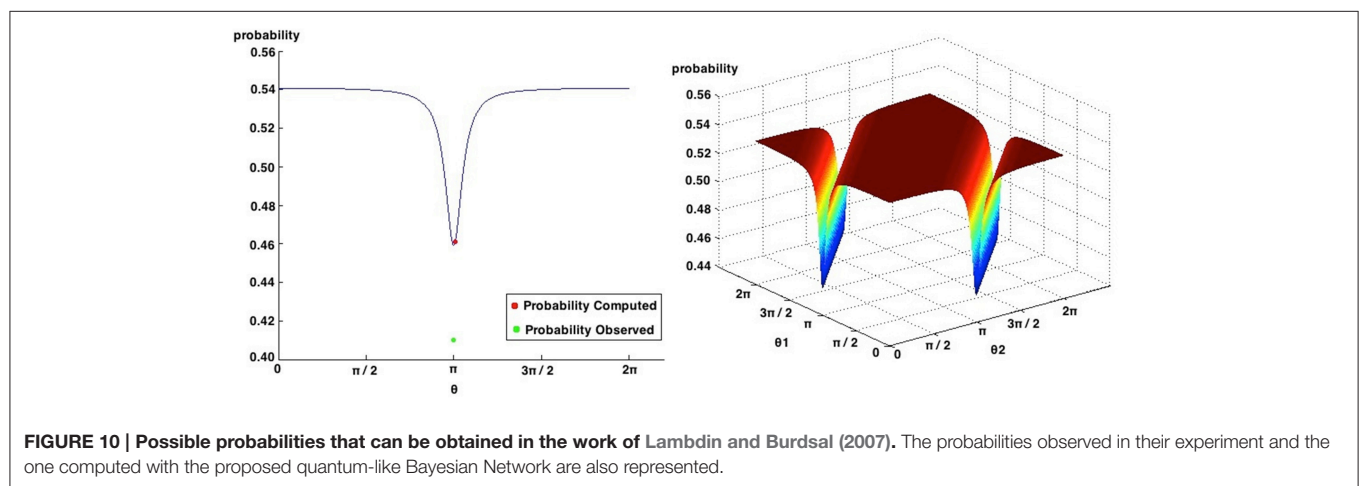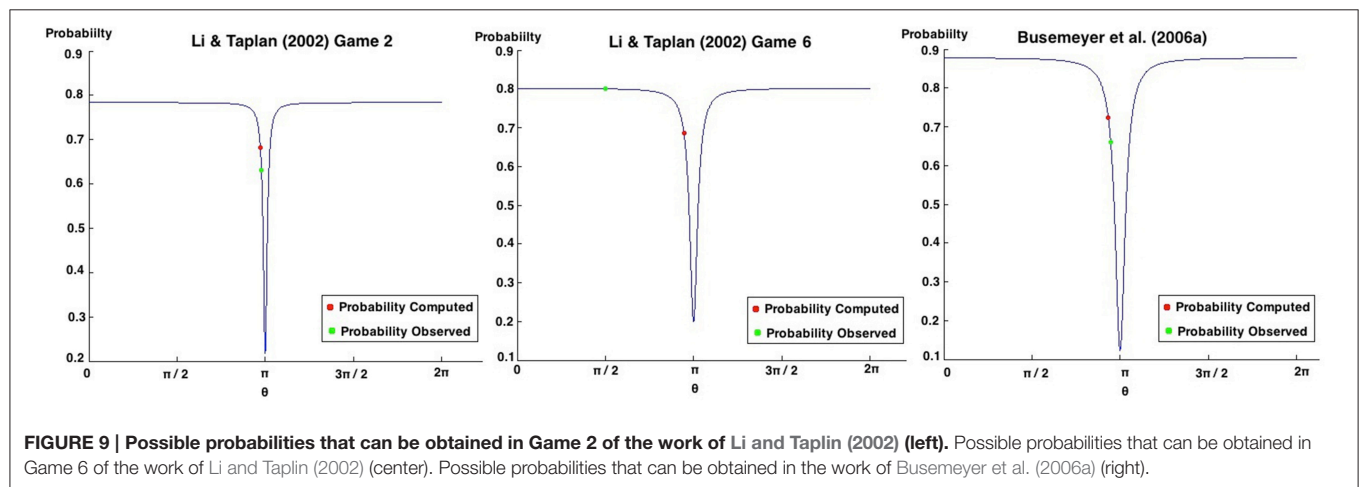## 7.2. Quantum Bayesian Network Applied to the Two Stage Gambling Game

For the Two Stage Gambling Game, the overall results reported very small fit errors. The highest fit error percentage achieved was 16.3% and corresponds to the work of Kuhberger et al. (2001). Once again, the work of Kuhberger et al. (2001) is not showing a violation to the Sure Thing Principle, enhancing the previous conclusion that the proposed quantum-like Bayesian Network works best in situation where this violation exists.

In what concerns the work of Lambdin and Burdsal (2007) the proposed Quantum-Like Bayesian Network could not make accurate predictions. **Figure 10** show all possible probabilities

Li and Taplin, 2002	Expected θ	Computed θ	Unknown	Unknown predicted	Fit error %
Game 1	3.0170	2.9845	0.6000	0.6313	5.21
Game 2	3.0758	3.0436	0.6300	0.7011	11.28
Game 3	**2.8052**	**2.9810**	**0.8667**	**0.8113**	**6.39**
Game 4	3.2313	3.0306	0.7000	0.7341	4.87
Game 5	2.8519	2.8511	0.7000	0.7006	0.08
Game 6	**1.5708**	**2.9350**	**0.8000**	**0.7169**	**10.39**
Game 7	**3.7812**	**2.7365**	**0.7667**	**0.7159**	**6.63**
Average	3.3033	2.9888	0.7200	0.7122	1.09

The entries highlighted correspond to games that are not violating the Sure Thing Principle. Expected θ corresponds to the quantum parameter that leads to the observed probability value in the experiment. Computed θ corresponds to the quantum parameter computed with the proposed heuristic.

FIGURE 9 | Possible probabilities that can be obtained in Game 2 of the work of Li and Taplin (2002) (left). Possible probabilities that can be obtained in Game 6 of the work of Li and Taplin (2002) (center). Possible probabilities that can be obtained in the work of Busemeyer et al. (2006a) (right).

FIGURE 10 | Possible probabilities that can be obtained in the work of Lambdin and Burdsal (2007). The probabilities observed in their experiment and the one computed with the proposed quantum-like Bayesian Network are also represented.

that can be obtained by varying the quantum parameters. As one can see, the minimum value that we can obtain corresponds to 0.4593. However, the observed probability reported by Lambdin and Burdsal (2007) corresponds to 0.41. This leads to a fit error of 12.02%.

In the work of Busemeyer et al. (2012), the authors applied the quantum dynamical model to reproduce the results obtained

for the Two Stage Gambling Game and also explored the use of Hierarchical Bayesian methods to estimate the values of quantum parameters to simulate the player's *personal profile*: risk aversion, loss aversion, memory and choice. In the recent work of Busemeyer et al. (2015), the authors also compare the quantum model with a classical model using Bayes factor. They concluded that the quantum approach was preferred by the Bayes Factor.

7.3. Comparison with Other Works of the Literature

In this section, we compare the results obtained with the proposed Quantum-Like Bayesian Network with the Quantum Prospect Decision Theory (Yukalov and Sornette, 2011). From all the analyzed models, this is the only one that can be called predictive due to its static heuristic: the Interference Quarter Law. The reason why we proposed a dynamic heuristic is because every decision problem is different and, consequently, quantum interference effects should also be different and not static. In the Quantum Prospect Decision Theory, the quantum interference term is fixed by the Interference Quarter Law, that is, the quantum interference term in the law of total probability is fixed to 0.25.

In the current model, since each decision problem is different, the proposed heuristic will compute a quantum θ parameter through similarities that the vector make between each other and these vectors are constructed from the experimental data. So, the vectors take into account the properties of each experiment, making it possible to compute different quantum interference terms for different decision problems.

Table 12 shows the results obtained for the Quantum Prospect Decision Theory and for the Quantum-Like Bayesian Network for the different works of the literature that tested violations to the Sure Thing Principle in the Prisoner's Dilemma Game and the Two Stage Gambling Game.

In the end, the results from **Table 12** demonstrate that, in general, the proposed Quantum-Like Bayesian Network together with the dynamic heuristic managed to fit the observed results in the several different experiments with an average fit error of 6.3%, whereas the Quantum Prospect Decision Theory achieved an average fit error of 16.51%.

One needs to take into account that in the Quantum Prospect Decision Theory and in the proposed Quantum-Like Bayesian Network, heuristics are used to estimate the quantum interference effects. This means that the heuristic can lead to a good fit of the data most of the times, but, in some cases, it can lead to completely wrong results. In the Quantum Prospect Theory, for instance, one can see the static Interference Quarter Law heuristic performed several estimations with big fit errors. The same is applied to the proposed Quantum-Like Bayesian

Network. The difference is that this last model makes use of dynamic heuristics. **Table 12** shows that the proposed dynamic heuristic overestimated the results in the works of Busemeyer et al. (2006a) and Kuhberger et al. (2001). This also happens due to the sensitivity of the θ parameters already discussed in **Figure 9**.

We also applied the Quantum Prospect Theory and the proposed Quantum-Like Bayesian Network to all experiments performed in the work of Li and Taplin (2002). **Table 13** shows again great discrepancies between the average fit error obtained with the static heuristic of the Quantum Prospect Decision Theory. In general, the proposed model manages to fit all the different seven experiments with an average fit error of 6.41%, whereas the Quantum Prospect Decision Theory achieved an error of 24.23%. Most of the times, the Interference Quarter Law managed to produce lower estimations of the results observed during the several experiments. This shows that having a dynamical heuristic that is able to adapt to the different decision problems brings advantages in terms of predictive effectiveness.

8. DISCUSSION AND CONCLUSION

In this work, we proposed an alternative quantum structure to perform quantum probabilistic inferences to accommodate the paradoxical findings of the Sure Thing Principle. We proposed a Quantum-Like Bayesian Network, which consists in replacing classical probabilities by quantum probability amplitudes. However, since this approach suffers from the problem of exponential growth of quantum parameters, we also proposed a similarity heuristic that automatically fits quantum parameters through vector similarities. This makes the proposed model general and predictive in contrast to the current state of the art models, which cannot be generalized for more complex decision scenarios and that only provide an explanatory nature for the observed paradoxes.

In Section 1.3, we established a set of research questions that we would like to address with the present research work. Their answers are detailed below.

1. Why do we need another quantum-like model to explain violations to the Sure Thing Principle?

TABLE 12 | Comparison between the Quantum Prospect Decision Theory (QPDT) model and the proposed Quantum-Like Bayesian Network (QLBN) for different works of the literature reporting violations to the Sure Thing Principle.

Literature	Pr(Defect) Observed	Pr(Defect) Computed (QPDT)	Fit error (QPDT)	Pr(Defect) Computed (QLBN)	Fit error (QLBN)
Shafir and Tversky, 1992	0.6300	0.6550	0.0397	0.6408	**0.0171**
Li and Taplin, 2002[b]	0.7200	0.5450	0.2431	0.7122	**0.0108**
Busemeyer et al., 2006a	0.6600	0.6250	**0.0531**	0.7995	0.2113
Hristova and Grinberg, 2008	0.8800	0.7000	0.2045	0.8968	**0.0191**
Tversky and Shafir, 1992	0.3700	0.3850	0.0405	0.3641	**0.0159**
Kuhberger et al., 2001	0.4800	0.3450	0.2813	0.4018	**0.1629**
Lambdin and Burdsal, 2007	0.4100	0.2900	0.2927	0.4085	**0.0037**
Average fit error	–	–	0.1651	–	**0.0630**

b corresponds to the average of all seven experiments reported. The values in bold represent the models that obtained the lowest Fit error.

TABLE 13 | Comparison between the Quantum Prospect Decision Theory (QPDT) model and the proposed Quantum-Like Bayesian Network (QLBN) for all the different experiments performed in the work of Li and Taplin (2002).

Literature	Pr(Defect) Observed	Pr(Defect) Computed (QPDT)	Fit error (QPDT)	Pr(Defect) Computed (QLBN)	Fit error (QLBN)
Game 1	0.6000	0.4502	0.2497	0.6313	**0.0522**
Game 2	0.6300	0.5333	0.1535	0.7011	**0.1129**
Game 3	0.8667	0.6334	0.2692	0.8113	**0.0639**
Game 4	0.7000	0.5667	0.1904	0.7341	**0.0487**
Game 5	0.7000	0.5333	0.2381	0.7006	**0.0009**
Game 6	0.8000	0.5500	0.3125	0.7169	**0.1039**
Game 7	0.7667	0.5500	0.2826	0.7159	**0.0663**
Average fit error	–	–	0.2423	–	**0.0641**

The values in bold represent the models that obtained the lowest Fit error.

Many of the models that have been proposed in the literature cannot be considered predictive. Most of these models require a set of quantum parameters to be fitted and, so far, the only way these models have to fit the parameters is to use the final outcome of the experiment to set the parameters in order to explain the experimental outcome. There is, however, one model in the literature that proposed a static heuristic to compute the quantum interference effects and can be called predictive. This model is the Quantum Prospect Decision Theory, proposed by Yukalov and Sornette (2011).

2. What is the advantage of the proposed approach? How can it make a difference toward the current well-established quantum models that have been proposed in the literature?

Since each decision problem is different, we believe that a quantum decision model would benefit from a dynamic heuristic that could take into account the decision problem's settings and come up with estimations for the quantum interference parameters. In the proposed model, quantum parameters are found based on the correlations that the vectors share between them. These correlations are explored through vector similarities that are computed using the Law of Cosines in a vector space. In this sense, we suggest that the quantum parameters that arise from interference effects might represent some degree of similarity between events. The previous work of Moreira and Wichert (2015) point out this semantic relation between vectors. In the end, the proposed model can be seen as a nonparametric method for estimating inference effects from a statistical point of view. It is a statistical model that is simpler than the previous Quantum Dynamical Model (Pothos and Busemeyer, 2009) and Quantum-Like Approach (Khrennikov, 2010) models proposed in the literature. The method makes use of the principles of Bayesian Networks, in order to obtain a more general and scalable model that can produce competitive results over the current state of the art models.

Experimental data demonstrated that the proposed heuristic managed to produce accurate fits to the data, overcoming the previously proposed Quantum Prospect Theory. This suggests that taking into account a dynamic estimation of quantum parameters is a good direction to build quantum-like predictive models.

ACKNOWLEDGMENTS

REFERENCES

Aerts, D., Broekaert, J., and Smets, S. (2004). A quantum structure description of the liar paradox. *Int. Theor. Phys.* 38, 3231–3239. doi: 10.1023/A:1026686316673

Birnbaum, M. (2008). New paradoxes of risky decision making. *Psychol. Rev.* 115, 463–501. doi: 10.1037/0033-295X.115.2.463

Busemeyer, J., and Bruza, P. (2012). *Quantum Model of Cognition and Decision.* New York, NY: Cambridge University Press. doi: 10.1017/cbo9780511997716

Busemeyer, J., Matthew, M., and Wang, Z. (2006a). "A quantum information processing explanation of disjunction effects," in *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (Vancouver, BC).

Busemeyer, J., Wang, Z., and Lambert-Mogiliansky, A. (2009). Empirical comparison of markov and quantum models of decision making. *J. Math. Psychol.* 53, 423–433. doi: 10.1016/j.jmp.2009.03.002

Busemeyer, J., Wang, Z., and Shiffrin, R. (2015). Bayesian model comparison favours quantum over standard decision theory account of dynamic inconsistency. *Decision* 2, 1–12. doi: 10.1037/dec0000017

Busemeyer, J., Wang, Z., and Townsend, J. (2006b). Quantum dynamics of human decision making. *J. Math. Psychol.* 50, 220–241. doi: 10.1016/j.jmp.2006.01.003

Busemeyer, J., Wang, Z., and Trueblood, J. (2012). "Hierarchical bayesian estimation of quantum decision model parameters," in *Proceedings of the 6th International Symposium on Quantum Interactions* (Paris). doi: 10.1007/978-3-642-35659-9_8

Croson, R. T. (1999). The disjunction effect and reason-based choice in games. *Organ. Behav. Hum. Decis. Process.* 80, 118–133. doi: 10.1006/obhd.1999.2846

Epstein, L. (1999). A definition of uncertainty aversion. *Rev. Econ. Stud.* 66, 579–608. doi: 10.1111/1467-937X.00099

Griffiths, T., Kemp, C., and Tenenbaum, J. (2008). "Bayesian models of inductive learning," in *Proceedings of the Annual Conference of the Cognitive Science Society* (Washington, DC).

Hristova, E., and Grinberg, M. (2008). "Disjunction effect in prisoner's dilemma: evidences from an eye-tracking study," in *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (Washington, DC).

Khrennikov, A. (1999). Classical and quantum mechanics on information spaces with applications to cognitive psychological, social and anomalous phenomena. *Found. Phys.* 29, 1065–1098. doi: 10.1023/A:1018885632116

Khrennikov, A. (2001). Linear representations of probabilistic transformations induced by context transitions. *Physics A* 34, 9965–9981. doi: 10.1088/0305-4470/34/47/304

Khrennikov, A. (2003a). Quantum-like formalism for cognitive measurements. *BioSystems* 70, 211–233. doi: 10.1016/S0303-2647(03)00041-8

Khrennikov, A. (2003b). Representation of the kolmogorov model having all distinguishing features of quantum probabilistic model. *Phys. Lett. A* 316, 279–296. doi: 10.1016/j.physleta.2003.07.006

Khrennikov, A. (2004). On quantum-like probabilistic structure of mental information. *Open Syst. Infor. Dyn.* 11, 267–275. doi: 10.1023/B:OPSY.0000047570.68941.9d

Khrennikov, A. (2005a). "From classical statistical model to quantum model through ignorance of information," in *Proceedings of the Third Conference on the Foundations of Information Science* (Paris).

Khrennikov, A. (2005b). Linear and nonlinear analogues of the Schrödinger equation in the contextual approach in quantum mechanics. *Doklady Math.* 72, 791–794. Available online at: https://inis.iaea.org/search/searchsinglerecord.aspx?recordsFor=SingleRecord&RN=38042995

Khrennikov, A. (2005c). Representation of the contextual statistical model by hyperbolic amplitudes. *Math. Phys.* 46, 1–13. doi: 10.1063/1.1931042

Khrennikov, A. (2006). Quantum-like brain: interference of minds. *BioSystems* 84, 225–241. doi: 10.1016/j.biosystems.2005.11.005

Khrennikov, A. (2009a). Description of composite quantum systems by means of classical random fields. *Found. Phys.* 40, 1051–1064. doi: 10.1007/s10701-009-9392-8

Khrennikov, A. (2009b). *Interpretations of Probability*. Berlin: Springer. doi: 10.1515/9783110213195

Khrennikov, A. (2009c). Quantum-like model of cognitive decision making and information processing. *BioSystems* 95, 179–187. doi: 10.1016/j.biosystems.2008.10.004

Khrennikov, A. (2010). *Contextual Approach to Quantum Formalism*. Dordrecht: Springer.

Khrennikov, A., Basieva, I., Dzhafarov, E., and Busemeyer, J. (2014). Quantum models for psychological measurements: an unsolved problem. *PLoS ONE* 9:e110909. doi: 10.1371/journal.pone.0110909

Koller, D., and Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA: The MIT Press.

Kuhberger, A., Komunska, D., and Josef, P. (2001). The disjunction effect: does it exist for two-step gambles? *Organ. Behav. Hum. Decis. Process.* 85, 250–264. doi: 10.1006/obhd.2000.2942

Lambdin, C., and Burdsal, C. (2007). The disjunction effect reexamined: relevant methodological issues and the fallacy of unspecified percentage comparisons. *Organ. Behav. Hum. Decis. Process.* 103, 268–276. doi: 10.1016/j.obhdp.2006.04.001

Leifer, M., and Poulin, D. (2008). Quantum graphical models and belief propagation. *Ann. Phys. J.* 323, 1899–1946. doi: 10.1016/j.aop.2007.10.001

Li, S., and Taplin, J. (2002). Examining whether there is a disjunction effect in prisoner's dilemma game. *Chin. Psychol.* 44, 25–46. Available online at: https://digital.library.adelaide.edu.au/dspace/handle/2440/3240?mode=full&submit_simple=Show+full+item+record

Moreira, C., and Wichert, A. (2014). Interference effects in quantum belief networks. *Appl. Soft Comput.* 25, 64–85. doi: 10.1016/j.asoc.2014.09.008

Moreira, C., and Wichert, A. (2015). The synchronicity principle under quantum probabilistic inferences. *NeuroQuantology* 13, 111–133. doi: 10.14704/nq.2015.13.1.788

Nielsen, M. A., and Chuang, I. L. (2000). *Quantum Computation and Quantum Information*. Cambridge, UK: Cambridge University Press.

Nyman, P. (2010). On consistency of the quantum-like representation algorithm. *Theor. Phys.* 49, 1–9. doi: 10.1007/s10773-009-0171-2

Nyman, P. (2011a). On hyperbolic interferences in the quantum-like representation algorithm for the case of triple-valued observables. arXiv:1108.2194v1 [quant-ph].

Nyman, P. (2011b). On the consistency of the quantum-like representation algorithm for hyperbolic interference. *Adv. Appl. Cliff. Algeb.* 21, 799–811. doi: 10.1007/s00006-011-0287-3

Nyman, P., and Basieva, I. (2011a). Quantum-like representation algorithm for trichotomous observables. *Theor. Phys.* 50, 3864–3881. doi: 10.1007/s10773-011-0934-4

Nyman, P., and Basieva, I. (2011b). "Representation of probabilistic data by complex probability amplitudes - the case of triple-valued observables," in *Proceedings of the International Conference on Advances in Quantum Theory* (Växjö).

Osherson, D. (1995). *Thinking*. Cambridge, MA: The MIT Press.

Pothos, E., and Busemeyer, J. (2009). A quantum probability explanation for violations of rational decision theory. *Proc. R. Soc. B* 276, 2171–2178. doi: 10.1098/rspb.2009.0121

Pothos, E., Busemeyer, J., and Trueblood, J. (2013). A quantum geometric model of similarity. *Psychol. Rev.* 120, 679–696. doi: 10.1037/a0033142

Rieffel, E., and Polak, W. (2011). *Quantum Computing: A Gentle Introduction*. Cambridge, MA: The MIT Press.

Russel, S., and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach, 3rd Edn.* Englewood Cliffs, NJ: Pearson Education.

Savage, L. (1954). *The Foundations of Statistics*. New York, NY: John Wiley.

Shafir, E., and Tversky, A. (1992). Thinking through uncertainty: nonconsequential reasoning and choice. *Cogn. Psychol.* 24, 449–474. doi: 10.1016/0010-0285(92)90015-T

Shah, A., and Oppenheimer, D. (2008). Heuristics made easy: an effort-reduction framework. *Psychol. Bull.* 134, 207–222. doi: 10.1037/0033-2909.134.2.207

Sterman, J. (1989). Deterministic chaos in an experimental economic system. *Econ. Behav. Organ.* 12, 1–28. doi: 10.1016/0167-2681(89)90074-7

Tucci, R. (1995). Quantum bayesian nets. *Int. Mod. Phys. B* 9, 295–337. doi: 10.1142/S0217979295000148

Tversky, A., and Kahneman, D. (1983). Extension versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychol. Rev.* 90, 293–315. doi: 10.1037/0033-295X.90.4.293

Tversky, A., and Kahnenman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131. doi: 10.1126/science.185.4157.1124

Tversky, A., and Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychol. Sci.* 3, 305–309. doi: 10.1111/j.1467-9280.1992.tb00678.x

Yukalov, V., and Sornette, D. (2008). Quantum decision theory as quantum theory of measurement. *Phys. Lett. A* 372, 6867–6871. doi: 10.1016/j.physleta.2008.09.053

Yukalov, V., and Sornette, D. (2009a). Physics of risk and uncertainty in quantum decision making. *Eur. Phys. J. B* 71, 533–548. doi: 10.1140/epjb/e2009-00245-9

Yukalov, V., and Sornette, D. (2009b). Processing information in quantum decision theory. *Entropy* 11, 1073–1120. doi: 10.3390/e11041073

Yukalov, V., and Sornette, D. (2010a). Entanglement production in quantum decision-making. *Phys. Atom. Nucl.* 73, 559–562. doi: 10.1134/S106377881003021X

Yukalov, V., and Sornette, D. (2010b). Mathematical structure of quantum decision theory. *Adv. Comp. Syst.* 13, 659–698. doi: 10.1142/S0219525910002803

Yukalov, V., and Sornette, D. (2011). Decision theory with prospect interference and entanglement. *Theor. Decis.* 70, 283–328. doi: 10.1007/s11238-010-9202-y

Zurek, W. (2005). Probabilities from entanglement, born's rule from envariance. *Phys. Rev. A* 71, 1–29. doi: 10.1103/PhysRevA.71.052105

Zurek, W. (2011). Entanglement symmetry, amplitudes, and probabilities: inverting born's rule. *Phys. Rev. Lett.* 106, 1–5. doi: 10.1103/PhysRevLett.106.250402

Progress and current challenges with the quantum similarity model

Emmanuel M. Pothos[1], Albert Barque-Duran[1]*, James M. Yearsley[1], Jennifer S. Trueblood[2], Jerome R. Busemeyer[3] and James A. Hampton[1]

[1] Department of Psychology, City University London, London, UK
[2] Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, USA
[3] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA
*Correspondence: albert.barque-duran@city.ac.uk

This opinion paper reviews progress with the quantum similarity model (QSM), which was proposed by Pothos et al. (2013). In the QSM, concepts are associated with subspaces, the mental state is a state vector in a Hilbert space, and similarity between two concepts is computed in terms of the sequential projection, between the corresponding subspaces. If there is a relevant context, this is incorporated as prior projections (e.g., **Figure 1**).

The QSM was developed as a way to primarily cover the empirical findings of Tversky (1977). Tversky (1977) reported a series of results for (mostly) simple (non-analogical, see below) pairwise similarity judgments. Tversky's (1977) research severely challenged the dominant similarity models at that time, based on metric spaces and distances. Such models are constrained to obey the metric axioms (as long as similarities are simple functions of distances). Yet, in his seminal work, Tversky reported violations of all three metric axioms (minimality, symmetry, triangle inequality), in the similarity judgments of naïve observers. Moreover, Tversky reported a so-called diagnosticity effect, where the same similarity judgment could change greatly, depending on which other stimuli were present in a (broadly) relevant context set.

All of Tversky's (1977) findings reveal intuitions about human similarity that are, initially at least, very surprising. For

example, how can it be possible that the similarity between (simplifying his example) China and Korea be less than Korea and China? Yet, some thought shows that we indeed prefer to judge a non-prominent object (e.g., Korea) as more similar to a prominent one (e.g., China), as compared to the reverse order. Equally, how can it be possible that Austria is seen as more similar to Sweden than to Hungary in the context of Poland, but more similar to Hungary than to Sweden in the context of Norway?

Tversky's findings have been a major focus of subsequent theoretical work on similarity judgments. Some of the most prominent models are the distance-density model (Krumhansl, 1978), general recognition theory (Ashby and Perrin, 1988), and the generalized context model (Nosofsky, 1988; this is a theory of categorization, rather than similarity, yet Nosofsky considered in his influential work how to accommodate Tversky's findings as well, e.g., Nosofsky, 1991). Limited space prevents us from a detailed analysis of this work. Overall, we think that while such work has provided many excellent intuitions regarding human similarity, its application to Tversky's (1977) findings is not uniformly satisfactory. This was a consideration that in part motivated the QSM.

Another motivating consideration has been the recently proposed model for the conjunction fallacy, based on quantum

theory (Busemeyer et al., 2011). The conjunction fallacy is a famous result in decision making, whereby naïve observers judge a hypothetical person, Linda, to be more likely to be both a Bank teller and a feminist, than just a bank teller (Tversky and Kahneman, 1983). Of course, such a result is paradoxical, if one employs the rules of classical probability theory. Tversky and Kahneman (1983) suggested that naïve observers in their experiment employed a so-called representativeness heuristic, judging Linda to be more similar to the category of bank tellers and feminists. Thus, at the heart of the explanation for the conjunction fallacy is the idea that participants employ a similarity process (see also Shafir et al., 1990, for further validations of this idea). The quantum model for the conjunction fallacy indeed reflects operations that involve the overlap of a state vector (representing the mental state of participants) and subspaces (which correspond to different concepts in the participants' knowledge space, e.g., the idea that a woman can be a feminist; cf. Sloman, 1993). Thus, we were interested in whether the quantum model for the conjunction fallacy could be extended, more or less as it is, to function as a model of some aspects of similarity. This was indeed the approach that was adopted by Pothos et al. (2013) and the QSM is structurally and procedurally nearly entirely equivalent to Busemeyer et al's (2011) model for the conjunction fallacy. That the same
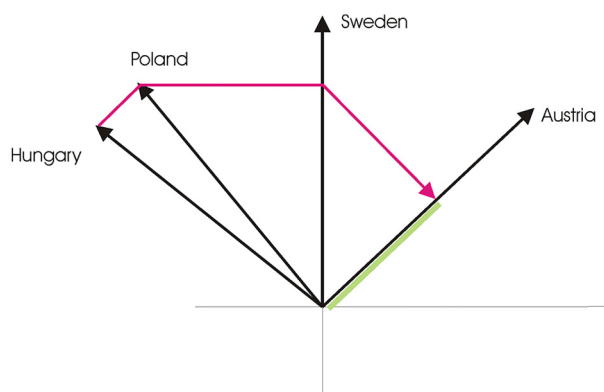
FIGURE 1 | An illustration of the series of projections, relevant to the similarity of Sweden to Austria, in the context of Poland and Hungary (assuming all countries are represented as rays). The red line shows the series of projections: $P_{Hungary}P_{Poland}P_{Sweden}P_{Austria} | \psi$). Similarity is the squared length of this projection (indicated in green). The last two projections correspond to the similarity comparison and the first two to the context. Note, also, that projections across context elements need be counterbalanced, but, for simplicity, in the example we illustrate only one order (from Hungary to Poland). Finally, with this context, the similarity judgment is going to be higher, compared to having context elements not grouped together.

principles can provide a route for explaining both aspects of decision-making and similarity enables the exciting possibility that a formal unification may be possible between these two seemingly disparate aspects of cognition.

One emphasis of the QSM has been the demonstration of asymmetries in similarity judgments. In the QSM this arises in part because concepts are represented as subspaces. Note that the use of subspaces as such is not a uniquely quantum feature of the QSM, but the lack of commutativities in projection sequence (which contributes to the emergence of asymmetries) is. Subspaces can have rich inner structure, corresponding e.g., to the characteristics of a concept. Thus, concepts for which we have more knowledge (such as China, if we imagine ourselves in the shoes of Tversky's participants in 1977) will be represented by a higher dimensionality subspace, contrasting with concepts for which we have less knowledge (such as Korea). Together with an assumption that the mental state prior to a (simple) similarity comparison is neutral between the two concepts to be compared, this enables a natural emergence of asymmetries in human similarity judgments, in the predicted direction. More generally, conceptually, we think that representations as subspaces are an important advance. This is because representations in the QSM can

have inner structure, not just in terms of a list of characteristics, but also in terms of how the characteristics relate to each other. By contrast, in traditional spatial representations, with concepts being represented as points or vectors, there is no possibility of such structure at all. This would be the case even in e.g., Latent Semantic Analysis approaches to representation, which have proved extremely useful and influential (e.g., Dumais, 2004; see also Kitsch, 2014; Kitsch, for an insightful comparison between the QSM and Latent Semantic Analysis; note that in Kitsch's (2014), approach, vectors are given variable length, and this can capture differences in degree of knowledge). But even in Tversky's (1977) feature-based approach, concepts would be lists of features, and Tversky (1977) did not consider how dependencies among features could be incorporated in his model.

The way violations of the triangle inequality arise in the QSM is very similar to how Tversky (1977) suggested such effects arise. Because in the QSM representations are subspaces, different regions in the overall space end up reflecting the features characteristic of the corresponding concepts. So, for example, imagine a region in the overall space with Russia and Cuba. This region will overall reflect the property of communism, noting that both Russia and Cuba are consistent with

this property (thinking again as participants in Tversky's experiment in 1977). Then, imagine a region different to the first one containing Cuba and Jamaica. The shared characteristic of Cuba and Jamaica is their geographical proximity (they are both in the Caribbean), so this second region will likewise correspond to this property. It should be hopefully straightforward to then see how, if Cuba is on the boundary of the communism and Caribbean regions in psychological space, we can have Cuba highly similar to Russia, Cuba highly similar to Jamaica, but Russia and Jamaica dissimilar from each other, thus violating the triangle inequality. It has to be noted, however, that the triangle inequality is not a challenge for standard (non-linear) distance-based models of similarity. This is because the triangle inequality is already violated if one relates distance and similarity, via a non-linear function (such as the standard exponentially decaying function; Nosofsky, 1984; Shepard, 1987). Nevertheless, it is clearly important for a similarity model to cover violations of the triangle inequality in a convincing manner. Note, violations of the triangle inequality have been the focus of an alternative similarity model, based on quantum theory (Aerts et al., 2011).

A great focus for further work with the QSM concerns the diagnosticity effect. This is because the diagnosticity effect has proved difficult to replicate (e.g., see Evers and Lakens, 2014). We are interested in exploring whether the QSM model can provide insight into why the diagnosticity effect has proved elusive in its replicability. The diagnosticity effect is also significant because the quantum formalism, overall, is often said to embody strong contextual influences. So, perhaps, quantum theory would be particularly suitable for modeling context effects in similarity judgments? Well, the diagnosticity effect does emerge fairly naturally from the QSM, but the mechanisms that allow this are not the traditional contextual mechanisms in quantum theory (e.g., relating to entanglement or incompatibility). In the QSM, the contextual influences relevant to the diagnosticity effect emerge from the way prior projections are used to capture sensitivity to the grouping of context elements. In other words, the difficulty lies in the fact that contextual

influences in similarity specifically depend on the degree of grouping of some of the options in the relevant choice set. For example, in Tversky's (1977) demonstration, participants were asked to decide which country is most similar to Austria, between Sweden, Hungary, and Poland. More participants chose Sweden, but when the choice set included Sweden, Hungary, and Norway, they chose Hungary. What we might call the "traditional" mechanisms for contextual influences in quantum theory are not sensitive to the similarity structure of the relevant options.

Contextual influences in the QSM arise in the following way. Similarity computations are based on projecting (laying down) the state vector (which represents the current mental state) onto different subspaces (which represents the concepts relevant in the similarity task; **Figure 1**). This projection operation can be highly order dependent in quantum theory. Of relevance, the outcome of a projection sequence is sensitive to the grouping of the subspaces across which projection takes place. If the subspaces are grouped together, then a projection sequence preserves the length of the state vector and vice versa. Thus, to account for the diagnosticity effect in the QSM, we postulated that, in a forced choice task (such as the one employed by (Tversky, 1977), in his diagnosticity formulation), prior to the projections corresponding to the elements in the similarity judgment, there would be projections corresponding to the other elements in the choice set. So, for example, if a participant is considering which between Sweden, Hungary, Poland is most similar to Austria, and is specifically evaluating the option of Sweden, then the similarity comparison would consist of projections from Sweden to Austria, but also there would be prior projections to Hungary and Poland. Using this scheme, with fairly minimal assumptions about the representation of the relevant stimuli, the diagnosticity effect emerges from the QSM.

One important challenge in further developing the QSM is further formalizing the way contextual influences are taken into account. The idea of incorporating context as prior projections works well, but it has a heuristic feel to it. Can the QSM be extended such that these prior

projections can be motivated in a more rigorous way (cf. Lambert-Mogiliansky et al.'s, 2009, quantum model of framing effects)? Moreover, as noted, can the QSM generate any new predictions regarding the emergence or suppression of the diagnosticity effect? Since Tversky's (1977) work, there has not really been much further examination (or little that has reached the journals), which is surprising (in the sense that the idea of context in similarity judgments seems like a vast topic). These questions are an important focus for our current work with the QSM.

Another important focus concerns so-called analogical similarity judgments (e.g., Goldstone, 1994; Gentner and Markman, 1997). Analogical similarity refers to the idea that, for example, if we are comparing two people, Jim and Jack, if they both have black hair, this will increase their similarity, but if Jim has black hair and Jack has black shoes (and blond hair), this will have less impact on their similarity. That is, work on analogical similarity recognizes that objects often consist of separate components. Commonalities on matching components (e.g., black hair) increase similarity more so than commonalities on mismatching components (e.g., black hair and black shoes). It is currently unclear whether there is a genuine distinction between cognitive processing corresponding to basic similarity tasks (as in Tversky, 1977) and analogical similarity ones (some researchers have suggested that different cognitive systems may mediate the two types of judgments; Casale et al., 2012). Nevertheless, there have been largely separate corresponding literatures, with different objectives. We think that the QSM can be extended to incorporate analogical similarity, because quantum theory already has extensive machinery in place for combining individual components into a whole (cf. Smolensky, 1990). We have been pursuing an approach based on tensor products and we are optimistic that a concrete proposal will be forthcoming soon (Pothos and Trueblood, 2015).

Finally, the QSM is only part of a broader effort within the quantum cognition community to understand similarity using quantum processes. A more challenging, though important objective, would be to examine the formal relation

between QSM and, for example, Aerts's (2009) model for conceptual combination or Lambert-Mogiliansky et al.'s (2009) model of framing effects.

REFERENCES

Aerts, D. (2009). Quantum structure in cognition. *J. Math. Psychol.* 53, 314–348. doi: 10.1016/j.jmp.2009.04.005

Aerts, S., Kitto, K., and Sitbon, L. (2011). "Similarity metrics within a point of view," in *Proceedings of the Quantum Interaction Conference*. Berlin; Heidelberg: Springer Verlag. doi: 10.1007/978-3-642-24971-6_3

Ashby, G. F., and Perrin, N. A. (1988). Towards a unified theory of similarity and recognition. *Psychol. Rev.* 95, 124–150. doi: 10.1037/0033-295X.95.1.124

Busemeyer, J. R., Pothos, E. M., Franco, R., and Trueblood, J. (2011). A quantum theoretical explanation for probability judgment errors. *Psychol. Rev.* 118, 193–218. doi: 10.1037/a0022542

Casale, M. B., Roeder, J. L., and Ashby, F. B. (2012). Analogical transfer in perceptual categorization. *Mem. Cognit.* 40, 434–449. doi: 10.3758/s13421-011-0154-4

Dumais, S. T. (2004). Latent semantic analysis. Microsoft research. *Annu. Rev. Inform. Sci. Technol.* 38, 188–230 doi: 10.1002/aris.1440380105

Evers, E. R. K., and Lakens, D. (2014). Revisiting Tversky's diagnosticity principle. *Front. Psychol.* 5:875. doi: 10.3389/fpsyg.2014.00875

Gentner, D., and Markman, A. B. (1997). Structure mapping in analogy and similarity. *Am. Psychol.* 52, 45–56. doi: 10.1037/0003-066X.52.1.45

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 3–28. doi: 10.1037/0278-7393.20.1.3

Kitsch, W. (2014). Similarity as a function of semantic distance and amount of knowledge. *Psychol. Rev.* 121, 559–561. doi: 10.1037/a0037017

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: the interrelationship between similarity and spatial density. *Psychol. Rev.* 85, 445–463. doi: 10.1037/0033-295X.85.5.445

Lambert-Mogiliansky, A., Zamir, S., and Zwirn, H. (2009). Type indeterminacy - A model of the KT (Kahneman Tversky)- man. *J. Math. Psychol.* 53, 349–361. doi: 10.1016/j.jmp.2009.01.001

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 104–114. doi: 10.1037/0278-7393.10.1.104

Nosofsky, R. M. (1988). Similarity, frequency, and category representation. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 54–65. doi: 10.1037/0278-7393.14.1.54

Nosofsky, R. M. (1991). Stimulus bias, asymmetric similarity, and classification. *Cogn. Psychol.* 23, 94–140. doi: 10.1016/0010-0285(91)90004-8

Pothos, E. M., Busemeyer, J. R., and Trueblood, J. S. (2013). A quantum geometric model of similarity. *Psychol. Rev.* 120, 679–696. doi: 10.1037/a0033142

Pothos, E. M., and Trueblood, J. S. (2015). Structured representations in a quantum probability model of similarity. *J. Math. Psychol.* 64, 35–43. doi: 10.1016/j.jmp.2014.12.001

Shafir, E. B., Smith, E. E., and Osherson, D. N. (1990). Typicality and reasoning fallacies. *Mem. Cognit.* 18, 229–239. doi: 10.3758/BF03213877

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science* 237, 1317–1323. doi: 10.1126/science.3629243

Sloman, S. A. (1993). Feature-based induction. *Cogn. Psychol.* 25, 231–280. doi: 10.1006/cogp.1993.1006

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artif. Intell.* 46, 159–216. doi: 10.1016/0004-3702(90)90007-M

Tversky, A. (1977). Features of similarity. *Psychol. Rev.* 84, 327–352. doi: 10.1037/0033-295X.84.4.327

Tversky, A., and Kahneman, D. (1983). Extensional versus intuitive reasoning: the conjunctive fallacy in probability judgment. *Psychol. Rev.* 90, 293–315. doi: 10.1037/0033-295X.90.4.293

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

A macroscopic violation of no-signaling in time inequalities? How to test temporal entanglement with behavioral observables

Patrizio E. Tressoldi[1]*, Markus A. Maier[2], Vanessa L. Buechner[2] and Andrei Khrennikov[3]

[1] *Dipartimento di Psicologia Generale, Università di Padova, Padova, Italy, [2] Psychology Department, Ludwig Maximilian University of Munich, Munich, Germany, [3] International Center for Mathematical Modeling in Physics, Engineering, Economics, and Cognitive Science, Linnaeus University, Växjö-Kalmar, Sweden*

In this paper we applied for the first time the no-signaling in time (*NSIT*) formalism discussed by Kofler and Brukner (2013) to investigate temporal entanglement between binary human behavioral unconscious choices at t1 with binary random outcomes at t2. *NSIT* consists of a set of inequalities and represents mathematical conditions for macro-realism which require only two measurements in time. The analyses of three independent experiments show a strong violation of *NSIT* in two out of three of them, suggesting the hypothesis of a quantum-like temporal entanglement between human choices at t1 with binary random outcomes at t2. We discuss the potentialities of using NSIT to test temporal entanglement with behavioral measures.

Keywords: no-signaling in time, temporal entanglement, non-local correlation in time, human choices, random events

Introduction

The possibility to use mathematical and statistical formalisms adopted in quantum mechanics for the study of biological (e.g., Engel et al., 2007; Blankenship and Engel, 2010) and cognitive phenomena (e.g., Wang et al., 2014) is not only a theoretical proposal but a rich field of empirical research (see Khrennikov, 2010; Busemeyer and Wang, 2014, for a review).

The application of quantum formalisms to domains other than quantum physics –such as biological or mental processes- is independent to the hypothesis that processing of information by biological systems is based on quantum physical processes within these systems. This approach known as "quantum biological information" is based on the *quantum-like paradigm*: biological systems of sufficiently high complexity process information in accordance with laws of quantum information theory (Khrennikov, 2010; Hameroff et al., 2014).

However, documenting the usefulness of such mathematical algorithms in modeling decision processes, memory, or consciousness, opens the possibility that the biological substrate constitutes the basis for the emergence of these quantum phenomena. This proposition is controversially discussed and only few researchers share this idea (see, e.g., Hameroff and Penrose, 2014). The main argument against the existence of quantum coherence or entanglement in biological systems like the brain refers to decoherence as a strong boundary condition of quantum phenomena (see, e.g., Tegmark, 2000; Jumper and Scholes, 2014). Decoherence of quantum states seems to occur with such a high frequency that these effects would be impossible to operate on macroscopically relevant spatial distances or time scales (Tegmark, 2000). This would imply that non-temporal correlations

between temporally separated events in the range of several 100 of milliseconds or even up to seconds would be highly unlikely. In other words, the brain or the parts of it that are involved in actual information processing constitute a macroscopic entity and non-temporal correlations for macroscopic events are quite rare or even impossible (see Tegmark, 2000; but see Hameroff et al., 2014).

Independently of the quantum mind discussion, recently, in psychology, non-temporal correlations between temporally separated events (from a few 100 ms up to several minutes) have been observed (see, e.g., Mossbridge et al., 2012; Bem et al., 2014; Maier et al., 2014). These phenomena usually involved a behavioral or physiological response at time 1 (RP t_1) and an activating event happening later at time 2 (AE t_2). In these studies a retro-causal influence and therefore temporally non-local correlations of AE t_2 on RP t_1 were reported.

Since the Maier et al.'s (2014) studies will be re-analyzed within this article, we will refer to their data in more detail here to illustrate the basic finding. In a series of four out of seven studies a selective key-press at time 1 (left or right) was affected by the random assignment of negative or non-negative picture presentations at time 2. On average the participants were able to avoid negative future events. The random assignment at t_2 was performed based on a pseudo random number generator (PRNG) in Studies 1, 2, and 3 and with a quantum based random number generator (RNG) in Study 4. In other words the events at t_1 and t_2 were classically uncorrelated. The findings, however, indicated that event t_1 was affected by event t_2 which could only be the case if these macroscopically occurring events were in a state of temporally non-local correlation. Although Maier et al. (2014) reported a significant avoidance effect at t_1 being affected by the event on t_2, a direct test of temporal non-locality has not been performed. The goal of the data presented here is to fill this gap by providing such a test.

Entanglement in time or temporal non-locality, that is a non-causal correlation between events measured at successive time frames, is one of the many "odd" phenomena studied in quantum physics and mathematical tools have been developed to test the existence of these effects within the empirical data.

Although a commonly accepted mathematical algorithm for a strict test of temporal non-locality does not exist, some mathematical inequalities that can be applied to temporally distinct physical or mental states have been developed to test the quantum-nature of the underlying physical or cognitive mechanisms. If the inequalities applied to the data are found to be violated, they would indicate the involvement of superposed states.

Contextual LG Inequality and No-Signaling in Time (NSIT) Inequality

The theoretical foundations were originally discussed by Leggett and Garg (1985) as a temporal variant of John Bell inequalities which mainly address entanglement or non-local correlations in space. A violation of the Leggett–Garg-equation would confirm quantum-like superposed states between temporally separated

events and is thus a pendant of the Bell inequalities for the time dimension. Whereas non-local temporal effects are intensely investigated in quantum physics (e.g., Olson and Ralph, 2012; Aharonov et al., 2014), there are still only few analyses of this type applied to human cognition. Atmanspacher and Filk (2010, 2012, 2013) were probably the first to test temporal non-locality to bistable perception applying their Necker–Zeno model which requires three different measures. Similarly, Asano et al. (2014), derived an analog of the Leggett and Garg (1985) inequality, "contextual LG inequality," and used it as a test of "quantum-likeness" of statistical data collected in a series of experiments on recognition of ambiguous figures. The Leggett–Garg approach has some limitations since this test can only be applied for situations involving three consecutively occurring events. For two event scenarios, as is the case in the Maier et al. (2014) research, the Leggett–Garg equation cannot be used. Fortunately, recently a test of non-local correlations for two consecutive events has been developed (Kofler and Brukner, 2013).

The No-Signaling in Time (NSIT) Inequality

Kofler and Brukner (2013), discuss *NSIT* as a further necessary condition to satisfy the Leggett–Garg inequalities to test macro-realism defined by the postulates that (a) macroscopic objects which may have two or more macroscopically different states, at any given time, are in a single specific state, (b) it is possible to measure this specific state without changing it, and, (c) the properties of this macroscopic object are determined exclusively by the initial conditions.

NSIT requires only two measurements in time of two dichotomous observables, A and B, that may assume only two distinct states ± 1. Hence, the basic scenario is: $A_{t1} = \pm 1$, $B_{t1} = \pm 1$ and $A_{t2} = \pm 1$, $B_{t2} = \pm 1$.

In accordance with the principle of *NSIT* the outcome probabilities for one part must not depend on the outcome probabilities of the second part and it is expressed by the following formula:

$$P(B_{t2} = +1) = P(A_{t1} = -1, B_{t2} = +1) +$$
$$P(A_{t1} = +1, B_{t2} = +1) \text{ and symmetrically}$$
$$P(B_{t2} = -1) = P(A_{t1} = +1, B_{t2} = -1) +$$
$$P(A_{t1} = -1, B_{t2} = -1) \qquad (1)$$

A violation of *NSIT* condition could be a first indicator that the mental state evolution cannot be described classically and may be explained by temporally distinct cognitive states existing in a state of superposition.

It is important to note that the temporal non-locality interpretation of NSIT is not straightforward and commonly accepted within the scientific community. The most accepted interpretation of violations of NSIT is that the data that violate these equalities are based on cognitive processes that most likely behaved quantum like. This includes the possibility that the underlying mechanisms are best described as information states that co-exist in a state of superposition. Such a quantum-like

behavior of cognitive states could be considered as being a pre-condition for temporal non-locality to occur. In the analyses presented here we tested this pre-condition. To our knowledge, this is the first attempt to test this formalism in human behavioral tasks.

Materials and Methods

Description of the Experimental Data

Here we report the analyses of the three formal experiments in Maier et al.'s (2014) work, Study 1, Study 2, and Study 4 carried out with participants in the laboratory and with identical conditions and instructions to the participants. Our selection was based on the fact that only in these studies a retro-causal effect of t_2 on t_1 was observed. One successful study, Study 3, was eliminated since it was completed by a web-based program and participants could not be monitored during their task execution. Thus, only methodologically rigorously obtained data were included. A more detailed description of these experiments is presented in Maier et al. (2014).

Participants

In all experiments participants were recruited among the undergraduate and graduate students of the University of Munich, Stony Brook NY, and Barcelona. The number of participants was 111, 201, and 327 for Study 1, Study 2, and Study 4, respectively.

Procedure

Each participant was tested individually in a quiet lab room. After the completion of two preliminary tasks, lasting approximately 20 min and being unrelated to the crucial study, which were devised in order to increase the cognitive fatigue for inducing a more intuitive approach, participants were informed about their new task. A written instruction was presented on the screen: '*In the following experiment you have to press two keys on the keyboard as simultaneously as possible. You will see this instruction on the monitor's screen: Please Press the Keys. While seeing this instruction, please press both keys as simultaneously as possible! Afterward colored stimuli will be presented which you should simply watch.*'

After the participants read the instructions, the experimenter explained that the participants should put their index fingers on the left and right cursor keys of the keyboard. Both keys were placed on the table in front of the participants exactly at the same horizontal position as the midpoint of the computer screen. The experimenter emphasized that both index fingers should slightly touch the cursor keys throughout the experiment, and once the command appears they should press both keys as simultaneously as possible. Participants were informed that there is no rush, but the response should be spontaneous, and that after the key-press they should simply watch the following presentation of a colored stimulus.

Each trial started with the key-press command presented on the screen. Once the key-press was performed, the command line disappeared and, after a 500 ms delay with a black screen, a masked positive (Study 1 or neutral, Studies 2 and 4) or negative picture was presented. The masked picture presentation consisted of three consecutive stimulus presentations.

First, a masking stimulus was presented for 72 ms, followed by the presentation of a negative or positive (neutral) picture for 18 ms, again followed by the same mask for 72 ms. Each negative and positive (neutral) picture was combined with an individual mask. The masking stimulus was constructed by dividing the original picture into small squares that were randomly rearranged. The resulting mask consisted of the same color and lightness properties as the original picture and could therefore effectively mask the content of the picture ensuring a subliminal presentation. According to our theoretical model, subliminal perception is critical to allow a superposition of the information states in time. After the second masking stimulus had disappeared, a 3000 ms inter-trial interval appeared before the key-press command initiated the next trial. A total of 60 trial presentations were used in all studies. The 60 experimental trials were preceded by three practice trials with neutral pictures helping the participants to familiarize themselves with the task. Pictures were taken from the International Affective Picture System (IAPS; Lang et al., 2008).

Although participants were told to press both keys simultaneously, due to the design of a typical computer keyboard, one of two keys is always triggered first. Thus, in any given trial, either a left or a right key-press was registered even though participants subjectively performed a simultaneous two-key-response. For Studies 1 and 2 a closed deck procedure was applied, that is in half of the trials, triggering a left key resulted in a positive (neutral) masked picture presentation and a right key in a negative one. In the other half, key and valence assignment were exactly reversed. The randomization procedure provided by E-Prime[TM] was used to randomize the order of trial presentation. The 10 positive and 10 negative pictures were randomly assigned to each trial with the restrictions that each picture could maximally be presented six times within a study [i.e., if a participant always 'chooses' a positive picture presentation, 60 (6×10) positive (neutral) pictures would be presented]. In Study 4 an open deck procedure was used, that is the exact assignment to left and right key press and neutral vs. negative picture presentation was abandoned. Also, in this study a quantum-based randomizer, i.e., a true RNG, was used for randomization. Randomized trial selection was performed at the beginning of each trial. After the completion of the 60 trials participants saw each masked picture presentation again and were asked after each whether they could recognize anything and, if so, what.

None of the participants in each of the experiments reported here could precisely name the content of any picture. Thus, the masking procedure met the criterion of subjective unawareness (from Maier et al., 2014, pp. 130–132).

In Studies 2 and 4, material, design, and procedure were the same as in Study 1 with the one difference that the 10 negative pictures from Study 1 were used together with 10 neutral instead of positive pictures. Again, the pictures were taken from the IAPS; (Lang et al., 2008).

In Study 4, the only difference with respect to Study 2 was that the randomization was obtained by using a quantum based number generator (QRNG) from www.idquantique.com.

Formal Mathematical Representation

There are two random variables $A = A_{t_1}$ and $B = B_{t_2}$. The first one corresponds to the first task where the right and left keys determine the values $A = +1$ and $A = -1$, respectively. The nature of another variable is more complicated. The task at t_2 determining B is the subliminal perception of a positive or a negative emotional picture. In psychology this task is considered a "response." Now if we assume that these random variables can be represented in the classical probabilistic framework, i.e., there can be introduced the joint probability distribution for their values $P(A = x, B = y)$, the additivity of probability implies that $P(B = y) = P(A = +1, B = y) + P(A = -1, B = y)$.

Typically in applications this equality is treated in the form of the formula of total probability

$$P(B = y) = P(A = +1)P(B = y/A = +1) +$$
$$P(A = -1)P(B = y/A = -1)$$

This formula is violated in a variety of psychological tasks related to disjunction, conjunction and order effects and various probability fallacies (see, for example, Khrennikov, 2010; Busemeyer and Bruza, 2012; Wang et al., 2014).

The main distinguishing feature of the present study is that we couple the violation of the formula of total for statistical data collected in experiments with humans with (non)signaling problem in quantum physics, i.e., time is fundamentally involved into the experimental scheme.

Application of NSIT Formalism

The left-hand side of equation (1) $P(B_{t2} = \pm 1)$ was estimated with a mean equal to 0.5 and a SD of 0.5 assuming a correct randomization.

The probabilities of the right-hand of equation 1, were empirically drawn cross-tabulating the data obtained in the three experiments (see Supplementary Material).

Following the suggestion of Khrennikov et al. (2014), we estimated the SEM of $P(B_{t2} = \pm 1)$ taking in account the number of trials of each experiment. The ratio of the observed NSIT with the SEM was used as an estimate of the NSIT violation.

Results

In **Table 1** we report the results of the application of the NSIT inequality and the standardized deviation with respect to the $P(B_{t2} = \pm 1)$ in SE.

The $\Delta\sigma$ values which represent the violation of NSIT inequality in term of the number of SE from the expected probability at t_2, 0.5 in our case, show a clear and strong NSIT violation both in the first two experiments and in the analysis of the total trials weighted for the number of trials. It is unclear to us why the NSIT analysis did not reveal a violation for Study 4.

TABLE 1 | Results of the three experiments.

Study	N trials	SE	*NSIT*	$\Delta\sigma$
Study 1	6660	0.006126	**0.068**	**11.1**
Study 2	8160	0.005535	**0.261**	**47.23**
Study 4	19611	0.003570	**0.000**	**0.00**
Total	34431	0.002694	**0.0757**	**29.09**
Weighted total			**0.0279**	**10.37**

SEM, standard error of mean; NSIT, no-signaling in time; $\Delta\sigma = NIST/SE$.

One reason could be the different approaches to realize the trial randomization. Although PRNGs have been used in Studies 1 and 2 and a true RNG was applied in Study 4, PRNG and trueRNG both equally produce random events especially when the seed number and the algorithm used for the PRNG procedure was unknown to the participants, which is the case for our Studies 1 and 2. Raw data for independent analyses are available on http://figshare.com/articles/No_Signaling_in_Time_Raw_Data/1383260.

Discussion

Applying quantum mathematical formalisms to test the quantum-likeness of cognitive and behavioral phenomena is becoming more and more popular within the scientific community. In this study we applied the NSIT formalism to investigate temporal entanglement between binary human behavioral unconscious choices at t1 with binary random outcomes at t2. This is, to our knowledge, the first time that NSIT formalism has successfully been applied to psychological data sets. The results of three independent experiments showed a strong violation of NSIT suggesting the hypothesis of a quantum-like temporal entanglement between the choices at t1 with binary random outcomes at t2 in Studies 1 and 2. However, a null result was observed in Study 4. Overall, it seems that for the majority of the data evidence for temporal entanglement could be found.

Our results therefore support the idea of exploring quantum phenomena with data obtained with psychological variables involving unconscious decision making based on automatic affective processes. NSIT could thus be a valuable tool to test quantum effects in similar paradigms since most psychological experiments consists of activating events and corresponding responses. The main goal of our analyses was to introduce this powerful set of inequalities to a broader psychologically interested scientific community.

In any event, it is too early to be able to draw firm conclusion about the effect of the differences between the studies on the outcome of the NSIT analysis. At the moment, a pre-registered replication of Study 4 is being undertaken and will be completed in about 1 year. An additional analysis of these data with NSIT will shed some more light on the usefulness and applicability of the NSIT theorem in psychology.

Acknowledgments

References

Aharonov, Y., Cohen, E., Grossman, D., and Elizutr, A. C. (2014). "Can a future choice affect a past measurement's outcome?" in *Proceedings of the EPJ Web of Conferences,* Vol. 70 (Les Ulis: EDP Sciences), 00038. doi: 10.1051/epjconf/20147000038

Asano, M., Hashimoto, T., Khrennikov, A., Ohya, M., and Tanaka, Y. (2014). Violation of contextual generalization of the Leggett-Garg inequality for recognition of ambiguous figures. *Phys. Scr.* T163:014006. doi: 10.1088/0031-8949/2014/T163/014006

Atmanspacher, H., and Filk, T. (2010). A proposed test of temporal nonlocality in bistable perception. *J. Math. Psychol.* 54, 314–321. doi: 10.1016/j.jmp.2009.12.001

Atmanspacher, H., and Filk, T. (2012). Contra classical causality: violating temporal Bell inequalities in mental systems. *J. Conscious. Stud.* 19, 95–116.

Atmanspacher, H., and Filk, T. (2013). The Necker–Zeno model for bistable perception. *Top. Cogn. Sci.* 5, 800–817. doi: 10.1111/tops.12044

Bem, D. J., Tressoldi, P. E., Rabeyron, T., and Duggan, M. (2014). *A Meta-Analysis of 90 Experiments on the Anomalous Anticipation of Random Future Events.* Available at: http://ssrn.com/abstract=2423692

Blankenship, R. E., and Engel, G. S. (2010). Long-lived quantum coherence in photosynthetic complexes at physiological temperature. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12766–12770. doi: 10.1073/pnas.1005484107

Busemeyer, J. R., and Bruza, P. D. (2012). *Quantum Models of Cognition and Decision.* Cambridge: Cambridge Press. doi: 10.1017/CBO9780511997716

Busemeyer, J. R., and Wang, Z. (2014). Quantum cognition: key issues and discussion. *Top. Cogn. Sci.* 6, 43–46. doi: 10.1111/tops.12074

Engel, G. S., Calhoun, T. R., Read, E. L., Ahn, T. K., Mancal, T., Cheng, Y. C., et al. (2007). Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature* 446, 782–786. doi: 10.1038/nature05678

Hameroff, S. R., Craddock, T. J., and Tuszynski, J. A. (2014). Quantum effects in the understanding of consciousness. *J. Integr. Neurosci.* 13, 229–252. doi: 10.1142/S0219635214400093

Hameroff, S., and Penrose, R. (2014). Consciousness in the universe: a review of the 'Orch OR'theory". *Phys. Life Rev.* 11, 39–78. doi: 10.1016/j.plrev.2013.08.002

Jumper, C. C., and Scholes, G. D. (2014). Life—Warm, wet and noisy? Comment on "Consciousness in the universe: a review of the 'Orch OR'theory" by Hameroff and Penrose. *Phys. Life Rev.* 11, 85–86. doi: 10.1016/j.plrev.2013.10.008

Khrennikov, A. (2010). *Ubiquitous Quantum Structure from Psychology to Finance.* Berlin: Springer-Verlag. doi: 10.1007/978-3-642-05101-2

Khrennikov, A., Ramelow, S., Ursin, R., Wittmann, B., Kofler, J., and Basieva, I. (2014). On the equivalence of Clauser-Horne and Eberhard inequality based tests. *Phys. Scr.* 2014:014019. doi: 10.1088/0031-8949/2014/T163/014019

Kofler, J., and Brukner, Č. (2013). Condition for macroscopic realism beyond the Leggett-Garg inequalities. *Phys. Rev. A* 87:052115. doi: 10.1103/PhysRevA.87.052115

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual Technical Report A-8.* Gainesville, FL: University of Florida Press.

Leggett, A. J., and Garg, A. (1985). Quantum mechanics versus macroscopic realism: is the flux there when nobody looks? *Phys. Rev. Lett.* 54, 857–860. doi: 10.1103/PhysRevLett.54.857

Maier, M. A., Buechner, V. L., Kuhbandner, C., Pflitsch, M., Fernández-Capo, M., and Gámiz-Sanfeliu, M. (2014). Feeling the future again: retroactive avoidance of negative stimuli. *J. Conscious. Stud.* 21, 121–152.

Mossbridge, J., Tressoldi, P., and Utts, J. (2012). Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis. *Front. Psychol.* 3:390. doi: 10.3389/fpsyg.2012.00390

Olson, S. J., and Ralph, T. C. (2012). Extraction of timelike entanglement from the quantum vacuum. *Phys. Rev. A* 85:012306. doi: 10.1103/PhysRevA.85.012306

Tegmark, M. (2000). Importance of quantum decoherence in brain processes. *Phys. Rev. E* 61:4194. doi: 10.1103/PhysRevE.61.4194

Wang, Z., Solloway, T., Shiffrin, R. M., and Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9431–9436. doi: 10.1073/pnas.1407756111

Supplementary Material

Unitary Transformations in the Quantum Model for Conceptual Conjunctions and Its Application to Data Representation

Tomas Veloz[1,2,3]*[†] and Sylvie Desjardins[1†]

[1] *Department of Mathematics, University of British Columbia, Kelowna, BC, Canada, [2] Center Leo Apostel, Vrije Universiteit Brussel, Brussels, Belgium, [3] Instituto de Filosofía y Ciencias de la Complejidad - IFICC, Ñuñoa, Chile*

Quantum models of concept combinations have been successful in representing various experimental situations that cannot be accommodated by traditional models based on classical probability or fuzzy set theory. In many cases, the focus has been on producing a representation that fits experimental results to validate quantum models. However, these representations are not always consistent with the cognitive modeling principles. Moreover, some important issues related to the representation of concepts such as the dimensionality of the realization space, the uniqueness of solutions, and the compatibility of measurements, have been overlooked. In this paper, we provide a dimensional analysis of the realization space for the two-sector Fock space model for conjunction of concepts focusing on the first and second sectors separately. We then introduce various representation of concepts that arise from the use of unitary operators in the realization space. In these concrete representations, a pair of concepts and their combination are modeled by a single conceptual state, and by a collection of exemplar-dependent operators. Therefore, they are consistent with cognitive modeling principles. This framework not only provides a uniform approach to model an entire data set, but, because all measurement operators are expressed in the same basis, allows us to address the question of compatibility of measurements. In particular, we present evidence that it may be possible to predict non-commutative effects from partial measurements of conceptual combinations.

Keywords: concept combination, quantum cognition, data representation, unitary transformation, conjunction

1. INTRODUCTION

1.1. Concept Combinations in Quantum Cognition

The application of quantum models to cognitive phenomena is an emergent field known as quantum cognition (Aerts, 2009; Pothos and Busemeyer, 2013). One of the areas in quantum cognition that has received much attention is the study of concepts and their combinations (Aerts and Gabora, 2005a,b; Aerts, 2007a,b; Aerts and Sozzo, 2011; Aerts et al., 2013). In a general setting, a cognitive situation might include multiple concepts forming aggregated structures (Rips, 1995; Fodor, 1998). For example, the concepts "*Fruit*" and "*Vegetable*" can be combined to form a new concept "*Fruit And Vegetable*" (Hampton, 1988a). This example of a concept combination

is built with the connective "And," which is also an operation mathematically defined in logic and probability. The question becomes, is it possible to apply the mathematical definition of the connective "*And*" to build the structure of "*Fruit And Vegetable*" from the structures of "*Fruit*" and "*Vegetable*"? Cognitive scientists have performed several experiments measuring various semantic estimations including typicality, membership, and similarity of concept combinations built with connectives such as "*And*," and "*Not*" (Hampton, 1997a, 1988a,b), and adjective-noun compounds such as "*Red Apple*" (Medin and Shoben, 1988; Medin, 1989; Kamp and Partee, 1995). The evidence collected during two decades of research suggests that it might not be possible to represent all the experimental data for concept combinations using the mathematical structures of fuzzy logic or probability theory. Quantum cognition proposes an alternative approach.

While traditional models based on classical logic, probability, or fuzzy set theory have failed to properly account for cognitive phenomena exhibiting non-classical probabilistic features, quantum models have consistently provided a framework that easily encompasses these and other so-called cognitive biases (Gilovich et al., 2002; Busemeyer et al., 2011) or paradoxical phenomena (Aerts et al., 2011a). Quantum inspired models have been successfully developed in the areas of decision making (Aerts et al., 2011b, 2012b; Busemeyer et al., 2011; Busemeyer and Bruza, 2012), psychology of categorization (Aerts and Aerts, 1995; Blutner et al., 2013; Sozzo, 2014), human memory (Bruza and Cole, 2005; Bruza et al., 2009, 2012), and finances (Khrennikov, 2009; Haven and Khrennikov, 2013). In this paper we will focus on the phenomena of concept conjunction. However, since our analysis and methodology is based on pure mathematical notions of the quantum mechanical framework, the results presented in this paper can be extended to other concept combinations (Veloz, 2015).

Aerts (2009) formally states the conditions that characterize the existence of a classical probability model for concept conjunction:

Definition 1. *Let $\mu(A)$, $\mu(B)$, and $\mu(AB)$ be the membership weights of an exemplar p with respect to a pair of concepts $\mathcal{A}$ and $\mathcal{B}$ and their conjunction $\mathcal{AB}$. We say that these membership weights are classical conjunction data if there exists a Kolmogorovian probability space $(\Omega, \sigma(\Omega), P)$, and events $E_A, E_B \in \sigma(\Omega)$ such that*

$$\begin{aligned} P(E_A) &= \mu(A), \\ P(E_B) &= \mu(B), \\ P(E_A \cap E_B) &= \mu(AB). \end{aligned} \quad (1)$$

Classical conjunction data characterizes the membership values of the conjunction of concepts that can be modeled in a classical probabilistic framework. It is therefore important to characterize the notion of classical conjunction data in terms of the membership weights.

Corollary 1. *The membership weights $\mu(A)$, $\mu(B)$, and $\mu(AB)$ of an exemplar p with respect to concepts $\mathcal{A}$, $\mathcal{B}$, and their conjunction*

$\mathcal{AB}$ *are classical conjunction data if and only if*

$$0 \leq \mu(AB) \leq \mu(A), \quad (2)$$
$$0 \leq \mu(AB) \leq \mu(B), \quad (3)$$
$$0 \leq \mu(A) + \mu(B) - \mu(AB) \leq 1. \quad (4)$$

A large body of experimental evidence and a considerable amount of data analysis indicate that the membership of exemplars with respect to concept combinations does not form classical conjunction data (Fodor and Lepore, 1996; Hampton, 1997a,b; Aerts and Gabora, 2005a,b). Namely, the membership with respect to the conjunction of concepts is generally larger than the membership of one of the former concepts, and thus violates either conditions (2) or (3). This phenomenon is called single overextension. When conditions (2) and (3) are violated simultaneously, it is called double overextension. The violation of condition (4) is called the Kolmogorovian factor violation. We refer to (Pitowsky, 1989; Aerts, 2009) for an explanation of this phenomenon.

In Supplementary Table 1, we show two cases reported in Hampton (1988b). In the first case, the membership weight $\mu_1(AB)$ of the item p_1 ="coffee table" with respect to the conjunction $\mathcal{A}_1 \mathcal{B}_1$ ="*Furniture And Household Appliances*" is single overextended with respect to the membership weights $\mu_1(A)$ and $\mu_1(B)$ of concepts $\mathcal{A}_1$ ="*Furniture*," and $\mathcal{B}_1$ ="*Household Appliances*," respectively. In the second case, membership weight $\mu_2(AB)$ of the item p_2 ="tree house" with respect to the conjunction $\mathcal{A}_2 \mathcal{B}_2$ ="*Building And Dwelling*" is doubly overextended with respect to the membership weigths $\mu_2(A)$ and $\mu_2(B)$ of the concepts $\mathcal{A}_2$ ="*Building*," and $\mathcal{B}_2$ ="*Dwelling*," respectively.

The phenomenon of overextension has also been demonstrated not only for membership estimations, but also in typicality (Smith and Osherson, 1981; Hampton, 1996; Storms et al., 1998), property relevance (Fodor and Lepore, 1996; Hampton, 1997a,b; Aerts and Gabora, 2005a,b), and probability estimations (Tversky and Kahneman, 1983; Moro, 2009).

1.2. The Quantum Approach to Concept Combination

The quantum approach to concepts introduces two fundamental assumptions that depart from classical approaches:

A1 Concepts are not represented by a set of instances. Instead, a concept is assumed to exist in a state. A Hilbert space $\mathcal{H}$ is introduced, and a unit vector $|\psi\rangle \in \mathcal{H}$ represents the state of the concept.
A2 Semantic estimations are not functions over the set of instances. Instead, a semantic estimation is a measurement operator, $\mathbf{M} : \mathcal{H} \to \mathcal{H}$, that projects onto a subspace of $\mathcal{H}$.

Concepts $\mathcal{A}$ and $\mathcal{B}$ are represented by the states $|A\rangle$ and $|B\rangle$, respectively. When we consider the conjunction $\mathcal{AB}$ of these two concepts, there are two different ways to combine the concepts (Aerts, 2009). The first considers the conjunction of concepts from an intuitive perspective in the sense that the connective *And* does not play a logical role in the combination

$\mathcal{AB}$; instead the conjunction $\mathcal{AB}$ is viewed as an emergent entity. In particular, the quantum model assumes that the state of the combined concept $|AB\rangle \in \mathcal{H}$ is given by a superposition of the states of concepts $\mathcal{A}$ and $\mathcal{B}$ as follows:

$$|AB\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle). \tag{5}$$

The second way considers the conjunction of concepts from a logical perspective, in the sense that *And* does play a logical role in the combination $\mathcal{AB}$. In particular, the quantum model assumes that the state of the combined concept $|C\rangle$ is modeled in the tensor product space $\mathcal{H} \otimes \mathcal{H}$, where each space in the product captures the representation of the concepts in the combination, while the entire space represents the conjunction. The two quantum models of concept combination are presented in Supplementary Material. These two modes can be unified in a mathematical framework developed in quantum mechanics called Fock space (Aerts, 2007a, 2009).

A Fock space is a direct sum of tensor products of Hilbert spaces, where each space in the sum represents the state space of a system having different numbers of particles (Meyer, 1995). For the case of concepts, we model the state of the combination of two concepts in the two-sector Fock space:

$$\mathcal{F} = \mathcal{H} \oplus (\mathcal{H} \otimes \mathcal{H}). \tag{6}$$

The first space, $\mathcal{H}$, also called the first sector, represents the concept combination as an emergent entity. The second space, $\mathcal{H} \otimes \mathcal{H}$, called the second sector, represents the concept combination as a logical entity. The state of the combined concept in the two-sector Fock space is hence a superposition of the two modes of combination.

For example, when $|C\rangle = |A\rangle \otimes |B\rangle$, the state $|\psi\rangle$ of the concept combination is

$$|\psi\rangle = \frac{n e^{\mathbf{i}\theta_1}}{\sqrt{2}}(|A\rangle + |B\rangle) + \sqrt{1 - n^2} e^{\mathbf{i}\theta_2} |A\rangle \otimes |B\rangle, \tag{7}$$

and the membership formula is given by

$$\mu(AB) = n^2 \left(\frac{\mu(A) + \mu(B)}{2} + \Re(\langle A|M|B\rangle) \right) + \sqrt{1 - n^2} \mu(A)\mu(B), \tag{8}$$

for $0 \leq n \leq 1$.

When $n = 1$, the membership weight $\mu(AB)$ corresponds to the sum of the average of $\mu(A)$ and $\mu(B)$, plus an interference term $\Re(\langle A|M|B\rangle)$ bounded by

$$-\sqrt{\mu(A)\mu(B)} \leq \Re(\langle A|M|B\rangle) \leq \sqrt{\mu(A)\mu(B)}.$$

In the absence of interference, i.e., when $\Re(\langle A|M|B\rangle) = 0$, the membership weight is simply the average of the former membership weights. This particular case, which has been shown to provide a good first approximation to exemplars of conceptual conjunction (Aerts et al., 2012a), is overextended, and therefore non-classical. When $n = 0$, the membership weight corresponds

to the product $\mu(A)\mu(B)$, which is equivalent to the probability of two joint classical events that are independent. When $0 < n < 1$, the state of the concept is in the superposition of the two modes of combination.

Finally, the membership operator for a certain exemplar with respect to the conjunction of two concepts is given by

$$\mathbf{M}^{\mathcal{F}} = \mathbf{M} \oplus (\mathbf{M} \otimes \mathbf{M}), \tag{9}$$

where $\mathbf{M}$ is the operator that measures membership of the exemplar in the first sector, and $\mathbf{M} \otimes \mathbf{M}$ measures the membership of the exemplar with respect to the two concepts simultaneously in the second sector.

In addition to providing a suitable mathematical framework for cognitive models, quantum cognition also offers a different perspective on cognitive phenomena: uncertainty is described by means of superposed states (Aerts et al., 2011b), non-logical coherence involves interference (Aerts, 2009), order effects are revealed by incompatible measurements (Wang and Busemeyer, 2013), and certain "verb-noun" conceptual combinations mimic the structure of physically entangled particles (Aerts and Sozzo, 2014).

1.3. The Representation of Data

One of the reasons why quantum models of concept combinations have not been widely used is that the issue of data representation has been overlooked. Scholars have studied the capacity of quantum models to fit semantic estimations of concept combinations, and have presented concrete representations of the different estimations to validate the models (Aerts, 2007a,b, 2009; Aerts et al., 2012a; Sozzo, 2014); these concrete representations, however, model the data in an exemplar-based fashion, where one operator is used for all exemplars, but the conceptual state varies with exemplars.

For example, Aerts (2009) builds a quantum model in the Hilbert space $\mathbb{C}^3$ to consider the exemplars "filing cabinet" and "heated waterbed" with respect to concepts $\mathcal{A}$ ="*Furniture*, $\mathcal{B}$ ="*Household Appliances*," and their conjunction $\mathcal{AB}$. For the first exemplar, we have $\mu(A) = 0.97$, $\mu(B) = 0.31$, and $\mu(AB) = 0.53$. This case is represented by the vectors

$$\begin{aligned} |A\rangle &= (-0.57 + 0.40\mathbf{i}, 0.29 - 0.63\mathbf{i}, 0.13 + 0.11\mathbf{i}), \\ |B\rangle &= (0.39, 0.39, 0.83). \end{aligned} \tag{10}$$

For the second exemplar, $\mu(A) = 1$, $\mu(B) = 0.49$, and $\mu(AB) = 0.78$, and the state vectors are given by

$$\begin{aligned} |A\rangle &= (0.71, 0.71, 0), \\ |B\rangle &= (0.49, 0.49, 0.71). \end{aligned} \tag{11}$$

In both cases $\mathbf{M}$ is defined by the projection operator

$$\mathbf{M}(x, y, z) \rightarrow (x, 0, 0). \tag{12}$$

Such concrete representations are useful to validate models, but unwieldy if one seeks to build a model that can be used for studying and comparing large amounts of data. Because the

state is independent of the exemplar, it must remain the same for all measurements. But if we require the state representing the concept to remain fixed, then the number of measurement operators is restricted by the dimension of the Hilbert space $\mathcal{H}$. In fact, because the membership operator is usually represented by the identity projector restricted to a smaller subspace, and the identity operator of the entire space and the null operator entail trivial measurements, the number of projectors available to represent membership measurements is restricted to $n - 1$, for $n = \dim(\mathcal{H})$. This implies that, if we consider n or more exemplars, then some exemplars will not have a unique membership operator. These issues become crucial in real-world situations involving concepts that entail thousands of exemplars (Tenenbaum et al., 2011).

In Section 2, we take a close look at the concrete representations of quantum models on each sector of the Fock space to identify the minimal dimensionality required to reach the modeling capacity of each of the sectors. In Section 3, we introduce the notion of unitary transformation for the first and second sectors of the Fock space separately, and propose concrete representations for concepts in these two models that require a single conceptual state, and a collection of exemplar-dependent operators. In Section 4, we use these representations to advance a conjecture concerning compatibility of measurements.

2. DIMENSIONALITY ANALYSIS OF THE TWO-SECTOR FOCK SPACE MODEL

In what follows we determine the dimension of $\mathcal{H}$ required to model concept combinations in the first and second sectors of the two-sector Fock space model. To explore this question, we assume $\mathcal{H} = \mathbb{C}^n$ equipped with the standard inner product, and analyze how n relates to the representation of concepts.

2.1. First Sector Dimension Analysis

The Hilbert space model for concept conjunction requires two vectors, $|A\rangle, |B\rangle \in \mathcal{H}$, and an orthogonal projector, $\mathbf{M} : \mathcal{H} \to \mathcal{H}$, such that

$$\langle A|A\rangle = \langle B|B\rangle = 1, \tag{13}$$

$$\langle A|B\rangle = 0, \tag{14}$$

$$\langle A|\mathbf{M}|A\rangle = \mu(A), \ \langle B|\mathbf{M}|B\rangle = \mu(B), \tag{15}$$

$$\mu(AB) = \frac{1}{2}(\mu(A) + \mu(B)) + \Re(\langle A|\mathbf{M}|B\rangle). \tag{16}$$

The next theorem shows that $n = 3$ is sufficient to build a model that satisfies conditions (13–16).

Theorem 1. *Let $\mu(A)$, $\mu(B)$, and $\mu(AB)$ denote the membership of an exemplar with respect to concepts $\mathcal{A}$, $\mathcal{B}$, and their conjunction $\mathcal{AB}$. The membership weights are compatible with a complex Hilbert space model $\mathcal{H} = \mathbb{C}^3$ if and only if*

$$\mu(AB) \in [\text{ave}(AB) - \text{dev}(AB), \text{ave}(AB) + \text{dev}(AB)], \tag{17}$$

where

$$\text{ave}(AB) = \frac{1}{2}(\mu(A) + \mu(B)), \ and$$
$$\text{dev}(AB) = \sqrt{\min(\mu(A)\mu(B), (1 - \mu(A))(1 - \mu(B))}. \tag{18}$$

Proof. We derive Equation (17) by applying conditions (13–16). First, if $\mathbf{M}$ is a zero- or three-dimensional projector, then

$$\mu(A) = \mu(B) = \mu(AB) = 0, \ \text{or}$$
$$\mu(A) = \mu(B) = \mu(AB) = 1, \tag{19}$$

respectively. Thus, Equation (17) holds, and Equations (13–16) are satisfied by choosing $|A\rangle$ and $|B\rangle$ to be any two mutually orthogonal unit vectors.

Next, we consider the cases where $\mathbf{M}$ is either a one- or two-dimensional projector. We apply conditions (13–16) to vectors $|A\rangle$ and $|B\rangle$ in these two cases separately, and combine the results to obtain (Equation 17).

If $\mathbf{M}$ is a one-dimensional projector, then without loss of generality, we can choose

$$\mathbf{M}(x, y, z) \to (x, 0, 0), \ \text{and}$$

$$|A\rangle = (a_1 e^{i\alpha_1}, a_2 e^{i\alpha_2}, a_3 e^{i\alpha_3}),$$
$$|B\rangle = (b_1 e^{i\beta_1}, b_2 e^{i\beta_2}, b_3 e^{i\beta_3}). \tag{20}$$

Note that conditions (13) and (15) are satisfied by choosing the coefficients in $|A\rangle$ and $|B\rangle$ as follows:

$$a_1 = \sqrt{\mu(A)}; \ a_2 = \sqrt{\lambda}\sqrt{1 - \mu(A)}; \ a_3 = \sqrt{1 - \lambda}\sqrt{1 - \mu(A)},$$
$$b_1 = \sqrt{\mu(B)}; \ b_2 = \sqrt{\kappa}\sqrt{1 - \mu(B)}; \ b_3 = \sqrt{1 - \kappa}\sqrt{1 - \mu(B)}, \tag{21}$$

with $0 \le \lambda \le 1$, and $0 \le \kappa \le 1$. Moreover, Equation (16) implies that $\mu(AB)$ is given by

$$\mu(AB) = \frac{1}{2}(\mu(A) + \mu(B)) + \sqrt{\mu(A)\mu(B)} \cos(\alpha_1 - \beta_1). \tag{22}$$

We then apply condition (14) to obtain

$$-\sqrt{\mu(A)\mu(B)} \cos(\gamma_1)$$
$$= \sqrt{(1 - \mu(A))(1 - \mu(B))} F(\lambda, \kappa, \cos(\gamma_2), \cos(\gamma_3)), \tag{23}$$
$$-\sqrt{\mu(A)\mu(B)} \sin(\gamma_1)$$
$$= \sqrt{(1 - \mu(A))(1 - \mu(B))} F(\lambda, \kappa, \sin(\gamma_2), \sin(\gamma_3)), \tag{24}$$

where

$$F(\lambda, \kappa, f(x), f(y)) = \left(\sqrt{\lambda\kappa}f(x) + \sqrt{(1 - \lambda)(1 - \kappa)}f(y)\right). \tag{25}$$

Since $F(\lambda, \kappa, \cos(\gamma_2), \cos(\gamma_3))$ is a convex combination of $\sqrt{\lambda\kappa}$ and $\sqrt{(1 - \lambda)(1 - \kappa)}$, we have

$$|F(\lambda, \kappa, \cos(\gamma_2), \cos(\gamma_3))| \le |\sqrt{\lambda\kappa}| + |\sqrt{(1 - \lambda)(1 - \kappa)}|. \tag{26}$$

We set

$$\sqrt{\lambda} = \cos(\theta_1), \quad \sqrt{\kappa} = \cos(\theta_2), \qquad (27)$$

for θ_1, θ_2 in $[0, \frac{\pi}{2}]$. Then

$$\sqrt{1-\lambda} = \sin(\theta_1),$$
$$\sqrt{1-\kappa} = \sin(\theta_2). \qquad (28)$$

Substituting Equations (27) and (28) in Equation (26), we obtain

$$|F(\lambda, \kappa, \cos(\gamma_2), \cos(\gamma_3))| \leq |\cos(\theta_1 - \theta_2)| \leq 1. \qquad (29)$$

Then Equation (23) implies that

$$|\sqrt{\mu(A)\mu(B)}\cos(\gamma_1)| \leq \sqrt{(1-\mu(A))(1-\mu(B))}. \qquad (30)$$

Therefore, the interference term is bounded as follows:

$$|\sqrt{\mu(A)\mu(B)}\cos(\gamma_1)| \leq \min(\sqrt{\mu(A)\mu(B)}, \\ \sqrt{(1-\mu(A))(1-\mu(B))}) \qquad (31) \\ = \mathrm{dev}(AB).$$

Next, combining Equations (23) and (24), we obtain

$$\mu(A)\mu(B) = (1-\mu(A))(1-\mu(B))\hat{F}(\lambda, \kappa, \gamma_2, \gamma_3), \qquad (32)$$

where

$$\hat{F}(\lambda, \kappa, \gamma_2, \gamma_3) = F^2(\lambda, \kappa, \cos(\gamma_2), \cos(\gamma_3)) \\ + F^2(\lambda, \kappa, \sin(\gamma_2), \sin(\gamma_3)). \qquad (33)$$

Hence,

$$\mu(A) + \mu(B) = 1 + \mu(A)\mu(B)\left(1 - \frac{1}{\hat{F}(\lambda, \kappa, \gamma_2, \gamma_3)}\right). \qquad (34)$$

We use the parametrization for λ and κ given by Equation (27), and apply Equations (29–33), to obtain

$$0 \leq \hat{F}(\lambda, \kappa, \gamma_2, \gamma_3) \leq \cos(\theta_1 - \theta_2)^2 + \sin(\theta_1 - \theta_2)^2 = 1. \qquad (35)$$

Combining Equations (35) and (34) yields

$$\mu(A) + \mu(B) \leq 1. \qquad (36)$$

Therefore, when $\mathbf{M}$ is a one-dimensional projector, conditions (13–16) imply

$$\mu(AB) \in [\mathrm{ave}(AB) - \mathrm{dev}(AB), \mathrm{ave}(AB) + \mathrm{dev}(AB)], \text{ and} \\ \mu(A) + \mu(B) \leq 1. \qquad (37)$$

Next, consider the case in which $\mathbf{M}$ is a two dimensional projector. Without loss of generality, we can assume

$$\mathbf{M}(x, y, z) \rightarrow (x, y, 0).$$

The requirements Equations (13) and (15) are satisfied by choosing the coefficients in $|A\rangle$, $|B\rangle$ as follows

$$a_1 = \sqrt{\lambda}\sqrt{\mu(A)}; \; a_2 = \sqrt{1-\lambda}\sqrt{\mu(A)} \, ; \, a_3 = \sqrt{1-\mu(A)}, \\ b_1 = \sqrt{\kappa}\sqrt{\mu(B)}; \; b_2 = \sqrt{1-\kappa}\sqrt{\mu(B)} \, ; \, b_3 = \sqrt{1-\mu(B)}, \qquad (38)$$

with $0 \leq \lambda \leq 1$, and $0 \leq \kappa \leq 1$. Moreover, Equation (16) implies that $\mu(AB)$ is given by

$$\mu(AB) = \frac{1}{2}(\mu(A) + \mu(B)) + \sqrt{\mu(A)\mu(B)}F(\lambda, \kappa, \cos(\gamma_1), \cos(\gamma_2)). \qquad (39)$$

We apply condition (14) to obtain

$$\sqrt{\mu(A)\mu(B)}F(\lambda, \kappa, \cos(\gamma_1), \cos(\gamma_2)) \\ = -\sqrt{(1-\mu(A))(1-\mu(B))}\cos(\gamma_3). \qquad (40)$$

Since $F(\lambda, \kappa, \cos(\gamma_1), \cos(\gamma_2)) \leq 1$, Equation (40) implies that

$$|\sqrt{\mu(A)\mu(B)}F(\lambda, \kappa, \cos(\gamma_1), \cos(\gamma_2))| \leq \min(\sqrt{\mu(A)\mu(B)}, \\ \sqrt{(1-\mu(A))(1-\mu(B))}) \\ = \mathrm{dev}(AB). \qquad (41)$$

We repeat the procedure used in the one-dimensional case to obtain

$$\mu(A)\mu(B)\hat{F}(\lambda, \kappa, \gamma_1, \gamma_2) = (1-\mu(A))(1-\mu(B)). \qquad (42)$$

Since $0 \leq \hat{F}(\lambda, \kappa, \gamma_1, \gamma_2) \leq 1$, Equation (42) yields

$$1 \leq \mu(A) + \mu(B). \qquad (43)$$

Therefore, when $\mathbf{M}$ is a two-dimensional projector, conditions (13–16) imply

$$\mu(AB) \in [\mathrm{ave}(AB) - \mathrm{dev}(AB), \mathrm{ave}(AB) + \mathrm{dev}(AB)], \text{ and} \\ 1 \leq \mu(A) + \mu(B). \qquad (44)$$

We complete the proof by merging Equations (37) and (44).

The general case, $\mathcal{H} = \mathbb{C}^n$ for $n > 3$, doesn't provide additional modeling power since the condition given by Equation (17) remains. Also, the case $\mathcal{H} = \mathbb{C}^2$ is more restrictive than the $\mathcal{H} = \mathbb{C}^3$ case. In fact, membership data compatible with conditions (13–16) for $\mathcal{H} = \mathbb{C}^2$ must satisfy $\mu(A) + \mu(B) = 1$ (Veloz, 2015).

2.2. Second Sector Dimension Analysis

The second sector of the two-sector Fock space requires a concept combination state $|C\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n$ and an operator $\mathbf{M} : \mathbb{C}^n \rightarrow \mathbb{C}^n$, such that $|C\rangle$ restricted to the first sector represents the concept $\mathcal{A}$, and $|C\rangle$ restricted to the second sector represents the concept $\mathcal{B}$. However, $|C\rangle$ cannot in general be decomposed as a tensor product of the type $|C_A\rangle \otimes |C_B\rangle$, for $|C_A\rangle, |C_B\rangle \in \mathbb{C}^n$. Therefore, $|C\rangle$ is usually a non-separable state.

To recover the probabilistic structure of the former concepts in the combination, the operators $\mathbf{M} \otimes \mathbf{1}$ and $\mathbf{1} \otimes \mathbf{M}$ are applied to $|C\rangle$ to obtain $\mu(A)$ and $\mu(B)$, respectively. Moreover, since $|C\rangle$ as a whole represents the concept combination $\mathcal{AB}$, then the operator $\mathbf{M} \otimes \mathbf{M}$ is applied to $|C\rangle$ to obtain $\mu(AB)$.

The following definition summarizes how data is represented in the second sector.

Definition 2. *Let* $\mu = \{\mu(A), \mu(B), \mu(AB)\}$ *be a triplet denoting the membership of concepts* $\mathcal{A}$, $\mathcal{B}$, *and their conjunction* $\mathcal{AB}$. *We say that the triplet* μ *admits a representation in* $\mathbb{C}^n \otimes \mathbb{C}^n$ *if there exists a unit vector* $|C\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n$, *and an operator* $\mathbf{M} : \mathbb{C}^n \to \mathbb{C}^n$ *such that*

$$\langle C|\mathbf{M}^A|C\rangle = \langle C|\mathbf{M} \otimes \mathbf{1}|C\rangle = \mu(A), \tag{45}$$

$$\langle C|\mathbf{M}^B|C\rangle = \langle C|\mathbf{1} \otimes \mathbf{M}|C\rangle = \mu(B), \tag{46}$$

$$\langle C|\mathbf{M}^\wedge|C\rangle = \langle C|\mathbf{M} \otimes \mathbf{M}|C\rangle = \mu(AB). \tag{47}$$

Let $\{|i\rangle\}_{i=1}^n$ be the canonical basis of $\mathbb{C}^n$. Without loss of generality, we can take $\mathbf{M}$ to be an orthogonal projector on the subspace of $\mathbb{C}^n$ spanned by the basis elements $|1\rangle, ..., |r\rangle$, with $r < n$. Hence,

$$\mathbf{M}(x_1, ..., x_n) \to (x_1, ..., x_r, 0, ..., 0).$$

Next, let $|C\rangle$ be a unit vector in $\mathbb{C}^n \otimes \mathbb{C}^n$. That is,

$$|C\rangle = \sum_{i=1}^n \sum_{j=1}^n c_{ij} e^{\dot{\mathbb{I}}\gamma_{ij}} |i\rangle \otimes |j\rangle, \tag{48}$$

and

$$\langle C|C\rangle = \sum_{i,j=1}^n c_{ij} e^{\dot{\mathbb{I}}\gamma_{ij}} \langle i| \otimes \langle j| \sum_{k,l=1}^n c_{kl} e^{\dot{\mathbb{I}}\gamma_{kl}} |k\rangle \otimes |l\rangle$$

$$= \sum_{i,j,k,l=1}^n c_{ij} c_{kl} e^{\dot{\mathbb{I}}(-\gamma_{ij}+\gamma_{kl})} \langle i|k\rangle \langle j|l\rangle \tag{49}$$

$$= \sum_{i,j=1}^n c_{ij}^2 = 1.$$

We now prove that the operator $\mathbf{M}$ and the vector $|C\rangle$ above satisfy Equations (45–47) if and only if $\mu(A)$, $\mu(B)$, and $\mu(AB)$ are classical conjunction data.

Theorem 2. *Let* $\mu = \{\mu(A), \mu(B), \mu(AB)\}$ *be a triplet denoting the membership of concepts* $\mathcal{A}$, $\mathcal{B}$, *and their conjunction* $\mathcal{AB}$. *The triplet* μ *is classical conjunction data if and only if it admits a representation in* $\mathbb{C}^n \otimes \mathbb{C}^n$ *with* $n = 2$.

Proof. If μ admits a representation in $\mathbb{C}^2 \otimes \mathbb{C}^2$, there exists a unit vector $|C\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ and an operator $\mathbf{M}$ such that Equations (45–47) are satisfied. If $\mu(A) = \mu(B) = \mu(AB) = 0$ or 1, we can choose $|C\rangle$ to be any unit vector in $\mathbb{C}^2 \otimes \mathbb{C}^2$, and $\mathbf{M}$ to be a zero- or two-dimensional projector, respectively. Otherwise, let $\{|1\rangle, |2\rangle\}$

be the canonical basis for $\mathbb{C}^2$. Without loss of generality, we can define $|C\rangle$ by

$$|C\rangle = c_{11} e^{\dot{\mathbb{I}}\gamma_{11}} |1\rangle \otimes |1\rangle + c_{12} e^{\dot{\mathbb{I}}\gamma_{12}} |1\rangle \otimes |2\rangle + c_{21} e^{\dot{\mathbb{I}}\gamma_{21}} |2\rangle \otimes |1\rangle$$
$$+ c_{22} e^{\dot{\mathbb{I}}\gamma_{22}} |2\rangle \otimes |2\rangle \tag{50}$$

and $\mathbf{M}$ by the one-dimensional projector into the subspace determined by $|1\rangle$. Note that

$$\mu(A) = \langle C|\mathbf{M} \otimes \mathbf{1}|C\rangle = c_{11}^2 + c_{12}^2,$$
$$\mu(B) = \langle C|\mathbf{1} \otimes \mathbf{M}|C\rangle = c_{11}^2 + c_{21}^2, \tag{51}$$
$$\mu(AB) = \langle C|\mathbf{M} \otimes \mathbf{M}|C\rangle = c_{11}^2.$$

Then, clearly $\mu(AB) \leq \mu(A)$, $\mu(AB) \leq \mu(B)$, and since $|C\rangle$ is a unit vector,

$$\mu(A) + \mu(B) - \mu(AB) = c_{11}^2 + c_{12}^2 + c_{21}^2 \leq 1. \tag{52}$$

Therefore, μ is classical conjunction data. The other implication is proven by taking $\mathbf{M}$ to be the same one-dimensional projector, and $|C\rangle$ such that

$$c_{11} = \sqrt{\mu(AB)},$$
$$c_{12} = \sqrt{\mu(A) - \mu(AB)},$$
$$c_{21} = \sqrt{\mu(B) - \mu(AB)},$$
$$c_{22} = \sqrt{1 - \mu(A) - \mu(B) + \mu(AB)},$$

and $\gamma_{ij} = 0$, for $i, j = 1, 2$.

Theorem 2 proves the strict equivalence between classical conjunction data and the model of conjunction built in $\mathbb{C}^2 \otimes \mathbb{C}^2$.

3. UNITARY TRANSFORMATIONS AND DATA REPRESENTATION

We now investigate how multiple exemplars can be concretely represented using a single concept state. To do so, we use unitary transformations to identify a basis of the realization space where multiple exemplars can be represented simultaneously. In this new framework, concrete representations are consistent with the cognitive principles of the quantum model of concepts. Namely, a concept exists in a single state for all exemplars, and the measurement of membership of an exemplar depends on the exemplar to be measured rather than on the concept state.

3.1. Data Representation in the First Sector

The following definition and theorem introduce the notion of data representation in the first sector that is consistent with the cognitive principles of the quantum model of concepts in $\mathbb{C}^3$.

Definition 3. *Let* $\mu = \{\mu_i(A), \mu_i(B), \mu_i(AB)\}_{i=1}^k$ *be a set of experimental data, where* $\mu_i(x)$ *is the semantic estimation of an exemplar* p_i *with respect to concepts* $\mathcal{A}$, $\mathcal{B}$, *and their conjunction* $\mathcal{AB}$. *A representation of* μ *in* $\mathbb{C}^3$ *is defined as a pair of unit vectors* $|A\rangle, |B\rangle \in \mathbb{C}^3$, *and a collection of orthogonal projectors* $\mathbf{M}_i : \mathbb{C}^3 \to \mathbb{C}^3$ *such that conditions (13–16) are satisfied for* $i = 1, ..., k$. *We say* $(|A\rangle, |B\rangle, \{\mathbf{M}_i\}_{i=1}^k)$ *is a representation of* μ *in* $\mathbb{C}^3$.

Theorem 3. *Let* $\mu = \{\mu_i(A), \mu_i(B), \mu_i(AB)\}_{i=1}^{k}$ *be a set of experimental data, where* $\mu_i(x)$ *is the semantic estimation of exemplar* p_i *with respect to concepts* $\mathcal{A}$, $\mathcal{B}$, *and their conjunction* $\mathcal{A}\mathcal{B}$. *The set of data* μ *has a representation in* $\mathbb{C}^3$ *if and only if for all* $i = 1, ..., k$

$$\mu_i(AB) \in [\text{ave}_i(AB) - \text{dev}_i(AB), \text{ave}_i(AB) + \text{dev}_i(AB)]. \quad (53)$$

Proof. Let $|A\rangle = (1, 0, 0)$, $|B\rangle = (0, 1, 0)$, and $|C\rangle = (0, 0, 1)$ be the canonical basis for $\mathbb{C}^3$. We prove that, if Equation (53) is satisfied for each $i = 1, ..., k$ then there exists an orthogonal projector $\mathbf{M}_i$ such that conditions (13–16) are satisfied for $|A\rangle$, $|B\rangle$, and $\mathbf{M}_i$.

Since $\mu_i(A)$, $\mu_i(B)$ and $\mu_i(AB)$ satisfy (Equation 53), by Theorem 1 for each $i \in \{1, ..., k\}$ there exist two vectors,

$$|A_i\rangle = (a_1 e^{\dot{\mathbb{i}}\alpha_1}, a_2 e^{\dot{\mathbb{i}}\alpha_2}, a_3 e^{\dot{\mathbb{i}}\alpha_3}), \ |B_i\rangle = (b_1 e^{\dot{\mathbb{i}}\beta_1}, b_2 e^{\dot{\mathbb{i}}\beta_2}, b_3 e^{\dot{\mathbb{i}}\beta_3}), \quad (54)$$

and an orthogonal projector $\hat{\mathbf{M}}_i$ such that Equations (13–16) are satisfied. Thus, the pair of vectors $|A_i\rangle$ and $|B_i\rangle$, as constructed in the proof of Theorem 1, are orthonormal. We set $|C_i\rangle = |A_i\rangle \times |B_i\rangle$ so that the set $\{|A_i\rangle, |B_i\rangle, |C_i\rangle\}$ forms an orthonormal basis for $\mathbb{C}^3$ for any $i \in \{1, ..., k\}$. Next, we define the operator U_i by

$$U_i = \begin{pmatrix} \langle A_i|A\rangle & \langle A_i|B\rangle & \langle A_i|C\rangle \\ \langle B_i|A\rangle & \langle B_i|B\rangle & \langle B_i|C\rangle \\ \langle C_i|A\rangle & \langle C_i|B\rangle & \langle C_i|C\rangle \end{pmatrix}. \quad (55)$$

U_i is a unitary matrix whose action induces a change from the basis $\{|A_i\rangle, |B_i\rangle, |C_i\rangle\}$ to the basis $\{|A\rangle, |B\rangle, |C\rangle\}$. Note that $U_i|A_i\rangle = |A\rangle$, $U_i|B_i\rangle = |B\rangle$, and $U_i|C_i\rangle = |C\rangle$.

We can also use the operator U_i to represent $\hat{\mathbf{M}}_i$ in the canonical basis $\{|A\rangle, |B\rangle, |C\rangle\}$ as follows:

$$\mathbf{M}_i = U_i \hat{\mathbf{M}}_i U_i^{-1}. \quad (56)$$

We use the fact that $\mathbf{I} = U_i^{-1} U_i = U_i U_i^{-1}$ to show that the remaining conditions are satisfied. That is, for each $i = 1, ..., k$,

$$\mu_i(A) = \langle A_i|\hat{\mathbf{M}}_i|A_i\rangle = \langle A_i U_i^{-1}|U_i \hat{\mathbf{M}}_i U_i^{-1}|U_i A_i\rangle = \langle A|\mathbf{M}_i|A\rangle,$$
$$\mu_i(B) = \langle B_i|\hat{\mathbf{M}}_i|B_i\rangle = \langle B_i U_i^{-1}|U_i \hat{\mathbf{M}}_i U_i^{-1}|U_i B_i\rangle = \langle B|\mathbf{M}_i|B\rangle, \quad (57)$$

and

$$\begin{aligned} \mu_i(AB) &= \frac{1}{2}(\mu(A) + \mu(B)) + \Re(\langle A_i|\hat{\mathbf{M}}_i|B_i\rangle) \\ &= \frac{1}{2}(\mu(A) + \mu(B)) + \Re(\langle A_i U_i^{-1}|U_i \hat{\mathbf{M}}_i U_i^{-1}|U_i A_i\rangle) \\ &= \frac{1}{2}(\mu(A) + \mu(B)) + \Re(\langle A|\mathbf{M}_i|B\rangle). \end{aligned} \quad (58)$$

Theorem 3 provides a data representation in terms of a single pair of vectors $|A\rangle$ and $|B\rangle$, and a set of projectors $\mathbf{M}_i$, for $i = 1, ..., k$, corresponding to the membership operator for each exemplar. Since the unitary transformations preserve the inner product

between vectors and operators, the values of the membership estimations $\mu_i(A)$, $\mu_i(B)$, and $\mu_i(AB)$ are preserved.

Consider for example the exemplars $p =$"filing cabinet" and $q =$"heated waterbed" mentioned in Section 1.3. These can now be represented by the states $|A\rangle = (1, 0, 0)$, $|B\rangle = (0, 1, 0)$ and the following measurement operators

$$\begin{aligned} \mathbf{M}_p &= \begin{pmatrix} 0.97 & -0.11 + 0.09\dot{\mathbb{i}} & 0.09 + 0.01\dot{\mathbb{i}} \\ -0.11 - 0.09\dot{\mathbb{i}} & 0.31 & 0.28 + 0.34\dot{\mathbb{i}} \\ 0.09 - 0.01\dot{\mathbb{i}} & 0.28 - 0.34\dot{\mathbb{i}} & 0.72 \end{pmatrix}, \\ \mathbf{M}_q &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.49 & 0.499 \\ 0 & 0.499 & 0.51 \end{pmatrix}. \end{aligned} \quad (59)$$

From a geometric perspective, the operators $\mathbf{M}_p$ and $\mathbf{M}_q$ correspond to rotations of the one-dimensional projector $\mathbf{M}(x, y, z) \to (x, 0, 0)$ in $\mathbb{C}^3$.

3.2. Data Representation in the Second Sector

We now apply unitary transformations in the concrete representations of the tensor product model in $\mathbb{C}^n \otimes \mathbb{C}^n$. We first define different types of representations for multiple exemplars, and then provide explicit representation theorems for the cases $n = 2$ and 3.

Definition 4. *A zero-type representation of* $\mu_{i=1}^{k}$ *on the tensor product space* $\mathbb{C}^n \otimes \mathbb{C}^n$ *is a unit vector* $|C\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n$, *and a collection of orthogonal projectors* $\{\mathbf{M}_i^A, \mathbf{M}_i^B\}_{i=1}^{k}$ *from* $\mathbb{C}^n \otimes \mathbb{C}^n$ *to* $\mathbb{C}^n \otimes \mathbb{C}^n$, *such that conditions (47)–(49) are satisfied with* $\mathbf{M}_i^\wedge = \mathbf{M}_i^A \mathbf{M}_i^B$, *for* $i = 1, ..., k$. *We say* $(|C\rangle, \{\mathbf{M}_i^A, \mathbf{M}_i^B\}_{i=1}^{k})$ *is a zero-type representation of* $\mu_{i=1}^{k}$ *in* $\mathbb{C}^n \otimes \mathbb{C}^n$.

The zero-type representation is, mathematically speaking, the most general representation in the tensor product model that is consistent with the modeling principles of quantum cognition because it assumes a single concept state $|C\rangle$, and a collection of measurements that represent the membership weight estimations. However, this representation cannot be appropriately interpreted because $\mathbf{M}_i^A$ and $\mathbf{M}_i^B$ can be entangled measurements, for $i = 1, ..., k$.

A more reasonable representation of data assumes that $\mathbf{M}_i^A = \mathbf{M}_i \otimes \mathbb{1}$, and $\mathbf{M}_i^B = \mathbb{1} \otimes \mathbf{M}_i$, for $i = 1, ..., k$. Therefore, these operators are not entangled because they act on different sides of $\mathbb{C}^3 \otimes \mathbb{C}^3$.

Definition 5. *A first-type representation of* $\mu_{i=1}^{k}$ *on the tensor product space* $\mathbb{C}^n \otimes \mathbb{C}^n$ *is a unit vector* $|C\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n$, *and a collection of orthogonal projectors* $\mathbf{M}_i$ *from* $\mathbb{C}^n$ *to* $\mathbb{C}^n$, *for* $i = 1, ..., k$, *such that* $(|C\rangle, \{\mathbf{M}_i \otimes \mathbb{1}, \mathbb{1} \otimes \mathbf{M}_i\}_{i=1}^{k})$ *is a zero-type representation of* $\mu_{i=1}^{k}$ *in* $\mathbb{C}^n \otimes \mathbb{C}^n$.

The first-type representation is a direct extension of the representation of individual exemplars in Definition 2, and thus it is interpreted according to such representation: The state $|C\rangle$ describes the situation of having two concepts and their combination, and $\mathbf{M}_i$ represents the semantic estimation of exemplar p_i, $i = 1, ..., k$.

The zero- and first-type representations require different conditions to model a collection of exemplars for a pair of concepts and their conjunction. While the first-type corresponds to the natural way to represent a pair of systems in quantum physics, and thus is the natural way to define a representation in the tensor product model for concepts, the zero-type provides a more general way to build concrete representations because it does not impose a product structure on the concept state or on the membership operators for the exemplars.

In fact, from Definitions 4–5 it is trivial to deduce that a first-type representation is also a zero-type representation.

The following theorem characterizes the cases when a set of data has a zero-type representation in $\mathbb{C}^2 \otimes \mathbb{C}^2$.

Theorem 4. *The set of data $\mu_{i=1}^k$ has a zero-type representation in $\mathbb{C}^2 \otimes \mathbb{C}^2$ if and only if μ_i is classical conjunction data for $i = 1, ..., k$.*

Proof. For each $i = 1, ..., k$, we use the construction in the proof of Theorem 2 to obtain a tensor $|\tilde{C}_i\rangle$ and a one-dimensional projector $\tilde{\mathbf{M}}$ such that $\tilde{\mathbf{M}}_i^A = \tilde{\mathbf{M}} \otimes \mathbb{1}$, $\tilde{\mathbf{M}}_i^B = \mathbb{1} \otimes \tilde{\mathbf{M}}$, and $\tilde{\mathbf{M}}_i^\wedge = \tilde{\mathbf{M}} \otimes \tilde{\mathbf{M}}$. This gives the tensor product representation of μ_i. Next, we use unitary transformations to change this representation so that $|\tilde{C}_i\rangle$ is a vector in the canonical basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$. To facilitate the notation, we will make use of the isomorphism $\mathbb{I}$ between $\mathbb{C}^2 \otimes \mathbb{C}^2$ and $\mathbb{C}^4$. Let

$$
\begin{aligned}
(1, 0, 0, 0) &= |e_1\rangle, \\
(0, 1, 0, 0) &= |e_2\rangle, \\
(0, 0, 1, 0) &= |e_3\rangle, \\
(0, 0, 0, 1) &= |e_4\rangle.
\end{aligned}
\tag{60}
$$

We define

$$
\begin{aligned}
\mathbb{I}(|1\rangle \otimes |1\rangle) &= |e_1\rangle, \\
\mathbb{I}(|1\rangle \otimes |2\rangle) &= |e_2\rangle, \\
\mathbb{I}(|2\rangle \otimes |1\rangle) &= |e_3\rangle, \\
\mathbb{I}(|2\rangle \otimes |2\rangle) &= |e_4\rangle.
\end{aligned}
\tag{61}
$$

The isomorphism $\mathbb{I}$ allows us to represent $|\tilde{C}_i\rangle$ by a vector $|C_i\rangle$ in $\mathbb{C}^4$.

We can prove the theorem by building a unitary transformation that takes $|C_i\rangle$ to one of the canonical basis vectors of $\mathbb{C}^4$, and use this transformation to represent the operators $\tilde{\mathbf{M}}^A$, $\tilde{\mathbf{M}}^B$, and $\tilde{\mathbf{M}}^\wedge$ by the operators $\mathbf{M}^A$, $\mathbf{M}^B$, and $\mathbf{M}^\wedge$ in $\mathbb{C}^4$. Next, we apply the the inverse isomorphism $\mathbb{I}^{-1}$ to map these new representations to $\mathbb{C}^2 \otimes \mathbb{C}^2$.

Let $|D_i\rangle, |E_i\rangle, |F_i\rangle$ be three vectors in $\mathbb{C}^4$ such that

$$
\begin{aligned}
\langle D_i | D_i \rangle &= \langle E_i | E_i \rangle = \langle F_i | F_i \rangle = 1, \\
\langle C_i | D_i \rangle &= \langle C_i | E_i \rangle = \langle C_i | F_i \rangle = 0, \\
\langle D_i | E_i \rangle &= \langle D_i | F_i \rangle = \langle E_i | F_i \rangle = 0.
\end{aligned}
\tag{62}
$$

The vectors $|C_i\rangle, |D_i\rangle, |E_i\rangle$, and $|F_i\rangle$ form an orthonormal basis for $\mathbb{C}^4$.

Let

$$
U_i = \begin{pmatrix}
\langle C_i | e_1 \rangle & \langle C_i | e_2 \rangle & \langle C_i | e_3 \rangle & \langle C_i | e_4 \rangle \\
\langle D_i | e_1 \rangle & \langle D_i | e_2 \rangle & \langle D_i | e_3 \rangle & \langle D_i | e_4 \rangle \\
\langle E_i | e_1 \rangle & \langle E_i | e_2 \rangle & \langle E_i | e_3 \rangle & \langle E_i | e_4 \rangle \\
\langle F_i | e_1 \rangle & \langle F_i | e_2 \rangle & \langle F_i | e_3 \rangle & \langle F_i | e_4 \rangle
\end{pmatrix}.
\tag{63}
$$

Note that U_i is a unitary matrix whose action induces a change from the basis $\{|C_i\rangle, |D_i\rangle, |E_i\rangle, |F_i\rangle\}$ to the basis $\{|e_j\rangle\}_{j=1}^4$. In fact,

$$
U_i |C_i\rangle = |e_1\rangle, \; U_i |D_i\rangle = |e_2\rangle, \; U_i |E_i\rangle = |e_3\rangle, \text{ and } U_i |F_i\rangle = |e_4\rangle.
$$

The operator U_i can now be used to change the basis in which $\mathbf{M}_i^A$, $\mathbf{M}_i^B$, and $\mathbf{M}_i^\wedge$ are represented, to the basis $\{|e_j\rangle\}_{j=1}^4$:

$$
\begin{aligned}
\bar{\mathbf{M}}_i^A &= U_i \mathbf{M}_i^A U_i^{-1}, \\
\bar{\mathbf{M}}_i^B &= U_i \mathbf{M}_i^B U_i^{-1}, \\
\bar{\mathbf{M}}_i^\wedge &= U_i \mathbf{M}_i^\wedge U_i^{-1}.
\end{aligned}
\tag{64}
$$

Since $\mathbb{1} = U_i^{-1} U_i = U_i U_i^{-1}$, we obtain

$$
\begin{aligned}
\mu_i(A) &= \langle C_i | \mathbf{M}_i^A | C_i \rangle = \langle C_i U_i^{-1} | U_i \mathbf{M}_i^A U_i^{-1} | U_i C_i \rangle = \langle e_1 | \bar{\mathbf{M}}_i^A | e_1 \rangle, \\
\mu_i(B) &= \langle C_i | \mathbf{M}_i^B | C_i \rangle = \langle C_i U_i^{-1} | U_i \mathbf{M}_i^B U_i^{-1} | U_i C_i \rangle = \langle e_1 | \bar{\mathbf{M}}_i^B | e_1 \rangle, \\
\mu_i(AB) &= \langle C_i | \mathbf{M}_i^\wedge | C_i \rangle = \langle C_i U_i^{-1} | U_i \mathbf{M}_i^\wedge U_i^{-1} | U_i C_i \rangle = \langle e_1 | \bar{\mathbf{M}}_i^\wedge | e_1 \rangle.
\end{aligned}
\tag{65}
$$

We then use the inverse isomorphism $\mathbb{I}^{-1}$ to obtain a zero-type representation in $\mathbb{C}^2 \otimes \mathbb{C}^2$:

$$
\begin{aligned}
|C\rangle &= \mathbb{I}^-1(|e_1\rangle) = |1\rangle \otimes |1\rangle, \\
\tilde{\mathbf{M}}_i^A &= \mathbb{I}^{-1} \bar{\mathbf{M}}_i^A \mathbb{I}, \\
\tilde{\mathbf{M}}_i^B &= \mathbb{I}^{-1} \bar{\mathbf{M}}_i^B \mathbb{I}, \\
\tilde{\mathbf{M}}_i^\wedge &= \mathbb{I}^{-1} \bar{\mathbf{M}}_i^\wedge \mathbb{I}.
\end{aligned}
\tag{66}
$$

We have constructed a zero-type representation $(|1\rangle \otimes |1\rangle, \{\mathbf{M}_i^A, \mathbf{M}_i^B\}_{i=1}^k)$ from a collection of representations $(|C_i\rangle, \mathbf{M})$ for the exemplars p_i with $\mathbf{M}(x, y) \rightarrow (x, 0)$ obtained from Theorem 2.

In the construction of Theorem 4, note that when Equation (66) entails operators $\mathbf{M}_i^A$ and $\mathbf{M}_i^B$ that are of the form $\mathbf{M}_A^i = \check{\mathbf{M}}_i \otimes \mathbb{1}$ and $\mathbf{M}_B^i = \mathbb{1} \otimes \check{\mathbf{M}}_i$, then the representation is also of the first-type.

Stating the necessary and sufficient conditions required for a set of data to have first-type representation is out of the scope of this paper. However, we now introduce another type of representation that is mathematically simpler, and can be used to obtain sufficient conditions for a first-type representation.

Definition 6. *A second-type representation of $\mu_{i=1}^k$ on the tensor product space $\mathbb{C}^n \otimes \mathbb{C}^n$ is a pair of unit vectors $|A\rangle, B\rangle \in \mathbb{C}^n$, and a collection of orthogonal projectors $\mathbf{M}_i$ from $\mathbb{C}^n$ to $\mathbb{C}^n$, for $i = 1, ..., k$, such that $(|A\rangle \otimes |B\rangle, \{\mathbf{M}_i \otimes \mathbb{1}, \mathbb{1} \otimes \mathbf{M}_i\}_{i=1}^k)$ is a zero-type representation of $\mu_{i=1}^k$ in $\mathbb{C}^n \otimes \mathbb{C}^n$.*

The second-type is a mathematical simplification of the first-type representation that assumes $|C\rangle$ to be a product state.

Lemma 1. *The set of data* $\mu_{i=1}^k$ *has a second-type representation in* $\mathbb{C}^2 \otimes \mathbb{C}^2$ *if and only if for each* $i = 1, ..., k$ *there exist* $|A_i\rangle$, $|B_i\rangle$, $\check{\mathbf{M}}_i^A$, *and* $\check{\mathbf{M}}_i^B$ *such that Equations (45–47) are satisfied.*

Proof. Let $U_i(A)$, $U_i(B)$: $\mathbb{C}^2 \rightarrow \mathbb{C}^2$ be the unitary transformations that map $|A_i\rangle$ to $|1\rangle$ and $|B_i\rangle$ to $|1\rangle$ respectively, for $i = 1, ..., k$. Then, it is straightforward to show that $(|1\rangle \otimes |1\rangle, \{\mathbf{M}_i^A \otimes \mathbb{1}, \mathbb{1} \otimes \mathbf{M}_i^B\}_{i=1}^k)$ is a second-type representation of $\mu_{i=1}^k$ with

$$\mathbf{M}_i^A = U_i(A)^{-1}\check{\mathbf{M}}_i^A U_i(A),$$
$$\mathbf{M}_i^B = U_i(B)^{-1}\check{\mathbf{M}}_i^B U_i(B). \tag{67}$$

Theorem 4 and Lemma 1 characterize the sets of data that have a zero- and second-type representations. Since the first-type representation is less general than the zero-type representation, but more general than the second-type representation, these results can be applied to obtain an upper and lower bound for the number of exemplars that have a first-type representation in a given set of data.

Note that Theorem 4 is built in $\mathbb{C}^2 \otimes \mathbb{C}^2$. We now extend our results to $\mathbb{C}^3 \otimes \mathbb{C}^3$ so they become compatible with the representation analysis developed in Section 3.1 for a Hilbert space model in $\mathbb{C}^3$. The next corollary extends the proof of Theorem 4 to the space $\mathbb{C}^3 \otimes \mathbb{C}^3$.

Corollary 2. *If the set of data* $\mu_{i=1}^k$ *has a zero-type representation in* $\mathbb{C}^2 \otimes \mathbb{C}^2$, *then* $\mu_{i=1}^k$ *has a zero-type representation in* $\mathbb{C}^3 \otimes \mathbb{C}^3$.

Proof. Let $(|C\rangle, \{\mathbf{M}_i^A, \mathbf{M}_i^B\}_{i=1}^k)$ be a zero-type representation of $\mu_{i=1}^k$ in $\mathbb{C}^2 \otimes \mathbb{C}^2$. We can create a vector

$$|C^*\rangle = \sum_{i,j=1}^3 c_{ij}^* |i\rangle \otimes |j\rangle \tag{68}$$

such that it is the trivial embedding of

$$|C\rangle = \sum_{i,j=1}^2 c_{ij} |i\rangle \otimes |j\rangle \tag{69}$$

in $\mathbb{C}^3 \otimes \mathbb{C}^3$ by choosing

$$c_{ij}^* = \begin{cases} c_{ij} & i,j \in \{1,2\} \\ 0 & \text{else.} \end{cases} \tag{70}$$

Similarly, we can also create operators $\mathbf{M}_i^{A*}$ and $\mathbf{M}_i^{B*}$ by using the trivial embedding in such a way that the actions of the operators $\mathbf{M}_i^A$ and $\mathbf{M}_i^B$ on $\mathbb{C}^2 \otimes \mathbb{C}^2$ are preserved. This completes the proof.

Since second-type representations are also first- and zero-type representations, we can apply Corollary 2 to obtain a first- and second-type representation in $\mathbb{C}^3 \otimes \mathbb{C}^3$.

4. A CONJECTURE ABOUT COMPATIBILITY OF EXEMPLARS

In quantum theory, measurement operators can be incompatible. That is, when we consider two different observables, the result of their sequential application can depend on the order in which they are applied. The fact that quantum measurements can be incompatible is related to fundamental differences between the quantum and classical realms, such as the observer phenomena, and the Heisenberg uncertainty principle (Heisenberg, 1927; Isham, 2001).

Definition 7. *Given two operators* $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *represented in the same basis. We say that* $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *represent compatible observables if and only if the commutator operator*

$$[\mathbf{M}_1, \mathbf{M}_2] = \mathbf{M}_1\mathbf{M}_2 - \mathbf{M}_2\mathbf{M}_1 = 0. \tag{71}$$

Otherwise, the operators represent incompatible observables.

In terms of cognitive phenomena, sequential measurements could be interpreted as consecutive cognitive actions where the previous action serves as a context for the next action (Busemeyer and Wang, 2007; Wang and Busemeyer, 2013). Since in our concrete representations membership operators are represented in the same basis for all exemplars $p_i = 1, ..., k$, it is now possible to test whether or not these measurement operators commute. If we find exemplars whose operators are non-commutative, then we can conjecture the existence of a fundamental limit to the precision with which the membership of these exemplars can be known simultaneously.

Note that we would expect that classical probabilistic models should be compatible, and because the classical probabilistic model and the tensor product model are equivalent, tensor product operators obtained from the data should also be compatible for the vector representing the conceptual situation. However, Hilbert space models could exhibit incompatible measurements for certain data on concept combination, as the Hilbert space model represents non-classical measurements.

We introduce the following definitions to characterize the compatibility of exemplars in $\mathbb{C}^3$ and in $\mathbb{C}^3 \otimes \mathbb{C}^3$:

Definition 8. *Let* $|A\rangle = (1, 0, 0)$, $|B\rangle = (0, 1, 0)$, *and* $\{\mathbf{M}_1, \mathbf{M}_2\}$ *be a representation in* $\mathbb{C}^3$ *of* $\{(\mu_i(A), \mu_i(B), \mu_i(AB))\}_{i=1}^2$, *and set*

$$c_A = \langle A|[\mathbf{M}_1, \mathbf{M}_2]|A\rangle,$$
$$c_B = \langle B|[\mathbf{M}_1, \mathbf{M}_2]|B\rangle,$$
$$c_{AB} = \frac{1}{2}(\langle A| + \langle B|)[\mathbf{M}_1, \mathbf{M}_2](|A\rangle + |B\rangle). \tag{72}$$

We say p_1 *and* p_2 *are compatible with respect to the concepts* $\mathcal{A}$, $\mathcal{B}$, *and* $\mathcal{AB}$ *if and only if* $c_A = 0$, $c_B = 0$, *and* $c_{AB} = 0$, *respectively.*

For simplicity, we will study compatibility for zero-type representations in $\mathbb{C}^3 \otimes \mathbb{C}^3$.

Definition 9. *Let* $|C\rangle = (1, 0, 0) \otimes (1, 0, 0)$, $\{\mathbf{M}_1^A, \mathbf{M}_1^B, \mathbf{M}_1^\wedge\}$, *and* $\{\mathbf{M}_2^A, \mathbf{M}_2^B, \mathbf{M}_2^\wedge\}$ *be a zero-type representation of data in* $\mathbb{C}^3 \otimes \mathbb{C}^3$ *of* $\{(\mu_i(A), \mu_i(B), \mu_i(AB))\}_{i=1}^2$, *and set*

$$
\begin{aligned}
c_A' &= \langle C|[\mathbf{M}_1^A, \mathbf{M}_2^A]|C\rangle, \\
c_B' &= \langle C|[\mathbf{M}_1^B, \mathbf{M}_2^B]|C\rangle, \\
c_{AB}' &= \langle C|[\mathbf{M}_1^\wedge, \mathbf{M}_2^\wedge]|B\rangle.
\end{aligned}
\tag{73}
$$

We say p_1 and p_2 are compatible with respect to concepts $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{AB}$ if and only if $c_A' = 0$, $c_B' = 0$, and $c_{AB}' = 0$, respectively.

We have verified the compatibility of exemplars for each conceptual combination that can be modeled in $\mathbb{C}^3$ and in $\mathbb{C}^3 \otimes \mathbb{C}^3$ using the data in Hampton (1988a,b). The results support our predictions. We have found that the tensor product model always leads to compatible measurements, and that the Hilbert space model leads to incompatible measurements in most cases.

For example, consider the concepts $\mathcal{A} = $ "*Machine*" and $\mathcal{B} = $"*Vehicle*," and the exemplars $p_5 = $"sailboat" and $p_{12} = $"skateboard." For the case of conceptual conjunction, we have

$$
\begin{aligned}
\mu_5(A) &= 0.56, \ \mu_5(B) = 0.8, \ \mu_5(AB) = 0.42, \ \text{and} \\
\mu_{12}(A) &= 0.28, \ \mu_{12}(B) = 0.84, \ \mu_{12}(AB) = 0.34.
\end{aligned}
\tag{74}
$$

Note that exemplar p_5 satisfies the conditions of Theorems 1 and 2. Thus, it can be represented in both $\mathbb{C}^3$ and in $\mathbb{C}^3 \otimes \mathbb{C}^3$. However, the exemplar p_{12} is singly overextended. Therefore, we can only represent the two exemplars simultaneously in $\mathbb{C}^3$.

When we apply Theorems 1 and 3, and Definition 8, on these data sets, we obtain

$$
c_A = 0.084\mathrm{i}, \ c_B = 0.097\mathrm{i}, \ \text{and} \ c_{AB} = 0.137\mathrm{i}.
\tag{75}
$$

Thus, exemplars p_5 and p_{12} are incompatible. Moreover, note that the incompatibility is larger for the conjunction of the concepts than for each of the former concepts.

As a second example, consider the concepts $\mathcal{A} = $"*Building*," and $\mathcal{B} = $"*Dwelling*," and the exemplars $p_2 = $"cave," and $p_{10} = $"synagogue," whose memberships are given by

$$
\begin{aligned}
\mu_2(A) &= 0.28, \ \mu_2(B) = 0.85, \ \mu_2(AB) = 0.28, \ \text{and} \\
\mu_{10}(A) &= 0.93, \ \mu_{10}(B) = 0.49, \ \mu_{10}(AB) = 0.45.
\end{aligned}
\tag{76}
$$

Both exemplars satisfy the conditions of Theorem 2. Applying Theorems 2 and 4, and Definition 9, we obtain

$$
c_A' = c_B' = c_{AB}' = 0.
$$

This is consistent with our expectations because the representation in the second sector $\mathbb{C}^3 \otimes \mathbb{C}^3$ correspond to classical (and thus compatible) measurements.

Since our data was collected presenting the exemplars in only one specific order (Hampton, 1988b), these computations demonstrate that we can predict order effects by determining the exemplars that are incompatible. The results presented here are, however, speculative since there is no experimental data where order effects have been recorded that could be used to contrast our computations. While our data set does not allow us to make a strong claim, we conjecture that order effects are predictable, and suggest that the concrete representations proposed in this paper could be used to develop Heisenberg-like uncertainty relations in the context of conceptual combinations.

5. CONCLUSION AND FUTURE WORK

In this paper, we have made some advances on the representational aspects of the quantum model for concept combinations. First, we proved that the first and second sectors of the two-sector Fock space model of concept conjunctions can be concretely represented in $\mathbb{C}^3$ and $\mathbb{C}^3 \otimes \mathbb{C}^3$, respectively. Next, we introduced unitary transformations to provide concrete representations that are consistent with the cognitive principles of the quantum model of concepts, and used these concrete representations to study the question of measurement compatibility.

The representations introduced here could be an important tool for future applications. First, since they are consistent with the cognitive principles of the quantum model of concepts, the model could easily be introduced to a wider audience, and extended to produce concrete representations in the two-sector Fock space model. Second, they can be adopted as a representational standard for different groups who seek to develop their own computational implementations of the model. Third, the fact that all the measurements are represented in a single basis constitutes a tremendous mathematical advantage for studying the probabilistic structure of concepts.

The evidence obtained in the application of our representations to the issue of exemplar compatibility is consistent with the assumptions of the model. Since the second sector entails logical reasoning, measurements in the tensor product model should be compatible. However, incompatible measurements are likely to be found in the Hilbert space model, since the first sector is associated with non-logical or intuitive reasoning. Moreover, this line of enquiry invites us to explore possible relations between the projector operator structure and the meaning of the exemplar.

In summary, the introduction of unitary transformations and the subsequent application to develop concrete representations of concepts and their combinations seems to be a promising line of research that has the potential to expand both theoretical and applied research in quantum cognition.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01734

REFERENCES

Aerts, D. (2007a). General quantum modeling of combining concepts: a quantum field model in fock space. *arXiv preprint arXiv:0705.1740*

Aerts, D. (2007b). Quantum interference and superposition in cognition: a theory for the disjunction of concepts. *arxiv.org/abs/0705.0975*

Aerts, D. (2009). Quantum structure in cognition. *J. Math. Psychol.* 53, 314–348. doi: 10.1016/j.jmp.2009.04.005

Aerts, D., and Aerts, S. (1995). Application of quantum statistics in psychological studies of decision processes. *Found. Sci.* 1, 85–97. doi: 10.1007/978-94-015-8816-4_11

Aerts, D., Broekaert, J., and Gabora, L. (2011a). A case for applying an abstracted quantum formalism to cognition. *New Ideas Psychol.* 29, 136–146. doi: 10.1016/j.newideapsych.2010.06.002

Aerts, D., Broekaert, J., Gabora, L., and Veloz, T. (2012a). "The guppy effect as interference," in *Quantum Interaction* (Berlin; Heidelberg: Springer), 36–47. doi: 10.1007/978-3-642-35659-9_4

Aerts, D., D'Hooghe, B., and Sozzo, S. (2011b). "A quantum cognition analysis of the Ellsberg paradox," in *Quantum Interaction* (Berlin; Heidelberg: Springer), 95–104. doi: 10.1007/978-3-642-24971-6_10

Aerts, D., and Gabora, L. (2005a). A state-context-property model of concepts and their combinations i: the structure of the sets of contexts and properties. *Kybernetes* 34, 151–175. doi: 10.1108/03684920510575799

Aerts, D., and Gabora, L. (2005b). A state-context-property model of concepts and their combinations ii: a hilbert space representation. *Kybernetes* 34, 176–204. doi: 10.1108/03684920510575807

Aerts, D., Gabora, L., and Sozzo, S. (2013). Concepts and their dynamics: a quantum-theoretic modeling of human thought. *Topics Cognit. Sci.* 5, 737–772. doi: 10.1111/tops.12042

Aerts, D., and Sozzo, S. (2011). "Quantum structure in cognition: why and how concepts are entangled," in *Quantum Interaction* (Berlin; Heidelberg: Springer), 116–127. doi: 10.1007/978-3-642-24971-6_12

Aerts, D., and Sozzo, S. (2014). Quantum entanglement in concept combinations. *Int. J. Theor. Phys.* 53, 3587–3603. doi: 10.1007/s10773-013-1946-z

Aerts, D., Sozzo, S., and Tapia, J. (2012b). "A quantum model for the Ellsberg and Machina paradoxes," in *Quantum Interaction* (Berlin; Heidelberg: Springer), 48–59. doi: 10.1007/978-3-642-35659-9_5

Blutner, R., Pothos, E. M., and Bruza, P. (2013). A quantum probability perspective on borderline vagueness. *Topics Cognit. Sci.* 5, 711–736. doi: 10.1111/tops.12041

Bruza, P., Kitto, K., Nelson, D., and McEvoy, C. (2009). Is there something quantum-like about the human mental lexicon? *J. Math. Psychol.* 53, 362–377. doi: 10.1016/j.jmp.2009.04.004

Bruza, P., Kitto, K., Ramm, L., Sitbon, L., Song, D., and Blomberg, S. (2012). Quantum-like non-separability of concept combinations, emergent associates and abduction. *Logic J. IGPL* 20, 445–457. doi: 10.1093/jigpal/jzq049

Bruza, P. D., and Cole, R. (2005). "Quantum logic of semantic space: an exploratory investigation of context effects in practical reasoning," in *We Will Show Them: Essays in Honour of Dov Gabbay, Vol. 1*, eds S. Artemov, H. Barringer, A. S. d'Avila Garcez, L. C. Lamb, and J. Woods (London: College Publications), 339–361.

Busemeyer, J., and Wang, Z. (2007). "Quantum information processing explanation for interactions between inferences and decisions," In *AAAI Spring Symposium: Quantum Interaction* (Stanford, CA: AAAI), 91–97.

Busemeyer, J. R., and Bruza, P. D. (2012). *Quantum Models of Cognition and Decision*. Cambridge: Cambridge University Press.

Busemeyer, J. R., Pothos, E. M., Franco, R., and Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychol. Rev.* 118, 193. doi: 10.1037/a0022542

Fodor, J. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

Fodor, J., and Lepore, E. (1996). The red herring and the pet fish: why concepts still can't be prototypes. *Cognition* 58, 253–270. doi: 10.1016/0010-0277(95)00694-X

Gilovich, T., Griffin, D., and Kahneman, D. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge: Cambridge University Press.

Hampton, J. (1988a). Disjunction of natural concepts. *Memory Cogn.* 16, 579–591. doi: 10.3758/BF03197059

Hampton, J. (1997a). Inheritance of attributes in natural concept conjunctions. *Memory Cogn.* 15, 55–71. doi: 10.3758/BF03197712

Hampton, J. A. (1997b). "Emergent attributes in combined concepts," in *Creative Thought: An Investigation of Conceptual Structures and Processes*, eds T. B. Ward, S. M. Smith, and J. Vaid (Washington, DC: American Psychological Association), 83–110.

Hampton, J. A. (1988b). Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 12. doi: 10.1037/0278-7393.14.1.12

Hampton, J. A. (1996). Conjunctions of visually based categories: overextension and compensation. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 378. doi: 10.1037/0278-7393.22.2.378

Haven, E., and Khrennikov, A. (2013). *Quantum Social Science*. Cambridge: Cambridge University Press.

Heisenberg, W. (1927). Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift Phys.* 43, 172–198. doi: 10.1007/BF01397280

Isham, C. J. (2001). *Lectures on Quantum Theory*. New Delhi: Allied Publishers.

Kamp, H., and Partee, B. (1995). Prototype theory and compositionality. *Cognition* 57, 121–191. doi: 10.1016/0010-0277(94)00659-9

Khrennikov, A. (2009). Quantum-like model of cognitive decision making and information processing. *Biosystems* 95, 179–187. doi: 10.1016/j.biosystems.2008.10.004

Medin, D. (1989). Concepts and conceptual structure. *Am. Psychol.* 44, 1469–1481. doi: 10.1037/0003-066X.44.12.1469

Medin, D. L., and Shoben, E. J. (1988). Context and structure in conceptual combination. *Cogn. Psychol.* 20, 158–190. doi: 10.1016/0010-0285(88)90018-7

Meyer, P. A. (1995). *Quantum Probability for Probabilists*, Vol. 1538. Heidelberg: Springer Science & Business Media.

Moro, R. (2009). On the nature of the conjunction fallacy. *Synthese* 171, 1–24. doi: 10.1007/s11229-008-9377-8

Pitowsky, I. (1989). *Quantum Probability – Quantum Logic*, Vol. 321. Berlin: Springer.

Pothos, E. M., and Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* 36, 255–274. doi: 10.1017/S0140525X12001525

Rips, L. J. (1995). The current status of research on concept combination. *Mind Lang.* 10, 72–104. doi: 10.1111/j.1468-0017.1995.tb00006.x

Smith, E., and Osherson, D. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition* 9, 35–38. doi: 10.1016/0010-0277(81)90013-5

Sozzo, S. (2014). A quantum probability explanation in fock space for borderline contradictions. *J. Math. Psychol.* 58, 1–12. doi: 10.1016/j.jmp.2013.11.001

Storms, G., De Boeck, P., Van Mechelen, I., and Ruts, W. (1998). Not guppies, nor goldfish, but tumble dryers, noriega, jesse jackson, panties, car crashes, bird books, and stevie wonder. *Memory Cognit.* 26, 143–145. doi: 10.3758/BF03211377

Tenenbaum, J., Kemp, C., Griffiths, T., and Goodman, N. (2011). How to grow a mind: statistics, structure, and abstraction. *Science* 331, 1279–1286. doi: 10.1126/science.1192788

Tversky, A., and Kahneman, D. (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychol. Rev.* 90, 293. doi: 10.1037/0033-295X.90.4.293

Veloz, T. (2015). *Towards a Quantum Theory of Cognition: History, Development, and Perspectives*. Ph.D. thesis, University of British Columbia - Okanagan.

Wang, Z., and Busemeyer, J. R. (2013). A quantum question order model supported by empirical tests of an A Priori and precise prediction. *Topics Cogn. Sci.* 5, 689–710. doi: 10.1111/tops. 12040

Reintroducing the Concept of Complementarity into Psychology

Zheng Wang [1]* and Jerome Busemeyer [2]

[1] School of Communication, The Ohio State University, Columbus, OH, USA, [2] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

Central to quantum theory is the concept of complementarity. In this essay, we argue that complementarity is also central to the emerging field of quantum cognition. We review the concept, its historical roots in psychology, and its development in quantum physics and offer examples of how it can be used to understand human cognition. The concept of complementarity provides a valuable and fresh perspective for organizing human cognitive phenomena and for understanding the nature of measurements in psychology. In turn, psychology can provide valuable new evidence and theoretical ideas to enrich this important scientific concept.

INTRODUCTION

Central to quantum theory is the concept of complementarity. This essay argues that complementarity is also central to the emerging field of quantum cognition (e.g., Aerts and Aerts, 1994; Khrennikov, 1999; Pothos and Busemeyer, 2013; Wang et al., 2013; Bruza et al., 2015; Busemeyer and Wang, 2015), which applies abstract, mathematical principles of quantum theory to shed light on cognitive structures and processes. The concept of complementarity provides a valuable and fresh perspective for organizing human cognitive phenomena and for understanding the nature of measurements in psychology. In turn, psychology can provide useful new evidence and theoretical ideas to enrich this important scientific concept.

COMPLEMENTARITY, COMMUTATIVITY, AND COMPATIBILITY

The general concept of complementarity was developed by Niels Bohr in a series of debates with Einstein, but the main idea can be summarized as follows (Plotnitsky, 2014, p. 5): Different measurement conditions for observing different phenomena are complementary when

(a) they are mutually exclusive, and only one can be applied at any time; and

(b) they are all necessary for a comprehensive account of these phenomena.

An important consequence of complementarity is that the sequence or order of the measurements matters (von Neumann, 1932, 1962; Atmanspacher and Römer, 2012; Wang et al., 2014). The above definition of complementarity is deliberately general so that it can permit many specific implementations. Below, we provide a way to implement this idea in psychology.

The essential idea of complementarity can be illustrated using the following example involving the measurements of attitudes toward politicians. In a 1997 poll in the United States, half of the 1,002 nationally sampled respondents were asked, "Do you generally think Clinton is honest?" Then they were asked the same question about Gore. The other half answered the same questions in the opposite

order. The results exhibited a striking order effect: The proportion saying "yes" to both questions was significantly higher when Gore was judged first (Moore, 2002).
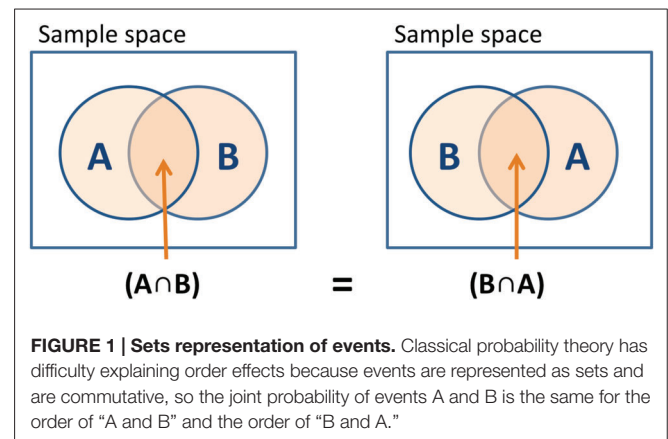
In this example, the phenomena of interest concern a survey respondent's beliefs about the honesty of different politicians. Complementarity arises when a person cannot have a well-defined position on each politician simultaneously. We can obtain a measurement of honesty concerning Clinton or concerning Gore, but we cannot measure both simultaneously, and the order in which we measure them affects the answers. Once we obtain a measurement on say, Clinton, that decision can create a definite position for Clinton, but then the opinion regarding the Gore must be uncertain. However, both measurements are needed to obtain a complete understanding of a respondent's attitude to the two politicians being considered. Therefore, these measurements satisfy the general requirements for complementarity.

This example captures another idea relevant to complementarity. The phenomena that we observe are products of an interaction between some object of investigation and our measurement instruments. The measurement does not simply record a phenomenon, but it creates one. This idea is consistent with the constructionist view of beliefs, attitudes, and intentions proposed by many psychologists (e.g., Feldman and Lynch, 1988; Schwarz, 2007). From this view, because of limited mental capacity and cognitive economy, beliefs, attitudes, and intentions do not exist in memory as properties ready to be recorded; instead, they are constructed when needed. When one is asked a subsequent question, information carried over from the preceding question provides a context for the construction of the second and influences the subsequent response.

Next, we will explain complementarity more specifically by providing a simple "toy" quantum model for this example. To do this, we need to first compare some concepts from classical and quantum probability theories (see Busemeyer and Bruza, 2012, for more detail).

Classical probability theory is concerned with the assignment of probabilities to events. Suppose, for example, we ask a survey participant to evaluate various politicians with regard to their honesty. For example, an event, A, might be that politician X is evaluated as honest. According to classical probability theory, events are represented as subsets of a universal set[1]. For example, the event that a politician is honest is a subset of the universe of all the features that a politician might have. Another event, B, might be that politician Y is evaluated as dishonest. The conjunction of two events is defined by set intersection—in this case, A and B. As shown in **Figure 1**, the combined event "A and B" is the same as the combined event "B and A," and therefore the order of the two events does not matter. Formally, we say that the intersection event is *commutative*, and the probability assigned to "A and B" must equal the probability assigned to "B and A."

Quantum probability theory is also concerned with the assignment of probabilities to events. However, according to quantum theory, events are represented as subspaces of a

FIGURE 1 | Sets representation of events. Classical probability theory has difficulty explaining order effects because events are represented as sets and are commutative, so the joint probability of events A and B is the same for the order of "A and B" and the order of "B and A."

universal vector space[2]. If events are defined as subspaces, then the conjunction of two events may or may not exist. The conjunction does not exist if the events are non-commutative so that the order of evaluating them matters. Events that are commutative are also called *compatible*, and events that are non-commutative are called *incompatible* (Atmanspacher and Römer, 2012). Classical probability theory essentially assumes that all events are compatible, but quantum probability theory allows some events to be incompatible.

Figure 2 illustrates how the projective geometry used by quantum probability theory naturally accounts for order effects. The "yes" answer to the "Do you generally think Clinton is honest?" question is represented by the horizontal ray (which forms one axis from the blue basis), and the "yes" answer to the "Do you generally think Gore is honest?" question is represented by an oblique ray (which forms one axis from the red basis). These two answers are *incompatible* because the subspaces (rays in this "toy" example) for these answers are not defined by a common basis. A person needs to evaluate the Clinton question using one pair of axes (the blue axes), and then must shift her or his viewpoint to another pair of axes (the red axes) to evaluate the Gore question. The final result depends on the order of the applications, because answering one question provides a new contextualized state that is used to generate responses to the second question. As a consequence of incompatibility, if a person is certain about an answer to one question, then the person must be uncertain about the answer to the other question (evidencing the uncertainty principle of quantum theory). In other words, when the questions are incompatible, one cannot be certain about the answers to both questions simultaneously (evidencing the superposition principle of quantum theory).

A key point here is that different bases (red vs. blue axes in **Figure 2**) are required to perform the Clinton and Gore measurements. According to quantum theory, two measurement conditions are complementary whenever we have to change the basis used to represent the outcomes of each measurement.

---

[1]The universal set is a set that contains a sigma algebra of subsets. The subsets are the events that can occur and the events are subsets of the universal set.

[2]The universal vector space is a vector space spanned by a set of basis vectors. The vector space contains subspaces, which are closed subsets of this vector space, and events are subspaces of the vector space.
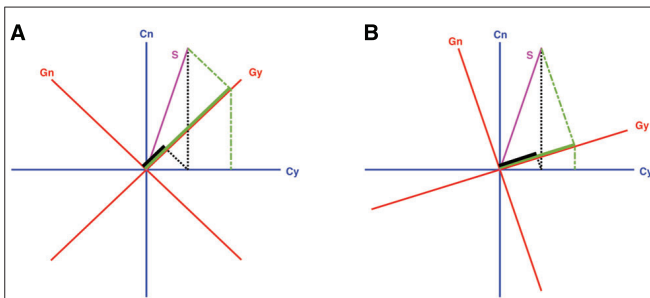
FIGURE 2 | A "toy" example of quantum probability model of sequential judgments. (A) Illustrates how quantum model provides a natural account for question order effects of the Clinton-Gore example in terms of incompatibility (Wang and Busemeyer, 2013; Wang et al., 2014). First, consider the probability of a person's answering "yes" to both questions when Clinton is judged first. This is obtained by first projecting (following the black dotted lines) the magenta-colored state S to the blue Cy ("yes" to Clinton) axis and then projecting the result up to the red Gy ("yes" to Gore) axis, which produces a small probability (as illustrated by the short length of the black projection on the Gy axis). Next, consider the probability of a person's answering "yes" to both questions when Gore is judged first. This is obtained by first projecting (following the green dash-dotted lines) the state S down to the red Gy ("yes" to Gore) axis and then projecting the result to the blue Cy ("yes" to Clinton) axis, which produces a much higher probability. In addition, note that the probability of saying "yes" to Gore in this Gore-Clinton order is much higher than it is in the Clinton-Gore order (as illustrated by the long length of the green projection on the Gy axis), producing the order effect for the Gore question when the question is asked before, as opposed to after, the Clinton question. **(B)** Illustrates similar order effects, but the order effects are much smaller because the basis vectors (i.e., the red and blue axes), which form the subspace for evaluating the Gore and Clinton questions, are more aligned with each other. This means that the rotation required to change from one evaluation basis (e.g., the blue Clinton axes) to the other (e.g., the red Gore axes) is smaller. In a sense, in psychology we could understand this as "reduced incompatibility" of the two questions, or two sets of projectors as "more nearly commutative," although in quantum physics, the original concept of complementarity does not have the notion of degrees of complementarity (events can be differentiated only by being complementary or not).

What makes two measures compatible in psychology? Two questions are compatible if the subspaces representing each question are defined by a common basis. In our example, to form a common basis for representing the Clinton and Gore questions, we must posit at least a four-dimensional space, with the four basis vectors (or axes) representing the four conjunctions: (1) "yes" to Clinton and "yes" to Gore, (2) "yes" to Clinton and "no" to Gore, (3) "no" to Clinton and "yes" to Gore, and (4) "no" to Clinton and "no" to Gore. The belief state would be a vector in this four-dimensional space, and each coordinate would indicate the belief about a conjunction (e.g., the belief in "yes" to Clinton and "no" to Gore). When a compatible representation is used, the order of questions does not matter, because the person eventually arrives at the same conjunction with the same probability when finished. Also, the person can be certain about the answers to both questions at the same time. This seems like a more ideal case of human cognition. This, however, all comes at a higher cost, because more cognitive resources are needed to increase and maintain the higher dimensionality of the compatible representation space (Wang and Busemeyer, 2013; Bruza et al., 2015).

FROM PSYCHOLOGY TO PHYSICS: THE HISTORY

It is an interesting twist of the history that the term "complementary" first appeared in the foundational work of psychology. In one of the most influential classic works in psychology, *The Principles of Psychology*, James (1890) wrote,

> "…in certain persons, at least, the total possible consciousness may be split into parts which coexist but mutually ignore each other, and share the objects of knowledge between them. More remarkable still, they are complementary." (p. 204)

Although there is still debate among philosophers and historians whether Bohr's concept of complementarity was influenced by James, many agree on the clear similarity between the concept of complementarity that James created for psychology in 1890 and that Bohr introduced into physics four decades later, and believe Bohr was at least indirectly affected by James's work (e.g., Stapp, 1993; Plotnitsky, 2012). The concept of complementarity emerged around 1926 and 1927 from the discussions between Bohr and Werner Heisenberg related to the discovery of the uncertainty principle. In a lecture in Como, Italy, in 1927, Bohr (1928) for the first time discussed complementarity in public, and the lecture was published the next year. By the time of the famous debate with Einstein regarding the Einstein-Podolsky-Rosen experiment, Bohr had developed a rather complete definition of complementarity (Plotnitsky, 2014):

> "Evidence obtained under different experimental conditions cannot be comprehended within a single picture, but must be regarded as complementary in the sense that only the totality of the phenomena exhaust the possible information about the objects…" (Bohr, 1987a, p. 40)

It is interesting that as an adolescent, Bohr had shown interest in describing human conscious processes (Folse, 1985, p. 175). Even in his earlier papers on complementarity and quantum physics, he tried to state how the concept of complementarity could be applied to psychology. For example, he ended his Como lecture,

> "I hope, however, that the idea of complementarity is suited to characterize the situation, which bears a deep-going analogy to the general difficulty in the formation of human ideas, inherent in the distinction between subject and object." (Bohr, 1928, p. 590)

A year later, in a paper he wrote for a Planck *Festschrift* in 1929, he stated his view on applying complementarity to psychology with greater clarity:

> "For describing our mental activity, we require, on one hand, an objectively given content to be placed in opposition to a perceiving subject, while, on the other hand, as is already implied in such an assertion, no sharp separation between object and subject can be maintained,

since the perceiving subject also belongs to our mental content. From these circumstances follows not only the relative meaning of every concept, or rather of every word, the meaning depending upon our arbitrary choice of view point, but also that we must, in general, be prepared to accept the fact that a complete elucidation of one and the same object may require diverse points of view which defy a unique description. Indeed, strictly speaking, the conscious analysis of any concept stands in a relation of exclusion to its immediate application. The necessity of taking recourse to a complementary, or reciprocal, mode of description is perhaps most familiar to us from psychological problems." (Bohr, 1987b, p. 96)

Complementarity is not limited to physics. Instead, it is a general concept that can be applied to any phenomena that are featured by "a participating observer." As Bohr recognized, these kinds of phenomena are typical in psychology. In the end, psychology is the field that studies "the participating observer"—the observer's perception, attention, emotion, motivation, memory, and decision-making, among other psychological processes.

The concept of complementarity applies naturally to psychological systems. Just like a physical system, a psychological system can be measured in different, mutually exclusive ways. Although all these measurements are essential for describing the system, they cannot be measured simultaneously, only sequentially. In this case, we say the different measurements are complementary (Stapp, 1993). Importantly, this means that the measurement is "an essential part of making a property definite" (Stapp, 1993, p. 234). In other words, measurements do not merely record the property of a system but construct it.

EMPIRICAL TESTABILITY OF COMPLEMENTARITY IN PSYCHOLOGY

Two criticisms are often raised in response to quantum cognition because of misunderstandings of this new research program. One we believe is a false alarm due to a general resistance to—and often a legitimate concern about—the loose, vague, metaphorical, speculative extension of quantum physics to cultural and social studies (Beller, 1998). However, differently from what is being argued against, the research program of quantum cognition rigorously uses mathematical principles of quantum probability theory to build new models of human cognition, develop specific new predictions, and empirically test the new predictions and compare new models against existing traditional models. Just like other cognitive models based on classical probability theory, quantum cognition models take advantage of quantum formalism to provide new theoretical and modeling tools that make precise predictions regarding human cognition.

The other typical criticism questions whether quantum cognition can ever provide the kind of rigor and precision that is shown by quantum mechanics. Unfortunately, it is true that compared to quantum physics, which provides rigorous and precise predictions about physical phenomena, psychological theories involve many more random variables that are hardly

controlled, resulting in lower precision in prediction. To be fair, this is a general challenge for any theories in the behavioral and social sciences. However, through rigorous model comparison, empirical studies have shown that quantum models provide an elegant new way to specify general and vague verbal theories in psychology, and better explain and predict many phenomena puzzling to classical models, leading to highly testable models (e.g., Bruza et al., 2015; Busemeyer and Wang, 2015; Busemeyer et al., 2015).

In fact, compared to many other psychological theories and models, quantum cognitive models may be more falsifiable. Because quantum cognitive models are based on a coherent set of axioms that are clearly stated, these models must stand up to strict tests of these axioms in addition to performance comparisons against competing classical models. Using our quantum question order model as an example again, the model provides clear theoretical predictions about when order effects will or will not occur as well as the pattern of order effects that do occur (Wang and Busemeyer, 2013). One of the most convincing examples illustrating the testability of quantum models has been an *a priori*, parameter-free, and precise test called the quantum question equality, or QQ equality (Wang and Busemeyer, 2013; Wang et al., 2014). This equality, derived from quantum theory, imposes a strong symmetry condition on the nature of order effects, and empirical results from more than 70 U.S. national surveys provided surprisingly strong support for this precise prediction (Wang et al., 2014). Rarely in social science research do we find *a priori* and parameter-free predictions being upheld with such high accuracy. Classical models cannot explain—in a principled and *a priori* manner—both the question order effects and the QQ equality observed in the empirical data (Wang et al., 2014).

EXTENDING THE CONCEPT OF COMPLEMENTARITY IN PSYCHOLOGY

Psychology provides an opportunity to extend and enrich the concept of complementarity beyond what is being formulated in physics. When applied to psychology as opposed to physics, compatibility may take on a more fluent and malleable role. Perhaps compatibility varies across individuals, develops across age, and changes with experience. For example, very young children do not seem to have the ability to take on the perspective of another person—this capability to change perspectives develops only after a critical developmental stage (e.g., Epley et al., 2004).

As another example, perhaps compatibility can be formed after an individual has had many experiences with combinations of events that permit the formation of conjunctive concepts. To be more specific, if a combination of questions is new or unusual, then an answer must be constructed on-line that relies on a simpler, incompatible, lower-cost representation. However, if a person has a great deal of experience with a combination, then the person may have sufficient knowledge to form a compatible representation as a result of cognitive adaptation to the environment. Therefore, order effects are expected to occur for uncommon or unfamiliar pairs of questions, whose answers must be (at least partially) constructed on the spot. Indeed, two

field experiments during the 1988 and 1992 presidential elections supported this possibility (Simmons et al., 1993). The authors found that the question order effects on issue opinions decreased as the election became closer, which would be predicted by the quantum model because the measurements on issue opinions might more frequently occur over time during media exposure or daily conversations—even if the measurements were not directly noted.

In sum, at this early stage of research, the concept of compatibility is new in psychology, and we can only speculate about which measures will be compatible or incompatible. Then the speculations or assumptions can be empirically tested based on order effects or interference effects of the measures, among other predicted effects that follow incompatibility. However, this will be a crucial question for future research in quantum cognition, which should enrich the concept of complementarity through psychological experiments and theories.

REFERENCES

Aerts, D., and Aerts, S. (1994). Applications of quantum statistics in psychological studies of decision processes. *Found. Sci.* 1, 85–97. doi: 10.1007/BF00208726

Atmanspacher, H., and Römer, H. (2012). Order effects in sequential measurements of non-commuting psychological observables. *J. Math. Psychol.* 56, 274–280. doi: 10.1016/j.jmp.2012.06.003

Beller, M. (1998). The Sokal hoax: at whom are we laughing? *Phys. Today* 51, 29–34. doi: 10.1063/1.882436

Bohr, N. (1928). The quantum postulate and the recent development of atomic theory. *Nature* 121, 580–590. doi: 10.1038/121580a0

Bohr, N. (ed.) (1987a). "Discussion with Einstein on epistemological problems in atomic physics," in *The Philosophical Writings of Niels Bohr*, Vol. 2, (Woodbridge, CT: Ox Bow Press), 32–66. (Original work published 1949.)

Bohr, N. (ed.) (1987b). "The quantum of action and the description of nature," in *The Philosophical Writings of Niels Bohr*, Vol. 1, (Woodbridge, CN: Ox Bow Press), 92–101. (Original work published 1929.)

Bruza, P. D., Wang, Z., and Busemeyer, J. R. (2015). Quantum cognition: a new theoretical approach to psychology. *Trends Cogn. Sci.* 19, 383–393. doi: 10.1016/j.tics.2015.05.001

Busemeyer, J. R., and Bruza, P. D. (2012). *Quantum Models of Cognition and Decision*. Cambridge: Cambridge University Press.

Busemeyer, J. R., and Wang, Z. (2015). What is quantum cognition, and how is it applied to psychology? *Curr. Dir. Psychol. Sci.* 24, 163–169. doi: 10.1177/0963721414568663

Busemeyer, J. R., Wang, Z., and Shiffrin, R. (2015). Bayesian comparison of a quantum versus a traditional model of human decision making. *Decision* 2, 1–12. doi: 10.1037/dec0000017

Epley, N., Morewedge, C. K., and Keysar, B. (2004). Perspective taking in children and adults: equivalent egocentrism but differential correction. *J. Exp. Soc. Psychol.* 40, 760–768. doi: 10.1016/j.jesp.2004.02.002

Feldman, J. M., and Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *J. Appl. Psychol.* 73, 421–435. doi: 10.1037/0021-9010.73.3.421

Folse, H. J. (1985). *The Philosophy of Niels Bohr: The Framework of Complementarity*. Amsterdam: Elsevier.

Grossberg, S. (2000). The complementary brain: unifying brain dynamics and modularity. *Trends Cogn. Sci.* 4, 233–246. doi: 10.1016/S1364-6613(00)01464-9

James, W. (1890). *The Principles of Psychology*. New York/London: Holt and Macmillan.

Khrennikov, A. Y. (1999). Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social, and anomalous

CONCLUDING COMMENTS

As we have described, the idea of complementarity was introduced into psychology by James (1890). Later, the idea was developed formally and became one of the centerpieces of Niels Bohr's interpretation of quantum mechanics. Unfortunately, for many years the concept appeared to be useful only in physics, and it almost disappeared from the psychological literature (for exceptions, see Grossberg, 2000). In this article, we have attempted to reintroduce the concept of complementarity to its original home in psychology. We think the concept provides an invaluable service toward understanding the fundamental nature of human cognition.

ACKNOWLEDGMENTS

phenomena. *Found. Phys.* 29, 1065–1098. doi: 10.1023/A:1018885632116

Moore, D. W. (2002). Measuring new types of question-order effects. *Public Opin. Q.* 66, 80–91. doi: 10.1086/338631

Plotnitsky, A. (2012). *Niels Bohr and Complementarity: An Introduction*. Berlin: Springer.

Plotnitsky, A. (2014). What is complementarity? Niels Bohr and the architecture of quantum theory. *Phys. Scr.* 2014, T163. doi: 10.1088/0031-8949/2014/T163/014002

Pothos, E. M., and Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* 36, 255–274. doi: 10.1017/S0140525X12001525

Schwarz, N. (2007). Attitude construction: evaluation in context. *Soc. Cogn.* 25, 638–656. doi: 10.1521/soco.2007.25.5.638

Simmons, C. J., Bickart, B. A., and Lynch, J. G. (1993). Capturing and creating public opinion in survey research. *J. Consum. Res.* 20, 316–329. doi: 10.1086/209352

Stapp, H. P. (1993). *Mind, Matter, and Quantum Mechanics*. Berlin: Springer.

von Neumann, J. (1932). *Mathematical Foundations of Quantum Theory*. Princeton, NJ: Princeton University Press.

von Neumann, J. (1962). "Logic, theory of sets and quantum mechanics," in *Collected Works*, Vol. 1, ed. A. H. Taub (New York, NY: Pergamon Press), 151–422.

Wang, Z., and Busemeyer, J. R. (2013). A quantum question order model supported by empirical tests of an *a priori* and precise prediction. *Top. Cogn. Sci.* 5, 689–710. doi: 10.1111/tops.12040

Wang, Z., Busemeyer, J. R., Atmanspacher, H., and Pothos, E. M. (2013). The potential of using quantum theory to build models of cognition. *Top. Cogn. Sci.* 5, 672–688. doi: 10.1111/tops.12043

Wang, Z., Solloway, T., Shiffrin, R. M., and Busemeyer, J. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9431–9436. doi: 10.1073/pnas.1407756111

Preference reversal in quantum decision theory

Vyacheslav I. Yukalov [1,2*†] and Didier Sornette [1,3†]

[1] Department of Management, Technology and Economics, ETH Zürich, Zürich, Switzerland, [2] Bogolubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, Dubna, Russia, [3] Swiss Finance Institute, University of Geneva, Geneva, Switzerland

We consider the psychological effect of preference reversal and show that it finds a natural explanation in the frame of quantum decision theory. When people choose between lotteries with non-negative payoffs, they prefer a more certain lottery because of uncertainty aversion. But when people evaluate lottery prices, e.g., for selling to others the right to play them, they do this more rationally, being less subject to behavioral biases. This difference can be explained by the presence of the attraction factors entering the expression of quantum probabilities. Only the existence of attraction factors can explain why, considering two lotteries with close utility factors, a decision maker prefers one of them when choosing, but evaluates higher the other one when pricing. We derive a general quantitative criterion for the preference reversal to occur that relates the utilities of the two lotteries to the attraction factors under choosing vs. pricing and test successfully its application on experiments by Tversky et al. We also show that the planning paradox can be treated as a kind of preference reversal.

Keywords: preference reversal, decision theory, uncertainty, behavioral quantum probability, planning paradox

1. Introduction

For many decades, psychologists and economists have been intrigued by a seemingly anomalous effect termed *preference reversal*. The simplest example illustrating this effect is as follows. First, subjects are asked to choose between two lotteries, say L_1 and L_2, such that L_1 has a high chance to win a relatively modest prize, while L_2 offers a lower chance of winning, but an essentially larger prize. The majority of subjects choose the more certain win of lottery L_1, despite the fact that lottery L_2 can enjoy a larger expected utility. Then subjects are asked to price each of the lotteries, as if they would own them and wish to sell the right to play them. Surprisingly, the majority of subjects price higher the less certain lottery L_2 in apparent contradiction with their previous choice. This example embodies the essence of the preference reversal effect.

Among the first scientists emphasizing the existence of this effect were Lindman (1971) and Lichtenstein and Slovic (1971, 1973). Their studies were followed by several authors demonstrating the occurrence of this effect in psychology and economics (Grether and Plott, 1979; Loomes and Sugden, 1983; Holt, 1986; Goldstein and Einborn, 1987; Karni and Safra, 1987; Segal, 1988; Tversky et al., 1988; Schkade and Johnson, 1989). Many other citations can be found in the review articles (Slovic and Lichtenstein, 1983; Tversky and Thaler, 1990; Tversky et al., 1990). The experimental studies have established the clear validity and robustness of the preference reversal phenomenon.

The preference reversal effect looks surprising because, according to the common understanding of utility, the choice among the given lotteries should be based on the objective values of the latter, thus, being procedure invariant. Since the lottery values are not changed, why then is the preference reversed?

It has been proved by Tversky and Thaler (1990) and Tversky et al. (1990) that it is the *breaking of procedure invariance* that is responsible for the preference reversal phenomenon. It turns out that subjects weight more heavily payoffs in pricing than in choice, so that the preference reversal is a purely psychological effect.

The origin of the preference reversal has been recently explained from the point of view of neurology by Kim et al. (2012). It has been experimentally shown that there exists correlation between visual fixation and preferences. Visual fixations both reflect and influence preferences. From one side, these fixations reflect which objects seem to be more important for the subject. And, from the other side, such fixations modulate the neural correlates of preferences, with activity in ventromedial prefrontal cortex and ventral striatum, reflecting the value of the fixated item compared to the value of the item not fixated. Kim et al. studied the process of decision making under risk and measured eye movements while people chose between gambles or bid in pricing gambles. Consistently with the previous work, they found that, for two gambles matched in expected value, people systematically chose the higher probability option, but requested a higher ask price for the option that offered the greater amount to win, thus demonstrating preference reversal.

This effect was accompanied by a shift in fixation of the two attributes, with people fixating more on probabilities during choices and more on amounts during selling. In this way, there exists *probability-vs.-amount dichotomy*: When choosing, one pays more attention to probabilities while, when selling, one better appreciates amounts.

Understanding the cause of the preference reversal is the first necessary step. The next step should be the description of this effect by a mathematical model. Previous suggested models were not successful, as was analyzed by Tversky and Thaler (1990) and Tversky et al. (1990). In the present paper, we show that the effect of preference reversal finds a simple and natural explanation in the frame of the Quantum Decision Theory developed by the authors (Yukalov and Sornette, 2008, 2009a,b, 2010, 2011, 2013, 2014a,b, 2015).

2. Basics of Quantum Decision Theory

There exists several approaches applying quantum notions to psychological sciences, as can be inferred from the books (Khrennikov, 2010; Busemeyer and Bruza, 2012; Bagarello, 2013; Haven and Khrennikov, 2013) and the review articles (Yukalov and Sornette, 2009b; Busemeyer et al., 2014; Sornette, 2014; Ashtiani and Azgomi, 2015), where numerous citations to the previous literature can be found. Quantum Decision Theory (QDT) principally differs from all those approaches in two aspects. First, QDT is based on a self-consistent mathematical foundation that is common for both quantum measurement theory and quantum decision theory. Starting from the von Neumann (1955) theory of quantum measurements, we have generalized it to the case of uncertain or inconclusive events, making it possible to characterize uncertain measurements and uncertain prospects. Second, the main formulas of QDT are derived from general principles, giving the possibility of

quantitative predictions, without fitting parameters. This is in contrast with the usual way of constructing particular models for describing some concrete experiments, with fitting the model parameters from empirical data.

We shall not repeat here the mathematical foundation of QDT that has been thoroughly expounded in our previous papers, but we will just briefly recall the resulting formulas that are necessary for describing the preference reversal effect.

Let us consider a composite event, called prospect,

$$\pi_n = A_n \bigotimes B . \qquad (1)$$

Here A_n is an operationally testable event, represented in a Hilbert space by an eigenstate $|n\rangle$. While $B = \{B_\alpha, \ b_\alpha\}$ is an inconclusive event that is a set of possible events B_α, represented in a Hilbert space by eigenstates $|\alpha\rangle$, and equipped with random amplitudes b_α, so that the inconclusive event is represented by a state $|B\rangle = \sum_\alpha b_\alpha |\alpha\rangle$.

The prospect operator is $\hat{P}(\pi_n) = |nB\rangle\langle nB|$, such that the prospect probability is given by the quantum formula

$$p(\pi_n) = \text{Tr}\hat{\rho}\hat{P}(\pi_n) , \qquad (2)$$

where $\hat{\rho}$ is a strategic state of a decision maker. By construction, the prospect probability enjoys the properties of a probability measure:

$$\sum_n p(\pi_n) = 1 , \qquad 0 \le p(\pi_n) \le 1 . \qquad (3)$$

It is easy to show that the prospect probability takes the form

$$p(\pi_n) = f(\pi_n) + q(\pi_n) , \qquad (4)$$

where the first term is called *utility factor*, characterizing the utility of the prospect, while the second term is *attraction factor* representing behavioral biases.

The intuitive explanation of the above probability expression (4) is straightforward: The definition of a quantum probability (2) for a composite event can be separated into a term containing diagonal matrix elements and a term including off-diagonal elements. The diagonal elements compose the term $f(\pi_n)$, while the off-diagonal elements define the term $q(\pi_n)$. The occurrence of an off-diagonal term is a typical feature of quantum theory, where this quantity is called *interference term* or *coherence term*. The existence of such an interference term constitutes the principal difference of the quantum approach from the classical consideration, where there are no interference terms. It is the appearance of interference terms that makes the structure of quantum expressions richer then the related classical ones and that allows one to explain those psychological phenomena that, otherwise, are inexplicable in classical decision making. Sometimes, the quantum approach even yields conclusions that are impossible in classical decision making, as, for instance, the possibility to agree on disagree (Khrennikov and Basieva, 2014). Below we show that this interference term, composing the attraction factor, is essential in explaining the existence of

the preference reversal effect that cannot be described in classical decision theory.

The prospect probability satisfies the quantum-classical correspondence principle.

$$p(\pi_n) \to f(\pi_n) , \qquad q(\pi_n) \to 0 . \qquad (5)$$

This defines the utility factor as a classical-type probability, with the standard properties

$$\sum_n f(\pi_n) = 1 , \qquad 0 \leq f(\pi_n) \leq 1 . \qquad (6)$$

This is equivalent to the normalization condition

$$\sum_{n\alpha} |b_\alpha|^2 \langle n\alpha \mid \hat{\rho} \mid n\alpha \rangle = 1 ,$$

imposing a constraint on the random quantities b_α.

When considering lotteries, an event $A_n \equiv A(L_n)$ implies the choice of a lottery L_n. Then the inconclusive set B characterizes the decision maker hesitations between uncertain events B_α, describing uncertainty with respect to the decision maker ability and with respect to the lottery formulation (Yukalov and Sornette, 2014b, 2015). The explicit form of the utility factor is given by minimizing the Kullback-Leibler information functional, which in the simple case of uncertainty yields

$$f(\pi_n) = \frac{U(L_n)}{\sum_n U(L_n)} , \qquad (7)$$

with $U(L_n)$ being the expected utility of a lottery L_n. Note that the minimization of the information functional results in expression (7) that might be familiar to psychologists as a Luce (1959) choice rule using utility as response strength.

The attraction factor reflects the effects of quantum coherence and interference, and in decision theory it represents the behavioral biases rendering the prospects more or less attractive from the subconscious point of view of decision maker. By their definition, attraction factors lie in the interval

$$-1 \leq q(\pi_n) \leq 1 \qquad (8)$$

and satisfy the *alternation property*

$$\sum_n q(\pi_n) = 0 . \qquad (9)$$

Also, in the case of non-informative priors, the attraction factors for the considered prospect lattice $\{\pi_n : n = 1, 2, \ldots, N\}$ obey the *quarter law*

$$\frac{1}{N} \sum_{n=1}^{N} |q(\pi_n)| = \frac{1}{4} . \qquad (10)$$

This law makes it admissible to estimate the attraction factors by the values ± 0.25, thus quantitatively predicting preferences.

The prospect lattice is ordered by the values of prospect probabilities. A prospect π_i is termed preferable to π_j if and only if

$$p(\pi_i) > p(\pi_j) \qquad (\pi_i > \pi_j) .$$

At the same time, a prospect π_i is more useful than π_j when $f(\pi_i) > f(\pi_j)$. A prospect π_i is more attractive than π_j, when $q(\pi_i) > q(\pi_j)$. In this way, a prospect can be more useful but less attractive, as a result being less preferable.

A necessary condition for the existence of a nonzero attraction factor is that the composite prospect be entangled (Yukalov and Sornette, 2014a, 2015). Otherwise, there is no need of involving quantum probabilities.

3. General Criterion of Preference Reversal

Preference reversal may naturally arise in the frame of quantum decision theory. In this section, we derive the general criterion for the occurrence of this effect.

Suppose a decision maker considers a lattice of just two prospects

$$\pi_n = A(L_n) \bigotimes B \qquad (n = 1, 2) , \qquad (11)$$

with the intention of choosing between them. Here $A(L_n)$ implies the action of choosing a lottery L_n. And B is a set incorporating uncertainties associated with this choice. Let one prefer the prospect π_1 against π_2, which means that

$$p(\pi_1) > p(\pi_2) \qquad (\pi_1 > \pi_2) . \qquad (12)$$

Taking into account the alternation property, we have

$$q(\pi_1) + q(\pi_2) = 0 . \qquad (13)$$

This tells us that the prospect π_1 is preferred to π_2 if and only if

$$f(\pi_2) - f(\pi_1) < 2q(\pi_1) . \qquad (14)$$

Now, assume that the decision maker plans to price the given lotteries, e.g., wishing to sell them. The lotteries remain the same as before. However, uncertainties in selling are of course different from those when choosing, hence, the uncertain set B', associated with selling, is different from the set B including uncertainties associated with choosing. Now, the decision maker evaluates the two different prospects

$$\pi_n = A(L_n) \bigotimes B' \qquad (n = 3, 4) , \qquad (15)$$

where $L_1 = L_3$ and $L_2 = L_4$.

Preference reversal implies that, contrary to the situation with choosing, now the decision maker evaluates higher the prospect π_4 compared to π_3, so that

$$p(\pi_3) < p(\pi_4) \qquad (\pi_3 < \pi_4) . \qquad (16)$$

In view of the alternation property

$$q(\pi_3) + q(\pi_4) = 0 , \qquad (17)$$

the preference of π_4 occurs only when

$$f(\pi_4) - f(\pi_3) < 2q(\pi_3) . \qquad (18)$$

Since the lotteries are the same ($L_1 = L_3$ and $L_2 = L_4$), their expected utilities are pairwise equal: $U(L_1) = U(L_3)$ and $U(L_2) = U(L_4)$. Therefore, the utility factors are also pairwise equal

$$f(\pi_1) = f(\pi_3) , \qquad f(\pi_2) = f(\pi_4) . \qquad (19)$$

Combining the above conditions, we obtain the *preference reversal criterion*:

$$2q(\pi_3) \; < \; f(\pi_2) - f(\pi_1) \; < \; 2q(\pi_1) . \qquad (20)$$

Let us stress that in classical decision making, where $q(\pi_1) = q(\pi_3) \equiv 0$, the inequalities (20) cannot hold, which means that it is impossible to suggest a self-consistent mathematical explanation of the preference reversal phenomenon in classical terms, which is in agreement with discussions by Tversky and Thaler (1990) and Tversky et al. (1990).

Criterion (20) not only explains the preference reversal phenomenon, but it also provides a quantitative estimate of how likely it may happen, as well as a posteriori confirmation of why it has happened. This is because the attraction factors are not just some additional arbitrary characteristics, but because their signs are prescribed by the risk aversion notion, while their values are constrained by conditions (8) – (10). Thus, due to risk aversion when facing several choices, the more certain lottery is more attractive, hence $q(\pi_1) > q(\pi_2)$, which, in view of the alternation property (9), implies that $q(\pi_1) > 0$, while $q(\pi_2) < 0$. Contrary to this, when pricing, risk aversion is absent, hence more attractive is the lottery that can provide the larger gain, so that $q(\pi_4) > q(\pi_3)$, which, again taking into account the alternation property (9), tells us that $q(\pi_3) < 0$ while $q(\pi_4) > 0$. Estimating the absolute values of the attraction factors by the quantity 0.25, which follows from the quarter law (10), we have the criterion

$$-\frac{1}{2} \; < \; f(\pi_2) - f(\pi_1) \; < \; \frac{1}{2} .$$

Therefore, if the given lotteries are such that their utility factors satisfy the above inequalities, we may expect that preference reversal can occur. And, vice versa, if preference reversal has happened, then the above inequalities must hold. Below we demonstrate that criterion (20) really provides a necessary and sufficient conditions for the preference reversal phenomenon.

4. Confirmation of Preference Reversal Criterion

To confirm the validity of the preference reversal criterion, let us test it with empirical data of decision-making experiments. We

shall consider pairs of lotteries with the notation of the previous section. The prospects, related to the choice between the lotteries L_1 and L_2, are denoted as π_1 and π_2, respectively. The prospects, corresponding to pricing of these lotteries, will be denoted by π_3 and π_4. The expected utility of a lottery $L = \{x_i, p(x_i)\}$, consisting of payoffs x_i, with their weights $p(x_i)$, will be calculated by the formula $U(L) = \sum_i x_i p(x_i)$. And the utility factors are given by expression (7).

Example 1. Let us start with the example given by Tversky and Thaler (1990). Consider two lotteries

$$L_1 = \left\{ 4, \frac{8}{9} \;|\; 0, \frac{1}{9} \right\} , \qquad L_2 = \left\{ 40, \frac{1}{9} \;|\; 0, \frac{8}{9} \right\} ,$$

whose payoffs 4 and 40 are given in some monetary units. The type of units, whether these are Dollars, or Euro, or Francs, is not of importance, since such units are canceled in definition (7) of utility factors. This is one of the advantage of employing the dimensionless utility factors that are invariant with respect to the type of payoff measures. The corresponding expected utilities

$$U(L_1) = \frac{32}{9} , \qquad U(L_2) = \frac{40}{9} ,$$

result in the utility factors

$$f(\pi_1) = \frac{4}{9} , \qquad f(\pi_2) = \frac{5}{9} ,$$

which show that the second lottery is more useful.

The experimental probabilities are defined as the fractions of subjects preferring the related lotteries. According to Tversky and Thaler (1990), in the case of choice, it was found that 71% of decision makers preferred the more certain lottery L_1, so that

$$p(\pi_1) = 0.71 \; > \; p(\pi_2) = 0.29 ,$$

despite that this lottery is less useful. In view of (4), this corresponds to the attraction factors

$$q(\pi_1) = 0.266 , \qquad q(\pi_2) = -0.266 .$$

However, when pricing, 67% of subjects found Lottery L_2 more valuable, so that

$$p(\pi_3) = 0.33 \; < \; p(\pi_4) = 0.67 ,$$

despite that the win in this lottery is less probable. The related attraction factors are

$$q(\pi_3) = -0.114 , \qquad q(\pi_4) = 0.114 .$$

Notice that, in the case of pricing, the attraction factor signs are reversed as compared to the case of choosing. This is in agreement with the probability-amount dichotomy (Kim et al., 2012): when choosing, one accepts as more attractive the lottery with a higher probability win, while when pricing, one treats as more attractive the lottery with a higher payoff amount. In the

process of pricing, decision makers usually are more pragmatic, evaluating higher the more useful lottery.

Combining the data of this experiment, the two inequalities (20) read

$$-0.228 < 0.111 < 0.452 \,,$$

which confirms the prediction of QDT.

Example 2. When there is no preference reversal, the criterion (20) does not hold. To illustrate this, let us consider an example treated by Tversky et al. (1990), taking the lotteries

$$L_1 = \{100, 0.97 \mid 0, 0.03\} \,, \qquad L_2 = \{400, 0.31 \mid 0, 0.69\} \,.$$

Their expected utilities are

$$U(L_1) = 97 \,, \qquad U(L_2) = 124 \,,$$

which yields the utility factors

$$f(\pi_1) = 0.439 \,, \qquad f(\pi_2) = 0.561 \,.$$

The first lottery is essentially more certain, and subjects overwhelmingly tend to prefer this lottery, so that

$$p(\pi_1) = 0.91 \;>\; p(\pi_2) = 0.09 \,.$$

According to (4), the related attractions factors are

$$q(\pi_1) = 0.471 \,, \qquad q(\pi_2) = -0.471 \,.$$

When pricing, subjects pay higher attention to the payoff amounts so that the fraction of decision makers preferring the first lottery is drastically reduced. However, the preference reversal does not occur *per se*, with the (more narrow) majority pricing the first lottery higher:

$$p(\pi_3) = 0.54 \;>\; p(\pi_4) = 0.46 \,.$$

The corresponding attraction factors are

$$q(\pi_3) = 0.101 \,, \qquad q(\pi_4) = -0.101 \,.$$

Since

$$f(\pi_2) - f(\pi_1) = 0.122 < 2q(\pi_3) = 0.202 \,,$$

criterion (20) is not fulfilled, which is the expected situation in absence of preference reversal.

This example demonstrates that, although in pricing, one pays a higher attention to payoff amounts, however, the focus is not exclusively on this amount. Probabilities can also influence decisions, together with amounts.

Example 3. Another example from Tversky et al. (1990) deals with the lotteries

$$L_1 = \{12, 0.92 \mid 0, 0.08\} \,, \qquad L_2 = \{175, 0.06 \mid 0, 0.94\} \,.$$

The first lottery is both more certain as well as more useful, with the expected utilities

$$U(L_1) = 11.04 \,, \qquad U(L_2) = 10.5$$

and the utility factors

$$f(\pi_1) = 0.513 \,, \qquad f(\pi_2) = 0.487 \,.$$

It is not surprising that, when choosing, decision makers prefer this lottery according to

$$p(\pi_1) = 0.81 \;>\; p(\pi_2) = 0.19 \,.$$

The related attraction factors are

$$q(\pi_1) = 0.297 \,, \qquad q(\pi_2) = -0.297 \,.$$

When pricing, subjects take into account that the second lottery can provide a much higher payoff, yet with too small a probability. As a result, the fraction of decision makers preferring the first lottery diminishes, but preference reversal does not happen:

$$p(\pi_3) = 0.58 \;>\; p(\pi_4) = 0.42 \,.$$

In pricing, the first lottery becomes less attractive than in choosing, but remains more attractive than the second lottery, with the attraction factors

$$q(\pi_3) = 0.067 \,, \qquad q(\pi_4) = -0.067 \,.$$

In view of the relations

$$f(\pi_2) - f(\pi_1) = -0.026 \;<\; 2q(\pi_3) = 0.134 \,,$$

criterion (20) does not hold, in agreement with the absence of preference reversal. Again, we see that payoff amounts as well as probabilities are considered in the process of pricing, although the role of payoff amounts, without doubt, is more important in pricing than in choosing.

We have also analyzed a large set of data presented by Tversky et al. (1990), demonstrating the effect of preference reversal. Pairs of lotteries were presented to 198 participants. In each pair, one of the lotteries, L_1, had a high probability, while the other, L_2, a higher payoff with lower probability. These lotteries are given in **Table 1**. In each lottery, the first number is a payoff and the next number is the probability of this payoff. A lottery is represented as a set $\{x, p(x)\}$, implying that one gets either the payoff x, with probability $p(x)$, or nothing, with probability $1 - p(x)$. The expected utilities and utility factors are shown. The first six lottery pairs include rather small payoffs. The following five pairs contain much larger payoffs by a factor of 25. And the last five pairs present a mixture of large and small payoffs. All the cases demonstrate the effect of preference reversal.

In **Table 2**, we show the prospect probabilities $p(\pi_1)$ and $p(\pi_3)$, with the corresponding attraction factors $q(\pi_1)$ and $q(\pi_3)$, demonstrating preference reversal, since $p(\pi_1) > p(\pi_2)$, although $p(\pi_3) < p(\pi_4)$. Those quantities that are not presented can be found from the relations

$$f(\pi_1) = f(\pi_3) \,, \qquad f(\pi_2) = f(\pi_4) \,,$$
$$p(\pi_2) = 1 - p(\pi_1) \,, \qquad p(\pi_4) = 1 - p(\pi_3) \,,$$
$$q(\pi_2) = -q(\pi_1) \,, \qquad q(\pi_4) = -q(\pi_3) \,.$$

TABLE 1 | Pairs of lotteries, with their expected utilities and utility factors.

L_1	L_2	$U(L_1)$	$U(L_2)$	$f(\pi_1)$	$f(\pi_2)$
4, 0.97	16, 0.31	3.88	4.96	0.439	0.561
2, 0.81	9, 0.19	1.62	1.71	0.486	0.514
3, 0.94	6.5, 0.50	2.82	3.25	0.465	0.535
4, 0.89	40, 0.11	3.56	4.4	0.447	0.553
2.5, 0.94	8.5, 0.39	2.35	3.315	0.415	0.585
2, 0.92	5, 0.50	1.84	2.5	0.424	0.576
50, 0.81	225, 0.19	40.5	42.75	0.486	0.514
75, 0.94	160, 0.50	70.5	80	0.468	0.532
100, 0.89	1000, 0.11	89	110	0.447	0.553
65, 0.94	210, 0.39	61.1	81.9	0.427	0.573
50, 0.92	125, 0.50	46	62.5	0.424	0.576
10, 0.78	100, 0.08	7.8	8	0.494	0.506
7, 0.69	40, 0.17	4.83	6.8	0.415	0.585
3, 0.86	13, 0.19	2.58	2.47	0.511	0.489
4, 0.94	150, 0.03	3.76	4.5	0.455	0.545
11, 0.89	135, 0.08	9.79	10.8	0.475	0.525

TABLE 2 | Probability $p(\pi_1)$ defined as the fraction of decision makers choosing the lottery L_1, and probability $p(\pi_3)$ defined as the fraction of subjects pricing the lottery L_1 higher.

$p(\pi_1)$	$p(\pi_3)$	$q(\pi_1)$	$q(\pi_3)$	$[f(\pi_2) - f(\pi_1)]/2$
0.83	0.26	0.391	−0.179	0.061
0.68	0.22	0.194	−0.266	0.014
0.71	0.30	0.245	−0.165	0.035
0.71	0.33	0.263	−0.117	0.053
0.73	0.17	0.315	−0.245	0.085
0.62	0.14	0.196	−0.284	0.076
0.86	0.48	0.374	−0.006	0.014
0.77	0.46	0.302	−0.008	0.032
0.84	0.47	0.393	0.023	0.053
0.82	0.48	0.393	0.053	0.073
0.70	0.32	0.276	−0.104	0.076
0.81	0.38	0.316	−0.114	0.006
0.68	0.21	0.265	−0.205	0.085
0.74	0.39	0.229	−0.121	−0.011
0.74	0.38	0.285	−0.075	0.045
0.79	0.46	0.315	−0.015	0.025

The corresponding attraction factors $q(\pi_1)$ and $q(\pi_3)$, and the combination $[f(\pi_2) - f(\pi_1)]/2$ that should be compared with those attraction factors according to criterion (20) obtained from QDT, which reads here $q(\pi_3) < [f(\pi_2) - f(\pi_1)]/2 < q(\pi_1)$.

We also show the value $[f(\pi_2) - f(\pi_1)]/2$ that has to be compared with $q(\pi_3)$ and $q(\pi_1)$ in order to check the validity of criterion (20). As is seen from **Table 2**, the preference reversal criterion (20) is always valid.

Since, in each pair of lotteries considered in the case of choosing or pricing, the utility factors do not change, the preference reversal effect can be interpreted within QDT as caused by the existence of the attraction factors. If one would evaluate the lotteries solely on the basis of rational utility,

no preference reversal would occur. However, preferences of decision makers involve irrational feelings and biases as well as other considerations not included in the utility, which are embodied in the attraction factors, accounting for the phenomenon of preference reversal. In order to characterize the deviation from rationality during decision making over a family of N trials, we can introduce the *irrationality measure*

$$\delta_j \equiv \frac{1}{N} \sum_{n=1}^{N} |q(\pi_j)| \, .$$

Then δ_1 measures the level of irrationality in the course of choosing, while δ_3 describes the degree of irrationality in the process of pricing. From **Table 2**, we find $\delta_1 = 0.299$ and $\delta_3 = 0.118$. Thus, people seem to be significantly more irrational when choosing, as compared to pricing. In other words, the evaluation of lotteries in pricing is more rational.

5. Discussion

We have shown that the phenomenon of preference reversal, which is treated as an anomaly in classical decision making, finds a natural explanation in the frame of quantum decision theory. In the latter, the preference probability consists of two terms, the utility factor quantifying the utility of a prospect, and the attraction factor characterizing behavioral biases of a decision maker. In that way, a prospect probability, defined as a quantum quantity, has the meaning of a behavioral probability taking into account both utility of the considered prospects, as well as their attractiveness for the decision maker, due to subconscious behavioral biases. We have formulated the criterion associated within QDT with preference reversal and we have illustrated its validity for a large set of empirical data.

We summarize the key steps of the logic we have followed.

1. We acknowledge the existence of risk aversion that leads human to prefer the more probable outcome ceteris paribus.
2. We formulate decisions in terms of QDT and derive the general fundamental expression (4) of QDT: $p = f + q$.
3. We interpret q as an "attraction factor" embodying the point resulting from risk aversion, which determines the sign of q.
4. The structure of QDT leads to criterion (20) for preference reversal to occur, which relates the utilities of the two lotteries to the attraction factors under choosing vs. pricing.
5. We showed that this criterion is verified by experiments.

We have thus demonstrated that QDT predicts the existence of two inequalities for the reversal to occur, that turn out to be confirmed.

It is worth noting that the effect of preference reversal does not only occur when choice is compared with pricing, but similar reversals can happen in other cases. As another illustration, we can mention the so-called planning paradox that can be represented by the following stylized example.

Suppose one is deliberating about stopping smoking. Let the imaginary plan to stop smoking be denoted as the prospect π_1, while continuing smoking corresponds to prospect π_2. The utility

of not smoking clearly overweights that of smoking because of evident health reasons. In contrast, the negative feelings, connected with addiction, are yet too imaginary to influence the mood of the decision maker. We thus expect that the related attraction factors should be rather small, so that the decision is based mainly on rational grounds. Hence, the preference in this plan π_1 is expressed by the inequality $p(\pi_1) > p(\pi_2)$, implying that the majority of subjects would like to stop smoking.

However, when one has to choose to really stop smoking now (but not in the future), then one actually meets another alternative: really stop smoking, which can be denoted as the prospect π_3, or continue smoking, the prospect π_4. Deciding whether to really stop smoking now, one immediately confronts negative feelings anticipating the suffering resulting from addiction. This translates into the appearance of a negative

attraction factor $q(\pi_3)$ devaluating the utility of not smoking. As a result, $p(\pi_3)$ becomes smaller than $p(\pi_4)$, which means that the majority of people do not really quit smoking.

This planning paradox gives a clear example of preference reversal, which cannot be understood in terms of classical utility considerations, since the utility of prospects does not change. But there is no paradox in quantum decision theory, where the effect of preference reversal is explained by the variation of attraction factors. Numerous data, collected by Walsh and Sanson-Fisher (2001) from the World Health Organization, confirm the robust existence of the preference reversal in the stop-smoking planning paradox. Thus the preference reversal is a rather general phenomenon that obtains a straightforward explanation in the framework of quantum decision theory.

References

Ashtiani, M., and Azgomi, M. A. (2015). A survey of quantum-like approaches to decision making and cognition. *Math. Soc. Sci.* 75, 49–50. doi: 10.1016/j.mathsocsci.2015.02.004

Bagarello, F. (2013). *Quantum Dynamics for Classical Systems*. Hoboken: Wiley.

Busemeyer, J. R., and Bruza, P. (2012). *Quantum Models of Cognition and Decision*. Cambridge: Cambridge University.

Busemeyer, J. R., Wang, Z., Khrennikov, A., and Basieva, I. (2014). Applying quantum principles to psychology. *Phys. Scripta* T163:014007. doi: 10.1088/0031-8949/2014/T163/014007

Goldstein, W. M., and Einhorn, H. J. (1987). Expression theory and the preference reversal phenomenon. *Psychol. Rev.* 94, 236–254. doi: 10.1037/0033-295X.94.2.236

Grether, D. M., and Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *Am. Econ. Rev.* 69, 623–638.

Haven, E., and Khrennikov, A. (2013). *Quantum Social Science*. Cambridge: Cambridge University.

Holt, C. A. (1986). Preference reversal and the independence axiom. *Am. Econ. Rev.* 76, 508–515.

Karni, E., and Safra, Z. (1987). Preference reversal and the observability of preferences by experimental methods. *Econometrica* 55, 675–685. doi: 10.2307/1913606

Kim, B. E., Seligman, D., and Kable, J. W. (2012). Preference reversals in decision making under risk are accompanied by changes in attention to different attributes. *Front. Neurosci.* 6:109. doi: 10.3389/fnins.2012.00109

Khrennikov, A. (2010). *Ubiquitous Quantum Structure*. Berlin: Springer.

Khrennikov, A., and Basieva, I. (2014). Possibility to agree on disagree from quantum information and decision making. *J. Math. Phsychol.* 62, 1–15. doi: 10.1016/j.jmp.2014.09.003

Lichtenstein, S., and Slovic, P. (1971). Reversal of preference between bids and choices in gambling decisions. *J. Exp. Psychol.* 89, 46–55. doi: 10.1037/h0031207

Lichtenstein, S., and Slovic, P. (1973). Response-induced reversals of preference in gambling. *J. Exp. Psychol.* 101, 16–20. doi: 10.1037/h0035472

Lindman, H. R. (1971). Inconsistent preferences among gambles. *J. Exp. Psychol.* 89, 390–397. doi: 10.1037/h0031208

Loomes, G., and Sugden, R. (1983). A rationale for preference reversal. *Am. Econ. Rev.* 73, 428–432.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York, NY: Wiley.

Schkade, D. A., and Johnson, E. J. (1989). Cognitive processes in preference reversals. *Org. Behav. Hum. Perform.* 44, 203–231. doi: 10.1016/0749-5978(89)90025-3

Segal, U. (1988). Does the preference reversal phenomenon necessarily contradicts the independence axiom? *Am. Econ. Rev.* 78, 233–236.

Slovic, P., and Lichtenstein, S. (1983). Preference reversals: a broader perspective. *Am. Econ. Rev.* 73, 596–605.

Sornette, D. (2014). Physics and financial economics (1776-2014): puzzles, Ising and agent-based models. *Rep. Prog. Phys.* 77:062001. doi: 10.1088/0034-4885/77/6/062001

Tversky, A., Sattath, S., and Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychol. Rev.* 95, 371–384. doi: 10.1037/0033-295X.95.3.371

Tversky, A., and Thaler, R. H. (1990). Anomalies: preference reversals. *J. Econ. Persp.* 4, 201–211. doi: 10.1257/jep.4.2.201

Tversky, A., Slovic, P., and Kahneman, D. (1990). The causes of preference reversal. *Am. Econ. Rev.* 80, 204–217.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton: Princeton University.

Walsh, R. A., and Sanson-Fisher, R. B. (2001). *Behavioral science learning modules: encouraging people to stop smoking*. Geneva: World Health Organization.

Yukalov, V. I., and Sornette, D. (2008). Quantum decision theory as quantum theory of measurement. *Phys. Lett. A* 372, 6867–6871. doi: 10.1016/j.physleta.2008.09.053

Yukalov, V. I., and Sornette, D. (2009a). Physics of risk and uncertainty in quantum decision making. *Eur. Phys. J. B* 71, 533–548. doi: 10.1140/epjb/e2009-00245-9

Yukalov, V. I., and Sornette, D. (2009b). Processing information in quantum decision theory. *Entropy* 11, 1073–1120. doi: 10.3390/e11041073

Yukalov, V. I., and Sornette, D. (2010). Mathematical structure of quantum decision theory. *Adv. Complex Syst.* 13, 659–698. doi: 10.1142/S0219525910002803

Yukalov, V. I., and Sornette, D. (2011). Decision theory with prospect interference and entanglement. *Theor. Decis.* 70, 283–328. doi: 10.1007/s11238-010-9202-y

Yukalov, V. I., and Sornette, D. (2013). Quantum probabilities of composite events in quantum measurements with multimode states. *Laser Phys.* 23:105502. doi: 10.1088/1054-660X/23/10/105502

Yukalov, V. I., and Sornette, D. (2014a). Conditions for quantum interference in cognitive sciences. *Top. Cogn. Sci.* 6, 79–90. doi: 10.1111/tops.12065

Yukalov, V. I., and Sornette, D. (2014b). Manipulating decision making of typical agents. *IEEE Trans. Syst. Man Cybern. Syst.* 44, 1155–1168. doi: 10.1109/TSMC.2014.2314283

Yukalov, V. I., and Sornette, D. (2015). Positive operator-valued measures in quantum decision theory. *Lect. Notes Comput. Sci.* 8951, 146–161. doi: 10.1007/978-3-319-15931-7_12

Noncontextuality with marginal selectivity in reconstructing mental architectures

Ru Zhang and Ehtibar N. Dzhafarov*

Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

We present a general theory of series-parallel mental architectures with selectively influenced stochastically non-independent components. A mental architecture is a hypothetical network of processes aimed at performing a task, of which we only observe the overall time it takes under variable parameters of the task. It is usually assumed that the network contains several processes selectively influenced by different experimental factors, and then the question is asked as to how these processes are arranged within the network, e.g., whether they are concurrent or sequential. One way of doing this is to consider the distribution functions for the overall processing time and compute certain linear combinations thereof (interaction contrasts). The theory of selective influences in psychology can be viewed as a special application of the interdisciplinary theory of (non)contextuality having its origins and main applications in quantum theory. In particular, lack of contextuality is equivalent to the existence of a "hidden" random entity of which all the random variables in play are functions. Consequently, for any given value of this common random entity, the processing times and their compositions (minima, maxima, or sums) become deterministic quantities. These quantities, in turn, can be treated as random variables with (shifted) Heaviside distribution functions, for which one can easily compute various linear combinations across different treatments, including interaction contrasts. This mathematical fact leads to a simple method, more general than the previously used ones, to investigate and characterize the interaction contrast for different types of series-parallel architectures.

Keywords: interaction contrast, mental architectures, noncontextuality, response time, selective influences, series-parallel network

1. Introduction

The notion of a network of mental processes with components selectively influenced by different experimental factors was introduced to psychology in Saul Sternberg's (1969) influential paper. Sternberg considered networks of processes $a, b, c, \ldots$ involved in performing a mental task. Denoting their respective durations by $A, B, C \ldots$, the hypothesis he considered was that the observed response time T is $A + B + C + \ldots$. One cannot test this hypothesis, Sternberg wrote, without assuming that there are some factors, $\alpha, \beta, \gamma, \ldots$, that selectively influence the durations $A, B, C \ldots$, respectively. Sternberg's analysis was confined to stochastically independent $A, B, C, \ldots$, and the consequences of the assumptions of seriality and selective influences were tested on the level of the mean response times only.

Subsequent development of these ideas was aimed at the entire distributions of the response times and at a greater diversity and complexity of mental architectures than just series of "stages." This development prominently includes Townsend (1984, 1990a,b); Schweickert and Townsend (1989); Townsend and Schweickert (1989); Roberts and Sternberg (1993); Townsend and Nozawa (1995); Schweickert et al. (2000), and several other publications, primarily by James Townsend and Richard Schweickert with colleagues. For an overview of these developments see Dzhafarov (2003) and Schweickert et al. (2012). In the present context we should separately mention the development of the ideas of *stochastic ordering* of processing times in Townsend (1984, 1990a) and Townsend and Schweickert (1989); as well as the idea of *marginal selectivity* (Townsend and Schweickert, 1989).

The notion of selective influences also underwent a significant development, having been generalized from stochastically independent random variables to arbitrarily interdependent ones (Dzhafarov, 2003; Dzhafarov and Gluhovsky, 2006; Kujala and Dzhafarov, 2008; Dzhafarov and Kujala, 2010, in press). The essence of the development is easy to understand using two random variables (e.g., process durations) A, B selectively influenced by two respective factors α, β. In Dzhafarov's (2003) notation, this is written $(A, B) \leftrightarrow (\alpha, \beta)$. According to the definition given in Dzhafarov (2003), this means that there are functions f and g and a random variable R (a *common source of randomness*) such that $f(\alpha, R) = A$ and $g(\beta, R) = B$. If such a choice of (f, g, R) exists, it is not unique. For instance, R can always be chosen to have any distribution that is absolutely continuous with respect to the usual Borel measure on the real line (e.g., a standard uniform, or standard normal distribution, see Dzhafarov and Gluhovsky, 2006). However, a triple (f, g, R) need not exist. It does not exist, e.g., if marginal selectivity (Townsend and Schweickert, 1989) is violated, i.e., if the distribution of, say, A at a given value of α changes in response to changing β. But marginal selectivity is not sufficient for the existence of a triple (f, g, R). Let, e.g., α and β be binary factors, with values 1, 2 each, and let the correlation ρ between A and B for a treatment (α, β) be denoted $\rho_{\alpha\beta}$. Then the triple in question does not exist if the correlations violated the "cosphericity test" (Kujala and Dzhafarov, 2008), also known in quantum mechanics as Landau's inequality (Landau, 1988):

$$|\rho_{11}\rho_{12} - \rho_{21}\rho_{22}| \leq \overline{\rho}_{11}\overline{\rho}_{12} + \overline{\rho}_{21}\overline{\rho}_{22}, \tag{1}$$

where $\overline{\rho}_{\alpha\beta} = \sqrt{1 - \rho_{\alpha\beta}^2}$. There are many other known conditions that must be satisfied for the existence of a triple (f, g, R) when marginal selectivity is satisfied (Dzhafarov and Kujala, 2010, 2012a,b, 2013, 2014a).

The allusion to quantum mechanics is not accidental: as shown in Dzhafarov and Kujala (2012a,b), the theory of selective influences in psychology can be viewed as a special application of the theory of (non)contextuality. This theory is interdisciplinary (Khrennikov, 2009; Dzhafarov and Kujala, 2014b,c,d), but its origins are in quantum theory, dating from Kochen and Specker (1967) and Bell's (1964, 1966) celebrated work. For the modern state of the theory see Dzhafarov et al. (in press). A simplified

account of the (non)contextuality analysis of the example given above is as follows. One labels each random variable in play *contextually*, i.e., by what property is being measured/recorded under what treatment (context):

$$\left(\underbrace{A_{\text{value of }\alpha}}_{\text{property: what is measured}} \right) \overbrace{\underbrace{(\text{value of }\alpha, \text{value of }\beta)}_{\text{context: under what treatment}}}, $$

$$\left(\underbrace{B_{\text{value of }\beta}}_{\text{property: what is measured}} \right) \overbrace{\underbrace{(\text{value of }\alpha, \text{value of }\beta)}_{\text{context: under what treatment}}}. \tag{2}$$

The notation here is, of course, redundant, because the context and property identifiers overlap, but we need now to emphasize the logic rather than achieve notational convenience. Once the labeling is done, one looks at all possible joint distributions imposable on all these random variables, for all properties and all treatments. A system is noncontextual if there exists such a joint distributions in which any two random variables that represent the same property ("what is measured") are equal with probability 1. The latter is possible only if the random variables representing the same property always have the same distribution: in our case

$$(A_\alpha)^{(\alpha,\beta)} \sim (A_\alpha)^{(\alpha,\beta')}, (B_\beta)^{(\alpha,\beta)} \sim (B_\beta)^{(\alpha',\beta)} \tag{3}$$

for any values $\alpha, \beta, \alpha', \beta'$ of the two factors. This is called consistent connectedness (Dzhafarov et al., in press), and in physics is known under a variety of names, including (in certain paradigms) "no-signaling condition" (Popescu and Rohrlich, 1994; Cereceda, 2000; Masanes et al., 2006). In psychology, this is marginal selectivity. The definition of noncontextuality just given is not the most general one, as the notion of contextuality can be extended to inconsistently connected (violating marginal selectivity) systems (Dzhafarov et al., in press), but we do not need this generality in this paper. What is important for us here is that the existence of a joint distribution mentioned in our definition is equivalent to the existence of a random variable R and the functions f, g mentioned in the introductory paragraph.

It is easy to show (Dzhafarov, 2003) that the existence of a triple (f, g, R) for given joint distributions of (A, B) under different treatments (α, β) is equivalent to the existence of a quintuple (f', g', S, S_A, S_B), where S, S_A, S_B are random variables, such that $f'(\alpha, S, S_A) = A$ and $g'(\beta, S, S_B) = B$. In such a representation, one can speak of a common source of randomness S and specific sources of randomness S_A, S_B. In Dzhafarov et al. (2004) this representation was used to investigate different series-parallel arrangements of the hypothetical durations A and B. The reason this representation has been considered convenient is that if one fixes the value $S = s$, then $f'(\alpha, s, S_A) = A_c$ and $g'(\beta, s, S_B) = B_c$ are stochastically independent random variables. One can therefore use theorems proved for stochastically independent selectively influenced components (Schweickert et al., 2000) to obtain a general result by averaging across possible values of s. For

instance, let α, β be binary factors (with values 1, 2 each), and let us assume that the observed duration $T_{\alpha\beta}$ is $\min\left(A_\alpha, B_\beta\right)$ for every treatment (α, β). Then $T_{\alpha\beta s} = \min\left(A_{\alpha s}, B_{\beta s}\right)$ for every value $S = s$, and it is known that, for the independent $A_{\alpha s}, B_{\beta s}$ (satisfying a prolongation condition, as explained below),

$$\Pr\left(T_{11s} \leq t\right) - \Pr\left(T_{12s} \leq t\right) - \Pr\left(T_{21s} \leq t\right) + \Pr\left(T_{22s} \leq t\right) \leq 0. \quad (4)$$

Since this should be true for every value $S = s$, then it should also be true that

$$C(t) = \Pr\left(T_{11} \leq t\right) - \Pr\left(T_{12} \leq t\right) - \Pr\left(T_{21} \leq t\right) \\ + \Pr\left(T_{22} \leq t\right) \leq 0. \quad (5)$$

This follows from the fact that

$$\Pr\left(T_{\alpha\beta} \leq t\right) = \int \Pr\left(T_{\alpha\beta s} \leq t\right) \mathrm{dm}(s), \quad (6)$$

where m (s) is the probability measure for S, and the integration is over the space of all possible s. The linear combination $C(t)$ in (5) is called the *interaction contrast of distributions functions*.

The Prolongation Assumption used in Dzhafarov et al. (2004), and derived from Townsend (1984, 1990a) and Townsend and Schweickert (1989), is that, for every $S = s$,

$$\Pr\left(A_{1s} \leq t\right) \geq \Pr\left(A_{2s} \leq t\right), \ \Pr\left(B_{1s} \leq t\right) \geq \Pr\left(B_{2s} \leq t\right). \quad (7)$$

For this particular architecture, $T = \min(A, B)$, this is the only assumption needed. To prove analogous results for more complex mental architectures, however, one needs additional assumptions, such as the existence of density functions for $A_{\alpha s}, B_{\beta s}$ at every s, and even certain ordering of these density functions in some vicinity $[0, \tau]$.

The same results, however, can be obtained without these additional assumptions, if one adopts the other, equivalent definition of selective influences: $f(\alpha, R) = A$ and $g(\beta, R) = B$, for some triple (f, g, R). If such a representation exists, then

$$a_{\alpha r} = f(\alpha, r), \ b_{\beta r} = g(\beta, r) \quad (8)$$

are deterministic quantities (real numbers), for every value $R = r$. Any real number x in turn can be viewed as a random variable whose distribution function is a shifted Heaviside function

$$h(t - x) = \begin{cases} 0, & \text{if } t < x, \\ 1, & \text{if } t \geq x. \end{cases} \quad (9)$$

In particular, the quantity $t_{\alpha\beta r} = \min\left(a_{\alpha r}, b_{\beta r}\right)$ for the simple architecture $T = \min(A, B)$ considered above is distributed according to

$$h\left(t - t_{\alpha\beta r}\right) = h_{\alpha\beta r}(t). \quad (10)$$

Let us see how inequality (5) can be derived using these observations.

We first formulate the *(conditional) Prolongation Assumption*, a deterministic version of (7): the assumption is that f, g, R can be so chosen that for every $R = r$,

$$a_{1r} \leq a_{2r}, \ b_{1r} \leq b_{2r}. \quad (11)$$

Without loss of generality, we can also assume, for any given r,

$$a_{1r} \leq b_{1r} \quad (12)$$

(if not, rename a into b and vice versa).

Remark 1.1. The Prolongation Assumption clearly implies (7). Conversely, if (7) holds, one can always find functions f, g, R for which the Prolongation Assumption holds in the form above. For instance, one can choose $R = (S, S_A, S_B)$, take S_A and S_B to be uniformly distributed between 0 and 1, and choose $f(\alpha, \ldots), g(\beta, \ldots)$ to be the quantile functions for the hypothetical distributions of A and B at the corresponding factor levels.

We next form the conditional (i.e., conditioned on $R = r$) interaction contrast

$$c_r(t) = h_{11r}(t) - h_{12r}(t) - h_{21r}(t) + h_{22r}(t). \quad (13)$$

Notation Convention. When r is fixed throughout a discussion, we omit this argument and write a_α, b_β, $t_{\alpha\beta}$, $h_{\alpha\beta}(t)$, $c(t)$ in place of $a_{\alpha r}$, $b_{\beta r}$, $t_{\alpha\beta r}$, $h_{\alpha\beta r}(t)$, $c_r(t)$. (For binary factors α, β, we also conveniently replace α, β in indexation with i, j.)

Following this convention, there are three different arrangements of a_1, a_2, b_1, b_2 (for a given $R = r$) satisfying (11)–(12):

$$\begin{aligned} &\text{(i)} \quad a_1 \leq b_1 \leq a_2 \leq b_2 \\ &\text{(ii)} \quad a_1 \leq a_2 \leq b_1 \leq b_2 \quad (14) \\ &\text{(iii)} \quad a_1 \leq b_1 \leq b_2 \leq a_2 \end{aligned}$$

In all three cases,

$$t_{11} = \min\left(a_1, b_1\right) = a_1 = \min\left(a_1, b_2\right) = t_{12}. \quad (15)$$

For arrangement (i) we have

$$\begin{array}{ll}
+h_{11}(t) = 1 & +h_{11}(t) = 1 \\
-h_{12}(t) = -1 & -h_{12}(t) = -1 \\
-h_{21}(t) = -0 & -h_{21}(t) = -1 \\
+h_{22}(t) = 0 & +h_{22}(t) = 0 \\
= c(t) = 0 & = c(t) = -1
\end{array}$$

$$
t_{11} = t_{12} = a_1 \qquad \leq t < \qquad t_{21} = b_1 \qquad \leq t < \qquad t_{22} = a_2
$$
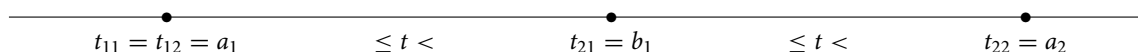
This diagram shows the values of $h_{ijr}(t)$ and the resulting values of $c_r(t)$ as t changes with respect to the fixed positions of t_{ijr} (with index r dropped everywhere). Analogously, for arrangements (ii) and (iii), we have, respectively

$$c(t) = 0$$

$$t_{11} = t_{12} = a_1 \qquad \leq t < \qquad t_{21} = t_{22} = a_2$$

and

$$c(t) = 0 \qquad\qquad c(t) < 0$$

$$t_{11} = t_{12} = a_1 \qquad \leq t < \qquad t_{21} = b_1 \qquad \leq t < \qquad t_{22} = b_2$$
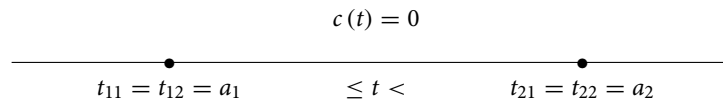
In all three cases, $c(t)$ is obviously zero for $t < t_{11}$ and $t \geq t_{22}$. We see that $c(t) = c_r(t) \leq 0$ for all t and every $R = r$. It follows that $C(t) \leq 0$, because

$$\Pr\left(T_{ij} \leq t\right) = \int h_{ijr}(t) \, d\mu(r), \tag{16}$$

for $i, j \in \{1, 2\}$, and

$$C(t) = \int c_r(t) \, d\mu(r) \leq 0, \tag{17}$$
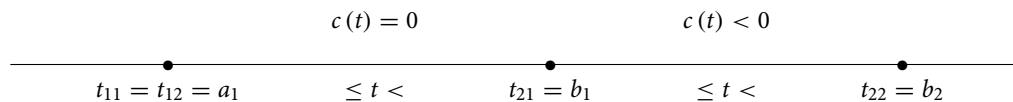
where μ is the probability measure associated with R and the integration is over all possible r. We obtain the same result as in (5), but in a very different way.

In this paper we extend this approach to other mental architectures belonging to the class of series-parallel networks, those involving other composition operations and possibly more than just two selectively influenced processes. In doing so we follow a long trail of work mentioned earlier. When dealing with multiple processes we follow Yang et al. (2013) in using high-order interaction contrasts. All our results are replications or straightforward generalizations of the results already known: the primary value of our work therefore is not in characterizing mental architectures, but rather in demonstrating a new theoretical approach and a new proof technique.

1.1. Definitions, Terminology, and Notation

Since we deal with the durations of processes rather than the processes themselves, we use the term *composition* to describe a function that relates the durations of the components of a network to the overall (observed) duration. Formally, a composition is a real-valued function $t = t(a, b, \ldots, z)$ of an arbitrary number of real-valued arguments. The arguments $a, b, \ldots, z$ are referred to as *durations* or *components*. In this article, we will use $X \wedge Y \wedge \ldots \wedge Z$ to denote $\min(X, Y, \ldots, Z)$, and $X \vee Y \vee \ldots \vee Z$ to denote $\max(X, Y, \ldots, Z)$.

A *series-parallel composition* (SP) is defined as follows.

Definition 1.2. (1) A single duration is an SP composition. (2) If X and Y are SP compositions with disjoint sets of arguments, then $X \wedge Y$, $X \vee Y$, and $X + Y$ are SP compositions. (3) There are no other SP compositions than those construable by Rules 1 and 2.

Remark 1.3. The requirement that X and Y in Rule 2 have disjoint sets of arguments prevents expressions like $X \wedge X$ or $X + X \vee Y$. But if the second X in $X \wedge X$ is renamed into X', or $X \vee Y$ in $X + X \vee Y$ is renamed into Z, then the resulting $X \wedge X'$ and $X + Z$ are legitimate SP compositions. This follows from the generality of our treatment, in which different components of an SP composition may have arbitrary joint distributions: e.g., X and X' in $X \wedge X'$ may very well be jointly distributed so that $\Pr\left[X = X'\right] = 1$. One should, however, always keep in minds the pattern of selective influences: thus, if X is influenced by α, then Z is also influenced by α in $X + Z$ above.

Any SP composition is obtained by a successive application of Rules 1 and 2 (the sequence being generally non-unique), and at any intermediate stage of such a sequence we also have an SP composition that we can term a *subcomposition*.

Definition 1.4. Two durations X, Y in an SP composition are said to be *parallel* or *concurrent* if there is a subcomposition of this SP composition of the form $\mathrm{SP}^1\left(X, X', \ldots\right) \wedge \mathrm{SP}^2\left(Y, Y', \ldots\right)$ (in which case X, Y are said to be *min-parallel*) or $\mathrm{SP}^1\left(X, X', \ldots\right) \vee \mathrm{SP}^2\left(Y, Y', \ldots\right)$ (X, Y are *max-parallel*). X, Y in an SP composition are said to be *sequential* or *serial* if there is a subcomposition of this SP composition of the form $\mathrm{SP}^1\left(X, X', \ldots\right) + \mathrm{SP}^2\left(Y, Y', \ldots\right)$.

Definition 1.5. An SP composition is called *homogeneous* if it does not contain both $\wedge$ and $\vee$ in it; if it does not contain $\wedge$, it is denoted $\mathrm{SP}_\vee$; if it does not contain $\vee$, it is denoted $\mathrm{SP}_\wedge$.

The only SP composition that is both $\mathrm{SP}_\wedge$ and $\mathrm{SP}_\vee$ is a purely serial one: $a + b + \ldots + z$. Most of the results previously obtained for mental networks are confined to homogeneous compositions. We will not need this constraint for the most part.

Since we will be dealing with compositions of more than just two components, we need to extend the definition of selective influences mentioned above. In the formulation below, $\sim$ stands for "has the same distribution as." A treatment $\phi = \left(\lambda_{i_1}^1, \ldots, \lambda_{i_n}^n\right)$ is a vector of values of the factors $\lambda^1, \ldots, \lambda^n$, the values of λ^k ($k = 1, \ldots, n$) being indicated by subscripts, $\lambda_{i_k}^k$.

Definition 1.6. Random variables $(X^1, \ldots, X^n)$ are *selectively influenced* by factors $(\lambda^1, \ldots, \lambda^n)$, respectively,

$$(X^1, \ldots, X^n) \leftarrowtail (\lambda^1, \ldots, \lambda^n), \tag{18}$$

if for some random variable R, whose distribution does not depend on $(\lambda^1, \ldots \lambda^n)$, and for some functions $g_1, \ldots, g_n$,

$$(X_\phi^1, \ldots, X_\phi^n) \sim (g_1(\lambda_{i_1}^1, R), \ldots, g_n(\lambda_{i_n}^n, R)), \tag{19}$$

for any treatment $\phi = \left(\lambda_{i_1}^1, \ldots, \lambda_{i_n}^n\right)$.

In the subsequent discussion we assume that all non-dummy factors involved are binary in a completely crossed design (i.e., the overall time T is recorded for all 2^n vectors of values for ϕ). When we have random variables not influenced by any of these factors, we will say they selectively influenced by an empty set of factors (we could also, equivalently, introduce for them dummy factors, with one value each).

2. SP Compositions Containing Two Selectively Influenced Processes

Consider two processes, with durations A and B in an SP composition. The overall duration of this SP composition can be written as a function of A, B and other components: $T = T(A, B, \ldots)$. We assume that A, B, and all other components are selectively influenced by α, β, and empty set, respectively: $(A, B, \ldots) \leftarrowtail (\alpha, \beta, \emptyset)$. Let each factor has two levels: $\alpha = 1, 2$ and $\beta = 1, 2$, with four allowable treatments $(1, 1)$, $(1, 2)$, $(2, 1)$, and $(2, 2)$. The corresponding overall durations (random variables) are written as T_{11}, T_{12}, T_{21}, and T_{22}.

By Definition 1.6 of selective influences, each process duration (a random variable) is a function of some random variable R and the corresponding factor: $A = a(\alpha, R)$, $B = b(\beta, R)$. For any given value $R = r$, the component durations are fixed numbers,

$$
\begin{array}{ll}
a(\alpha = 1, r) = a_{1r}, & a(\alpha = 2, r) = a_{2r}, \\
b(\beta = 1, r) = b_{1r}, & b(\beta = 2, r) = b_{2r}, \\
\multicolumn{2}{c}{x(\emptyset, r) = x_r,}
\end{array} \tag{20}
$$

where x is the value of any duration X in the composition other than A and B. We assume that R is chosen so that the Prolongation Assumption (11) holds, with the convention (12).

The overall duration T at $R = r$ is also a fixed number, written as (recall that we replace α, β in indexation with i, j)

$$T(a_{ir}, \beta_{jr}, \ldots) = t_{ijr}, \quad i, j \in \{1, 2\}. \tag{21}$$

The distribution function for t_{ijr} is the shifted Heaviside function $h_{ijr}(t) = h(t - t_{ijr})$,

$$\tag{22}$$
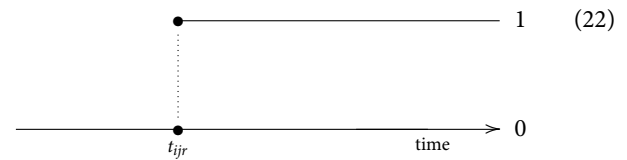
The *conditional interaction contrast* $c_r(t)$ is defined by (13). Denoting by $H_{ij}(t)$ the distribution function of T_{ij}, we have

$$H_{ij}(t) = \int_{\mathcal{R}} h_{ijr}(t) \, d\mu_r, \tag{23}$$

with $\mathcal{R}$ denoting the set of possible values of R. For the observable (i.e., estimable from data) interaction contrast

$$C(t) = H_{11}(t) - H_{12}(t) - H_{21}(t) + H_{22}(t), \tag{24}$$

we have then

$$C(t) = \int_{\mathcal{R}} c_r(t) \, d\mu_r. \tag{25}$$

Note that it follows from our Prolongation Assumption that

$$
\begin{aligned}
H_{11}(t) &\geq H_{12}(t), & H_{21}(t) &\geq H_{22}(t), \\
H_{11}(t) &\geq H_{21}(t), & H_{12}(t) &\geq H_{22}(t).
\end{aligned} \tag{26}
$$

We also define two conditional cumulative interaction contrasts (conditioned on $R = r$):

$$c(0, t) = \int_0^t c(\tau) \, d\tau. \tag{27}$$

$$c(t, \infty) = \int_t^\infty c(\tau) \, d\tau = \lim_{u \to \infty} \int_t^u c(\tau) \, d\tau. \tag{28}$$

The corresponding observable cumulative interaction contrasts are

$$
\begin{aligned}
C(0, t) &= \int_{\mathcal{R}} c(0, t) \, d\mu_r = \int_{\mathcal{R}} \left(\int_0^t c(\tau) \, d\tau \right) d\mu_r \\
&= \int_0^t \left(\int_{\mathcal{R}} c(\tau) \, d\mu_r \right) d\tau = \int_0^t C(\tau) \, d\tau.
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
C(t, \infty) &= \int_{\mathcal{R}} c(t, \infty) \, d\mu_r = \int_{\mathcal{R}} \left(\int_t^\infty c(\tau) \, d\tau \right) d\mu_r \\
&= \int_t^\infty \left(\int_{\mathcal{R}} c(\tau) \, d\mu_r \right) d\tau = \int_t^\infty C(\tau) \, d\tau.
\end{aligned} \tag{30}
$$

In these formulas we could switch the order of integration by Fubini's theorem, because, for any interval of reals I,

$$\int_{I \times \mathcal{R}} |c(\tau)| \, d(\tau \times \mu_r) \leq \int_{I \times \mathcal{R}} 2 \, d(\tau \times \mu_r) \leq 2. \tag{31}$$

2.1. Four lemmas

Recall the definition of $c_r(t)$ in (13). We follow our Notation Convention and drop the index r in $c_r(t)$ and all other expressions for a fixed r.

Lemma 2.1. *In any* SP *architecture, for any r,*

$$t_{11} \leq t_{12} \wedge t_{21} \leq t_{12} \vee t_{21} \leq t_{22}.$$

Proof. Follows from the (nonstrict) monotonicity of the SP composition in all arguments. □

Lemma 2.2. *In any* SP *architecture, for any r, $c(t)$ equals 0 for all values of t except for two cases:*
(Case$^+$) *if $t_{11} \leq t < t_{12} \wedge t_{21}$, then $c(t) = 1 - 0 - 0 + 0 > 0$, and*
(Case$^-$) *if $t_{12} \vee t_{21} \leq t < t_{22}$, then $c(t) = 1 - 1 - 1 + 0 < 0$.*

Proof. By direct computation. □

Lemma 2.3. *In any* SP *architecture, for any r, $c(t) \leq 0$ for all values of t if and only if $t_{11} = t_{12} \wedge t_{21}$; $c(t) \geq 0$ for all values of t if and only if $t_{12} \vee t_{21} = t_{22}$.*

Proof. Immediately follows from Lemma 2.2. □

Lemma 2.4. *In any* SP *architecture, for any r,*
(i) $c(0, t) = \int_0^t c(\tau)\, d\tau \geq 0$ *for any t if and only if $-t_{11} + t_{12} + t_{21} - t_{22} \geq 0$,*
(ii) $c(t, \infty) = \int_t^\infty c(\tau)\, d\tau \leq 0$ *for any t if and only if $-t_{11} + t_{12} + t_{21} - t_{22} \leq 0$,*
(iii) $\lim_{t \to \infty} c(0, t) = 0$ *if and only if $-t_{11} + t_{12} + t_{21} - t_{22} = 0$.*
(iv) $\lim_{t \to 0} c(t, \infty) = 0$ *if and only if $-t_{11} + t_{12} + t_{21} - t_{22} = 0$.*

Proof. Without loss of generality, put $t_{12} \leq t_{21}$. We have

$$c(0, t) = \begin{cases} 0 & \text{if } t < t_{11} \\ (t - t_{11}) & \text{if } t_{11} \leq t < t_{12} \\ (t - t_{11}) - (t - t_{12}) = t_{12} - t_{11} & \text{if } t_{12} \leq t < t_{21} \\ (t - t_{11}) - (t - t_{12}) - (t - t_{21}) \\ \qquad = -t_{11} + t_{12} + t_{21} - t & \text{if } t_{21} \leq t < t_{22} \\ (t - t_{11}) - (t - t_{12}) - (t - t_{21}) + (t - t_{22}) \\ \qquad = -t_{11} + t_{12} + t_{21} - t_{22} & \text{if } t \geq t_{22} \end{cases}$$

The expressions for the first three cases are obviously nonnegative. If $-t_{11} + t_{12} + t_{21} - t_{22} \geq 0$, then $c(0, t) \geq 0$ for all t in the last case ($t \geq t_{22}$). With $-t_{11} + t_{12} + t_{21} - t_{22} \geq 0$, we have $-t_{11} + t_{12} + t_{21} - t \geq t_{22} - t \geq 0$ for the fourth case ($t_{21} \leq t < t_{22}$). Hence $c(0, t) \geq 0$ for all t if $-t_{11} + t_{12} + t_{21} - t_{22} \geq 0$. Conversely, if $c(0, t) \geq 0$ for all t, then it is also true for $t \geq t_{22}$, whence $-t_{11} + t_{12} + t_{21} - t_{22} \geq 0$.

The proof for $c(t, \infty) = \int_t^\infty c(\tau)\, d\tau$ requires replacing it first with $\int_t^u c(\tau)\, d\tau \leq 0$ for some $u > t_{22}$. We have

$$\int_t^u c(\tau)\, d\tau =$$
$$\begin{cases} (u - t_{11}) - (u - t_{12}) - (u - t_{21}) + (u - t_{22}) \\ \qquad = -t_{11} + t_{12} + t_{21} - t_{22} & \text{if } t < t_{11} \\ (u - t) - (u - t_{12}) - (u - t_{21}) + (u - t_{22}) \\ \qquad = -t + t_{12} + t_{21} - t_{22} & \text{if } t_{11} \leq t < t_{12} \\ (u - t) - (u - t) - (u - t_{21}) + (u - t_{22}) \\ \qquad = t_{21} - t_{22} & \text{if } t_{12} \leq t < t_{21} \\ (u - t) - (u - t) - (u - t) + (u - t_{22}) \\ \qquad = t - t_{22} & \text{if } t_{21} \leq t < t_{22} \\ (u - t) - (u - t) - (u - t) + (u - t) \\ \qquad = 0 & \text{if } t \geq t_{22} \end{cases}$$

The expressions for the last three cases are obviously nonpositive. If $-t_{11} + t_{12} + t_{21} - t_{22} \leq 0$, then $\int_t^u c^{(2)}(\tau)\, d\tau \leq 0$ for all t in the first case ($t < t_{11}$). With $-t_{11} + t_{12} + t_{21} - t_{22} \leq 0$, we have $-t + t_{12} + t_{21} - t_{22} \leq t_{11} - t < 0$ for the second case ($t_{11} \leq t < t_{12}$). Hence $\int_t^u c(\tau)\, d\tau \leq 0$ for all t if $-t_{11} + t_{12} + t_{21} - t_{22} \leq 0$. Since in all expressions u is algebraically eliminated, they remain unchanged as $u \to \infty$. Conversely, if $c(t, \infty) \leq 0$ for all t, then it is also true for $t < t_{11}$, whence $-t_{11} + t_{12} + t_{21} - t_{22} \leq 0$.

The statements (iii) and (iv) follow trivially. □

2.2. Parallel Processes
2.2.1. Simple Parallel Architectures of Size 2
A simple parallel architecture corresponds to one of the two compositions: $T = A \wedge B$ or $T = A \vee B$, with $(A, B) \hookleftarrow_P (\alpha, \beta)$. Recall the definition of $C(t)$ in (24).

Theorem 2.5. *For $T = A \wedge B$, we have $c(t) \leq 0$ for any r, t; consequently, $C(t) \leq 0$ for any t. For $T = A \vee B$, we have $c(t) \geq 0$ for any r, t; consequently, $C(t) \geq 0$ for any t.*

Proof. For $T = A \wedge B$ with the Prolongation Assumption (11)–(12), we have

$$t_{11} = a_1 \wedge b_1 = a_1, \quad t_{12} = a_1 \wedge b_2, \quad t_{21} = a_2 \wedge b_1.$$

It follows that

$$t_{12} \wedge t_{21} = a_1 \wedge b_2 \wedge a_2 \wedge b_1 = a_1 = t_{11}.$$

By Lemma 2.3, $c(t) \leq 0$. As $C(t)$ in (25) preserves the sign of $c(t)$, we have $C(t) \leq 0$. For $T = A \wedge B$, we have

$$t_{22} = a_2 \vee b_2, \quad t_{12} = a_1 \vee b_2, \quad t_{21} = a_2 \vee b_1.$$

It follows that

$$t_{12} \vee t_{21} = a_1 \vee b_2 \vee a_2 \vee b_1 = t_{22},$$

whence, by Lemma 2.3, $c(t) \geq 0$ and therefore $C(t) \geq 0$. □

2.2.2. Two Parallel Processes in an Arbitrary SP Network

Consider now a composition $\mathrm{SP}(A, B, \ldots)$ with $(A, B, \ldots) \leftarrow\!\!\!\!\!\!\!\!\raise1pt\hbox{$\scriptscriptstyle\nearrow$}\; (\alpha, \beta, \emptyset)$.

Lemma 2.6. *If A, B in $\mathrm{SP}(A, B, \ldots)$ are parallel, then $\mathrm{SP}(A, B, \ldots)$ can be presented as $A' \wedge B'$ if they are min-parallel, or as $A' \vee B'$ if they are max-parallel, so that $(A', B') \leftarrow\!\!\!\!\!\!\!\!\raise1pt\hbox{$\scriptscriptstyle\nearrow$}\; (\alpha, \beta)$ and, for any fixed $R = r$, the Prolongation Assumption holds.*

Proof. By Definitions 1.2 and 1.4, if A, B are min-parallel, then $\mathrm{SP}_\wedge(A, B, \ldots)$ can be presented either as

$$\mathrm{SP}^1(A, \ldots) \wedge \mathrm{SP}^2(B, \ldots)$$

or

$$\left(\mathrm{SP}^1(A, \ldots) \wedge \mathrm{SP}^2(B, \ldots) + X\right) \wedge Y,$$

or else

$$\left(\mathrm{SP}^1(A, \ldots) \wedge \mathrm{SP}^2(B, \ldots) \wedge X\right) + Y,$$

where B does not enter in SP^1 and A does not enter in SP^2. On renaming

$$\underbrace{\mathrm{SP}^1(A, \ldots)}_{=A'} \wedge \underbrace{\mathrm{SP}^2(B, \ldots)}_{=B'},$$

$$\left(\mathrm{SP}^1(A, \ldots) \wedge \mathrm{SP}^2(B, \ldots) + X\right) \wedge Y$$
$$= \underbrace{\left(\mathrm{SP}^1(A, \ldots) + X\right)}_{=A'} \wedge \underbrace{\left(\mathrm{SP}^2(B, \ldots) + X\right) \wedge Y}_{=B'},$$

and

$$\left(\mathrm{SP}^1(A, \ldots) \wedge \mathrm{SP}^2(B, \ldots) \wedge X\right) + Y$$
$$= \underbrace{\left(\mathrm{SP}^1(A, \ldots) + Y\right)}_{=A'} \wedge \underbrace{\left(\mathrm{SP}^2(B, \ldots) \wedge X + Y\right)}_{=B'},$$

we have, obviously, $(A', B') \leftarrow\!\!\!\!\!\!\!\!\raise1pt\hbox{$\scriptscriptstyle\nearrow$}\; (\alpha, \beta)$. Fixing $R = r$, by the (nonstrict) monotonicity of SP compositions,

$$a_1' = \mathrm{SP}^1(a_1, \ldots) \leq \mathrm{SP}^1(a_2, \ldots) = a_2'$$

and

$$b_1' = \mathrm{SP}^2(b_1, \ldots) \leq \mathrm{SP}^2(b_2, \ldots) = b_2'$$

We can also put $a_1' = \mathrm{SP}^1(a_1, \ldots) \leq \mathrm{SP}^2(b_1, \ldots) = b_1'$ (otherwise we can rename the variables). The proof for the max-parallel case is analogous. $\square$

Theorem 2.7. *If A, B in $\mathrm{SP}(A, B, \ldots)$ are min-parallel, then $c(t) \leq 0$ for any r, t; consequently, $C(t) \leq 0$ for any t. If A, B in $\mathrm{SP}(A, B, \ldots)$ are max-parallel, then $c(t) \geq 0$ for any r, t; consequently, $C(t) \geq 0$ for any t.*

Proof. Immediately follows from Lemma 2.6 and Theorem 2.5. $\square$

2.3. Sequential Processes
2.3.1. Simple Serial Architectures of Size 2

Simple serial architectures of size 2 corresponds to the SP composition $T = A + B$, with $(A, B) \leftarrow\!\!\!\!\!\!\!\!\raise1pt\hbox{$\scriptscriptstyle\nearrow$}\; (\alpha, \beta)$. Recall the definitions of the two cumulative interaction contrasts: (27)–(30).

Theorem 2.8. *If $T = A + B$, then $c(0, t) \geq 0$ and $c(t, \infty) \leq 0$ for any r, t; moreover,*

$$\lim_{t \to \infty} c(0, t) = \lim_{t \to 0} c(t, \infty) = 0,$$

for any r, t. Consequently, $C(0, t) \geq 0$, $C(t, \infty) \leq 0$ for any t, and

$$\lim_{t \to \infty} C(0, t) = \lim_{t \to 0} C(t, \infty) = 0$$

Proof. Follows immediately from Lemma 2.4, since

$$-t_{11} + t_{12} + t_{21} - t_{22} = -\left(a_1 + b_1\right) + \left(a_1 + b_2\right)$$
$$+ \left(a_2 + b_1\right) - \left(a_2 + b_2\right) = 0.$$

$\square$

2.3.2. Two Sequential Processes in an Arbitrary SP Network

Consider now a composition $\mathrm{SP}(A, B, \ldots)$ with $(A, B, \ldots) \leftarrow\!\!\!\!\!\!\!\!\raise1pt\hbox{$\scriptscriptstyle\nearrow$}\; (\alpha, \beta, \emptyset)$.

Theorem 2.9. *If A and B are sequential in an $\mathrm{SP}(A, B, \ldots)$ composition, then one or both of the following statements hold:*
(i) $c(0, t) \geq 0$ for any r, t, and $C(0, t) \geq 0$ for any t,
(ii) $c(t, \infty) \leq 0$ for any r, t, and $C(t, \infty) \leq 0$ for any t.

Proof. In accordance with Definitions 1.2 and 1.4, $\mathrm{SP}(A, B, \ldots)$ with sequential A, B can be presented as either

$$\left(\mathrm{SP}^1(A, \ldots) + \mathrm{SP}^2(B, \ldots)\right) \wedge X + Y \tag{32}$$

or

$$\left(\mathrm{SP}^1(A, \ldots) + \mathrm{SP}^2(B, \ldots)\right) \vee X + Y \tag{33}$$

(note that any Z in $\mathrm{SP}^1(A, \ldots) + \mathrm{SP}^2(B, \ldots) + Z$ can be absorbed by either of the first two summands). For both cases, by the monotonicity of SP compositions, for any $R = r$, $\mathrm{SP}^1(a_1, \ldots) \leq \mathrm{SP}^1(a_2, \ldots)$, $\mathrm{SP}^2(b_1, \ldots) \leq \mathrm{SP}^2(b_2, \ldots)$, and we can always assume $\mathrm{SP}^1(a_1, \ldots) \leq \mathrm{SP}^1(b_1, \ldots)$. Denoting the durations of $\mathrm{SP}^1(a_i, \ldots) + \mathrm{SP}^2(b_j, \ldots)$ by t_{ij}', we have therefore, by Theorem 2.8, $-t_{11}' + t_{12}' + t_{21}' - t_{22}' = 0$. Denoting the durations of X and Y by t' and t'', respectively, in the case (32) we have

$$t_{ij} = t_{ij}' \wedge t' + t''.$$

By Lemma 2.4, all we have to show is that $-t_{11}+t_{12}+t_{21}-t_{22} \geq 0$. It is easy to see that t'' does not affect this linear combination, and its value is (assuming $t'_{12} \leq t'_{21}$, without loss of generality)

$$
\begin{cases}
0 & \text{if } t' < t'_{11} \\
-t'_{11} + t' & \text{if } t'_{11} \leq t' < t'_{12} \\
-t'_{11} + t'_{12} & \text{if } t'_{12} \leq t' < t'_{21} \\
-t'_{11} + t'_{12} + t'_{21} - t' & \text{if } t'_{21} \leq t' < t'_{22} \\
-t'_{11} + t'_{12} + t'_{21} - t'_{22} & \text{if } t' \geq t'_{22}.
\end{cases}
$$

The nonnegativity of the first three expressions is obvious, the fifth one is zero, and the forth expression is larger than the fifth because $t' < t'_{22}$.

The proof for the case (33) is analogous. □

If the SP composition with sequential A, B is homogeneous (Definition 1.5), the statement of theorem can be made more specific.

Theorem 2.10. *If A and B are sequential in an $\mathrm{SP}_\wedge(A, B, \ldots)$ composition, then $c(0, t) \geq 0$ for any r, t, and $C(0, t) \geq 0$ for any t; if the composition is $\mathrm{SP}_\vee(A, B, \ldots)$, then $c(t, \infty) \leq 0$ for any r, t, and $C(t, \infty) \leq 0$ for any t.*

3. Multiple Processes

We now turn to networks containing $n \geq 2$ processes with durations $(X^1, \ldots, X^n)$, selectively influenced by factors $(\lambda^1, \ldots, \lambda^n)$. In other words, we deal with compositions $\mathrm{SP}(X^1, \ldots, X^n, \ldots)$ such that $(X^1, \ldots, X^n, \ldots) \leftarrowtail (\lambda^1, \ldots, \lambda^n, \emptyset)$, where each λ^k is binary, with values 1,2. There are 2^n allowable treatments and 2^n corresponding overall durations, $T_{11\ldots1}, T_{11\ldots2}, \ldots, T_{22\ldots2}$. According to Definition 1.6 of selective influences, each process duration here is a function of some random variable R and of the corresponding factor, $X^k = x^k(R, \lambda^k)$. For any fixed value $R = r$, these durations are fixed numbers for any given treatment, and so is the overall, observed value of the SP composition. We denote them

$$x^k(r, \lambda^k = 1) = x^k_{1r}, \ x^k(r, \lambda^k = 2) = x^k_{2r}, \quad (34)$$

and

$$T(x^1_{i_1 r}, x^2_{i_2 r}, \ldots, x^n_{i_n r}, \ldots), \ldots = t_{i_1 i_2 \ldots i_n r}, \quad (35)$$

where $i_1, i_2, \ldots, i_n \in \{1, 2\}$. The distribution function for $t_{i_1 i_2 \ldots i_n r}$ is a shifted Heaviside function

$$h_{i_1 i_2 \ldots i_n r}(t) = \begin{cases} 0, & \text{if } t < t_{i_1 i_2 \ldots i_n r} \\ 1, & \text{if } t \geq t_{i_1 i_2 \ldots i_n r} \end{cases}. \quad (36)$$

Denoting by $H_{i_1 i_2 \ldots i_n}(t)$ the distribution function of $T_{i_1 i_2 \ldots i_n}$, we have

$$H_{i_1 i_2 \ldots i_n}(t) = \int_{\mathcal{R}} h_{i_1 i_2 \ldots i_n r}(t) \, d\mu_r. \quad (37)$$

Conditioned on $R = r$, the *n-th order interaction contrast* is defined in terms of mixed finite differences as

$$c^{(n)}_r(t) = \Delta_{i_1} \Delta_{i_2} \ldots \Delta_{i_n} h_{i_1 i_2 \ldots i_n r}(t), \quad (38)$$

which, with some algebra can be shown to be equal to

$$c^{(n)}_r(t) = \sum_{i_1, i_2, \ldots, i_n} (-1)^{n + \sum_{k=1}^n i_k} h_{i_1 \ldots i_n r}(t). \quad (39)$$

Thus,

$$
\begin{aligned}
c^{(1)}_r(t) &= \Delta_{i_1} h_{i_1 r}(t) = h_{1r}(t) - h_{2r}(t) \\
&= \sum_{i_1} (-1)^{1 + i_1} h_{i_1 r}(t),
\end{aligned} \quad (40)
$$

$$
\begin{aligned}
c^{(2)}_r(t) &= \Delta_{i_1} \Delta_{i_2} h_{i_1 i_2 r}(t) = \left[h_{11r}(t) - h_{12r}(t) \right] \\
&\quad - \left[h_{21r}(t) - h_{22r}(t) \right] \\
&= h_{11r}(t) - h_{12r}(t) - h_{21r}(t) + h_{22r}(t) \\
&= \sum_{i_1, i_2} (-1)^{2 + i_1 + i_2} h_{i_1 i_2 r}(t),
\end{aligned} \quad (41)
$$

$$
\begin{aligned}
c^{(3)}_r(t) &= \Delta_{i_1} \Delta_{i_2} \Delta_{i_3} h_{i_1 i_2 i_3 r}(t) \\
&= \left\{ \left[h_{111r}(t) - h_{112r}(t) \right] - \left[h_{121r}(t) - h_{122r}(t) \right] \right\} \\
&\quad - \left\{ \left[h_{211r}(t) - h_{212r}(t) \right] - \left[h_{221r}(t) - h_{222r}(t) \right] \right\} \\
&= h_{111r}(t) - h_{112r}(t) - h_{121r}(t) - h_{211r}(t) \\
&\quad + h_{122r}(t) + h_{212r}(t) + h_{221r}(t) - h_{222r}(t) \\
&= \sum_{i_1, i_2, i_3} (-1)^{3 + i_1 + i_2 + i_3} h_{i_1 i_2 i_3 r}(t),
\end{aligned} \quad (42)
$$

etc. The observable distribution function interaction contrast of order n is defined as

$$C^{(n)}(t) = \int_{\mathcal{R}} c^{(n)}_r(t) \, d\mu_r. \quad (43)$$

By straightforward calculus this can be written in extenso as

$$C^{(n)}(t) = \sum_{i_1, i_2, \ldots, i_n} (-1)^{n + \sum_{k=1}^n i_k} H_{i_1 \ldots i_n}(t), \quad (44)$$

or, in terms of finite differences,

$$C^{(n)}(t) = \Delta_{i_1} \Delta_{i_2} \ldots \Delta_{i_n} H_{i_1 i_2 \ldots i_n}(t). \quad (45)$$

This is essentially the high-order interaction contrast used by Yang et al. (2013), the only difference being that they use survivor functions $1 - H(t)$ rather than the distribution functions $H(t)$. We see that $c_r(t)$ and $C(t)$ in the preceding analysis correspond to $c^{(2)}_r(t)$ and $C^{(2)}(t)$, respectively.

We also introduce *n-th order cumulative contrasts*. Conditioned on $R = r$, we define

$$c^{[1]}_r(0, t) = c^{[1]}_r(t, \infty) = h_{1r}(t) - h_{2r}(t), \quad (46)$$

$$c_r^{[2]}(0,t) = \int_0^t c_r^{(2)}(t_1)\, dt_1, \quad c_r^{[2]}(t,\infty) = \int_t^\infty c_r^{(2)}(t_1)\, dt_1, \tag{47}$$

$$c_r^{[3]}(0,t) = \int_0^t \int_0^{t_1} c_r^{(3)}(t_2)\, dt_2 dt_1,$$

$$c_r^{[3]}(t,\infty) = \int_t^\infty \int_{t_1}^\infty c_r^{(3)}(t_2)\, dt_2 dt_1, \tag{48}$$

etc. Generalizing,

$$c_r^{[n]}(0,t) = \int_0^t \left(\int_0^{t_1} \cdots \int_0^{t_{n-2}} c_r^{(n)}(t_{n-1})\, dt_{n-1} \ldots dt_2 \right) dt_1, \tag{49}$$

$$c_r^{[n]}(t,\infty) = \int_t^\infty \left(\int_{t_1}^\infty \cdots \int_{t_{n-2}}^\infty c_r^{(n)}(t_{n-1})\, dt_{n-1} \ldots dt_2 \right) dt_1. \tag{50}$$

The corresponding unconditional cumulative contrasts of the n-th order are, as always, defined by integration of the conditional ones:

$$C^{[n]}(0,t) = \int_{\mathcal{R}} c_r^{[n]}(0,t)\, d\mu_r$$
$$= \int_0^t \left(\int_0^{t_1} \cdots \int_0^{t_{n-2}} C^{(n)}(t_{n-1})\, dt_{n-1} \ldots dt_2 \right) dt_1, \tag{51}$$

$$C^{[n]}(t,\infty) = \int_{\mathcal{R}} c_r^{[n]}(t,\infty)\, d\mu_r$$
$$= \int_t^\infty \left(\int_{t_1}^\infty \cdots \int_{t_{n-2}}^\infty C^{(n)}(t_{n-1})\, dt_{n-1} \ldots dt_2 \right) dt_1. \tag{52}$$

In the proofs below we will make use of the recursive representation of the conditional cumulative contrasts $c_r^{[n]}$. It is verified by straightforward calculus. Denoting

$$c_{i_w r}^{(n-1)}(t) =$$
$$\sum_{i_1,\ldots,i_{w-1},i_{w+1},\ldots,i_n} (-1)^{n-1-i_w+\sum_{k=1}^n i_k}\, h_{i_1 \ldots i_{w-1} i_w i_{w+1} \ldots i_n r}(t), \tag{53}$$

where $w \in \{1, \ldots, n\}$ and i_w is fixed at 1 or 2, we have:

$$c_r^{[1]}(0,t) = c_r^{[1]}(t,\infty) = h_{1r}(t) - h_{2r}(t), \tag{54}$$

$$c_r^{[2]}(0,t) = \int_0^t c_r^{(2)}(\tau)\, d\tau$$
$$= \int_0^t \left(h_{11r}(\tau) - h_{12r}(\tau) - h_{21r}(\tau) + h_{22r}(\tau) \right) d\tau$$
$$= \int_0^t \left[c_{i_w=1,r}^{(1)}(\tau) - c_{i_w=2,r}^{(1)}(\tau) \right] d\tau \tag{55}$$
$$= \int_0^t c_{i_w=1,r}^{[1]}(0,\tau)\, d\tau - \int_0^t c_{i_w=2,r}^{[1]}(0,\tau)\, d\tau,$$

$$c_r^{[2]}(t,\infty) = \int_t^\infty c_r^{(2)}(\tau)\, d\tau$$
$$= \int_t^\infty \left(h_{11r}(\tau) - h_{12r}(\tau) - h_{21r}(\tau) + h_{22r}(\tau) \right) d\tau \tag{56}$$
$$= \int_t^\infty \left[c_{i_w=1,r}^{(1)}(\tau) - c_{i_w=2,r}^{(1)}(\tau) \right] d\tau$$
$$= \int_t^\infty c_{i_w=1,r}^{[1]}(\tau,\infty)\, d\tau - \int_t^\infty c_{i_w=2,r}^{[1]}(\tau,\infty)\, d\tau,$$

$$c_r^{[3]}(0,t) = \int_0^t \int_0^{t_1} c_r^{(3)}(t_2)\, dt_2 dt_1$$
$$= \int_0^t \int_0^{t_1} \left[c_{i_w=1,r}^{(2)}(t_2) - c_{i_w=2,r}^{(2)}(t_2) \right] dt_2 dt_1$$
$$= \int_0^t \left[\int_0^{t_1} c_{i_w=1,r}^{(2)}(t_2)\, dt_2 - \int_0^{t_1} c_{i_w=2,r}^{(2)}(t_2)\, dt_2 \right] dt_1 \tag{57}$$
$$= \int_0^t c_{i_w=1,r}^{[2]}(0,\tau)\, d\tau - \int_0^t c_{i_w=2,r}^{[2]}(0,\tau)\, d\tau,$$

$$c_r^{[3]}(t,\infty) = \int_t^\infty \int_{t_1}^\infty c_r^{(3)}(t_2)\, dt_2 dt_1$$
$$= \int_t^\infty \int_{t_1}^\infty \left[c_{i_w=1,r}^{(2)}(t_2) - c_{i_w=2,r}^{(2)}(t_2) \right] dt_2 dt_1$$
$$= \int_t^\infty \left[\int_{t_1}^\infty c_{i_w=1,r}^{(2)}(t_2)\, dt_2 - \int_{t_1}^\infty c_{i_w=2,r}^{(2)}(t_2)\, dt_2 \right] dt_1 \tag{58}$$
$$= \int_t^\infty c_{i_w=1,r}^{[2]}(\tau,\infty)\, d\tau - \int_t^\infty c_{i_w=2,r}^{[2]}(\tau,\infty)\, d\tau,$$

and generally, for $n > 1$,

$$c_r^{[n]}(0,t) = \int_0^t c_{i_w=1,r}^{[n-1]}(0,\tau)\, d\tau - \int_0^t c_{i_w=2,r}^{[n-1]}(0,\tau)\, d\tau, \tag{59}$$

$$c^{[n]}(t,\infty) = \int_t^\infty c_{i_w=1,r}^{[n-1]}(\tau,\infty)\, d\tau - \int_t^\infty c_{i_w=2,r}^{[n-1]}(\tau,\infty)\, d\tau. \tag{60}$$

Also we have, by substitution of variables under integral,

$$c_{i_w r}^{[n-1]}(0,t) = c_r^{[n-1]}\left(0, t - x_{i_w r}^w\right), \tag{61}$$

$$c_{i_w r}^{[n-1]}(t,\infty) = c^{[n-1]}\left(t - x_{i_w r}^w, \infty\right). \tag{62}$$

The Prolongation Assumption generalizing (11)–(12) is formulated as follows.

Prolongation Assumption. R and functions $x^1, \ldots, x^n$ in (34) can be chosen so that $x_{1r}^k \leq x_{2r}^k$ for all $R = r$ and for all $k = 1, \ldots, n$. Without loss of generality, we can also assume $x_{1r}^1 \leq x_{1r}^2 \leq \ldots \leq x_{1r}^n$ (if not, rearrange $x_{1r}^1, \ldots, x_{1r}^n$).

Notation Convention. As we did before for $n = 2$, when r is fixed throughout a discussion, we omit this argument and write $x_{i_1}^1, \ldots, x_{i_n}^n$, $t_{i_1 i_2 \ldots i_n}$, $h_{i_1 i_2 \ldots i_n}(t)$, $c^{(n)}(t)$ in place of $x_{i_1 r}^1, \ldots, x_{i_n r}^n$, $t_{i_1 i_2 \ldots i_n r}$, $h_{i_1 i_2 \ldots i_n r}(t)$, $c_r^{(n)}(t)$.

3.1. Parallel Processes
3.1.1. Simple Parallel Architectures of Size *n*

Theorem 3.1. *If $T = X^1 \wedge \ldots \wedge X^n$, then for any r, t, $c^{(n)}(t) \leq 0$ if n is even and $c^{(n)}(t) \geq 0$ if n is odd. Consequently, for any t, $C^{(n)}(t) \leq 0$ if n is even and $C^{(n)}(t) \geq 0$ if n is odd.*

Proof. By induction on n, the case $n = 1$ being true by the Prolongation Assumption:

$$c^{(1)}(t) = h_1(t) - h_2(t) \geq 0.$$

Let the statement of the theorem be true for $c^{(n-1)}(t)$, with $n - 1 \geq 1$. By the Prolongation Assumption,

$$t_{1 i_2 \ldots i_n} = x_1^1 \wedge x_{i_2}^2 \wedge \ldots \wedge x_{i_n}^n = x_1^1,$$

for any $i_2 \ldots i_n$, whence

$$h_{1 i_2 \ldots i_n}(t) = \begin{cases} 0, & \text{if } t < x_1^1 \\ 1, & \text{if } t \geq x_1^1 \end{cases}.$$

Therefore $c_{i_1=1}^{(n-1)}(t) = 0$, and, applying the induction hypothesis to $c_{i_1=2}^{(n-1)}(t)$,

$$c^{(n)}(t) = c_{i_1=1}^{(n-1)}(t) - c_{i_1=2}^{(n-1)}(t) = -c_{i_1=2}^{(n-1)}(t)$$
$$= \begin{cases} \leq 0, & \text{if } n \text{ is even} \\ \geq 0, & \text{if } n \text{ is odd} \end{cases}.$$

That $C^{(n)}(t) \leq 0$ if n is even and $C^{(n)}(t) \geq 0$ if n is odd follows by the standard argument. $\square$

Theorem 3.2. *If $T = X^1 \vee \ldots \vee X^n$, then for any r, t, $c^{(n)}(t) \geq 0$. Consequently, for any t, $C^{(n)}(t) \geq 0$.*

Proof. By induction on n, the case $n = 1$ being true by the Prolongation Assumption:

$$c^{(1)}(t) = h_1(t) - h_2(t) \geq 0.$$

Let the theorem be true for $c^{(n-1)}(t)$, where $n - 1 \geq 1$. Let

$$x_2^1 \vee x_2^2 \vee \ldots \vee x_2^n = x_2^m,$$

where $1 \leq m \leq n$. We have then

$$t_{i_1 i_2 \ldots i_{m-1} 2 i_{m+1} \ldots i_n} = x_2^m,$$

and

$$h_{i_1 \ldots i_{m-1} 2 i_{m+1} \ldots i_n}(t) = \begin{cases} 0, & \text{if } t < x_2^m \\ 1, & \text{if } t \geq x_2^m \end{cases},$$

for all $i_1 \ldots i_{m-1}, i_{m+1} \ldots i_n$. Then $c_{i_m=2}^{(n-1)}(t) = 0$, and

$$c^{(n)}(t) = c_{i_m=1}^{(n-1)}(t) - c_{i_m=2}^{(n-1)}(t) = c_{i_m=1}^{(n-1)}(t) \geq 0.$$

Consequently, $C^{(n)}(t) \geq 0$, for any t. $\square$

3.1.2. Multiple Parallel Processes in Arbitrary SP Networks

In a composition $\mathrm{SP}(X^1, \ldots, X^n, \ldots)$, the components $X^1, \ldots, X^n$ are considered parallel if any two of them are parallel. We assume selective influences $(X^1, \ldots, X^n, \ldots) \nleftrightarrow (\lambda^1, \ldots, \lambda^n, \emptyset)$. We do not consider the complex situation when some of the selectively influenced processes $X^1, \ldots, X^n$ are min-parallel and some are max-parallel. However, if they are all (pairwise) min-parallel or all max-parallel, we have essentially the same situation as with a simple parallel arrangement of n durations.

Lemma 3.3. *If $X^1, \ldots, X^n$ are all min-parallel or max-parallel in an SP composition, this composition can be presented as $T = A^1 \wedge \ldots \wedge A^n$ or $T = A^1 \vee \ldots \vee A^n$, respectively. In either case, $(A^1, \ldots, A^n) \nleftrightarrow (\lambda^1, \ldots, \lambda^n)$ and the Prolongation Assumption holds for any $R = r$.*

Proof. For the min-parallel case, by a minor modification of the proof of Lemma 2.6 we present the SP composition as

$$\underbrace{\mathrm{SP}^1(X^1, \ldots)}_{=A^1} \wedge \mathrm{SP}^2(X^2, \ldots, X^n, \ldots),$$

or

$$\left(\mathrm{SP}^1(X^1, \ldots) \wedge \mathrm{SP}^2(X^2, \ldots, X^n, \ldots) + X\right) \wedge Y$$
$$= \underbrace{\left(\mathrm{SP}^1(X^1, \ldots) + X\right)}_{=A^1} \wedge \left(\mathrm{SP}^2(X^2, \ldots, X^n, \ldots) + X\right) \wedge Y,$$

or else

$$\left(\mathrm{SP}^1(X^1, \ldots) \wedge \mathrm{SP}^2(X^2, \ldots, X^n, \ldots) \wedge X\right) + Y$$
$$= \underbrace{\left(\mathrm{SP}^1(X^1, \ldots) + Y\right)}_{=A^1} \wedge \left(\mathrm{SP}^2(X^2, \ldots, X^n, \ldots) \wedge X + Y\right).$$

Then we analogously decompose $\mathrm{SP}^2(X^2, \ldots, X^n, \ldots)$ achieving $A^1 \wedge A^2 \wedge \mathrm{SP}^3(X^3, \ldots, X^n, \ldots)$, and proceed in this fashion until we reach the required $A^1 \wedge \ldots \wedge A^n$. The pattern of selective influences is seen immediately, and the Prolongation Assumption follows by the monotonicity of the SP compositions. The proof for the max-parallel case is analogous. $\square$

Theorem 3.4. *If $X^1, \ldots, X^n$ are min-parallel in an SP composition, then for any r, t, $c^{(n)}(t) \leq 0$ if n is even and $c^{(n)}(t) \geq 0$ if n is odd. Consequently, for any t, $C^{(n)}(t) \leq 0$ if n is even and $C^{(n)}(t) \geq 0$ if n is odd. If $X^1, \ldots, X^n$ are max-parallel, then for any r, t, $c^{(n)}(t) \geq 0$, and for any t, $C^{(n)}(t) \geq 0$.*

Proof. Follows from Lemma 3.3 and Theorems 3.1 and 3.2. $\square$

3.2. Sequential Processes
3.2.1. Simple Serial Architectures of Size *n*

Theorem 3.5. *If $T = X^1 + \ldots + X^n$, then for any r, t, $c^{[n]}(0, t) \geq 0$, while $c^{[n]}(t, \infty) \leq 0$ if n is even and $c^{[n]}(t, \infty) \geq 0$ if n is odd; moreover, $c^{[n]}(0, \infty) = 0$ for any r. Consequently, for any t, $C^{[n]}(0, t) \geq 0$, while $C^{[n]}(t, \infty) \leq 0$ if n is even and $C^{[n]}(t, \infty) \geq 0$ if n is odd; moreover, $C^{[n]}(0, \infty) = 0$.*

Proof. By induction on n, the case $n = 1$ being true by the Prolongation Assumption:

$$c^{[1]}(0, t) = c^{[1]}(t, \infty) = h_1(t) - h_2(t) \geq 0,$$

and

$$\lim_{t \to \infty} c^{[1]}(0, t) = \lim_{t \to 0} c^{[1]}(t, \infty) = 0.$$

Let the statement of the theorem hold for all natural numbers up to and including $n - 1 \geq 1$. Using the recursive representations (59)–(60),

$$
\begin{aligned}
c^{[n]}(0, t) &= \int_0^t c_{i_w=1}^{[n-1]}(0, \tau)\, d\tau - \int_0^t c_{i_w=2}^{[n-1]}(0, \tau)\, d\tau \\
&= \int_0^{t-x_1^w} c^{[n-1]}(0, \tau)\, d\tau - \int_0^{t-x_2^w} c^{[n-1]}(0, \tau)\, d\tau \quad , \\
&= \int_{t-x_2^w}^{t-x_1^w} c^{[n-1]}(0, \tau)\, d\tau
\end{aligned}
\tag{63}
$$

which is ≥ 0 since $c^{[n-1]}(0, \tau) \geq 0$ and $t - x_2^w \leq t - x_1^w$. Analogously,

$$
\begin{aligned}
c^{[n]}(t, \infty) &= \int_t^\infty c_{i_w=1}^{[n-1]}(\tau, \infty)\, d\tau - \int_t^\infty c_{i_w=2}^{[n-1]}(\tau, \infty)\, d\tau \\
&= \int_{t-x_1^w}^\infty c^{[n-1]}(\tau, \infty)\, d\tau - \int_{t-x_2^w}^\infty c^{[n-1]}(\tau, \infty)\, d\tau \quad , \\
&= -\int_{t-x_2^w}^{t-x_1^w} c^{[n-1]}(\tau, \infty)\, d\tau
\end{aligned}
\tag{64}
$$

which is ≤ 0 if n is even and ≥ 0 if n is odd. Applying the mean value theorem to the results of (63) and (64), we get, for some $t - x_2^w < t', t'' < t - x_1^w$

$$\int_{t-x_2^w}^{t-x_1^w} c^{[n-1]}(0, \tau)\, d\tau = c^{[n-1]}(0, t')\left(-x_1^w + x_2^w\right),$$

$$\int_{t-x_2^w}^{t-x_1^w} c^{[n-1]}(\tau, \infty)\, d\tau = c^{[n-1]}(t'', \infty)\left(-x_1^w + x_2^w\right),$$

and, as $c^{[n-1]}(0, \infty) = 0$, both expressions tend to zero as, respectively, $t \to \infty$ (implying $t' \to \infty$) and $t \to 0$ (implying $t'' \to 0$). □

3.2.2. Multiple Sequential Processes in Arbitrary SP Networks

In a composition $\mathrm{SP}\left(X^1, \ldots, X^n, \ldots\right)$, the components $X^1, \ldots, X^n$ are considered sequential if any two of them are sequential. By analogy with Theorem 2.9 for two sequential processes and with Theorem 3.4 for parallel $X^1, \ldots, X^n$, one might expect that the result for the simple sequential arrangement $X^1 + \ldots + X^n$ will also extend to n sequential components of more complex compositions $\mathrm{SP}\left(X^1, \ldots, X^n, \ldots\right)$. However, this is not the case, as one can see from the following counterexample.

Consider the composition

$$\mathrm{SP}(X^1, X^2, X^3, Y) = \left(X^1 + X^2 + X^3\right) \wedge (Y = 2), \tag{65}$$

with $\left(X^1, X^2, X^3\right)$ selectively influenced by binary factors, so that

$$
\begin{aligned}
x_1^1 = x_1^2 = x_1^3 = 0, \\
x_2^1 = x_2^2 = x_2^3 = 1.
\end{aligned}
\tag{66}
$$

It follows that

$$
\begin{aligned}
t_{111} &= 0, \\
t_{112} &= t_{121} = t_{211} = 1, \\
t_{122} &= t_{212} = t_{221} = t_{222} = 2.
\end{aligned}
\tag{67}
$$

This is clearly a sequential arrangement of the three durations X^1, X^2, X^3, but one can easily check that $c^{[3]}(0, t)$ here is not nonnegative for all t. For instance, at t=3 we have, by direct computation, $c^{[3]}(0, t) = -1$. We conclude that there is no straightforward generalization of Theorems 3.5 to arbitrary SP compositions.

4. Conclusion

The work presented in this paper is summarized in the abstract. By proving and generalizing most of the known results on the interaction contrast of distribution functions, we have demonstrated a new way of dealing with SP mental architectures. It is based on conditioning all hypothetical components of a network on a fixed value of a common source of randomness R (the "hidden variable" of the contextuality analysis in quantum theory), which renders these components deterministic quantities, and then treating these deterministic quantities as random variables with shifted Heaviside distribution functions.

The potential advantage of this method can be seen in the fact that the shifted Heaviside functions have the simplest possible arithmetic among distribution functions: for every time moment it only involves 0's and 1's. As a result, the complexity of this arithmetic does not increase with nonlinearity of the relations involved. Thus, Dzhafarov and Schweickert (1995); Cortese and Dzhafarov (1996), and Dzhafarov and Cortese (1996) argued that composition rules for mental architectures need not be confined to +, ∧, ∨. They analyzed architectures involving other associative and commutative operations, such as multiplication. Due to mathematical complexity, however, this work was confined to networks consisting of two components that are either stochastically independent or monotone functions of each other. It remains to be seen whether the approach presented here, *mutatis mutandis*, will lead to significant generalizations in this line of work.

The limitations of the approach, however, are already apparent. Thus, we were not able to achieve any progress over known results in applying it to Wheatstone bridges (Schweickert and Giorgini, 1999; Dzhafarov et al., 2004). The possibility that the "architecture" (composition rule) itself changes as one changes experimental factors makes the perspective of a general theory based on our approach even more problematic (e.g., Townsend and Fific, 2004). It seems, however, that these problems are not specific for just our approach.

Acknowledgments

References

Bell, J. (1964). On the Einstein-Podolsky-Rosen paradox. *Physics* 1, 195–200.

Bell, J. (1966). On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.* 38, 447–453. doi: 10.1103/revmodphys.38.447

Cereceda, J. (2000). Quantum mechanical probabilities and general probabilistic constraints for einstein–podolsky–rosen–bohm experiments. *Found. Phys. Lett.* 13, 427–442. doi: 10.1023/A:1007828731477

Cortese, J. M., and Dzhafarov, E. N. (1996). Empirical recovery of response time decomposition rules II: discriminability of serial and parallel architectures. *J. Math. Psychol.* 40, 203–218. doi: 10.1006/jmps.1996.0021

Dzhafarov, E. N. (2003). Selective influence through conditional independence. *Psychometrika* 68, 7–26. doi: 10.1007/bf02296650

Dzhafarov, E. N., and Cortese, J. M. (1996). Empirical recovery of response time decomposition rules I: sample-level Decomposition tests. *J. Math. Psychol.* 40, 185–202. doi: 10.1006/jmps.1996.0020

Dzhafarov, E. N., and Gluhovsky, I. (2006). Notes on selective influence, probabilistic causality, and probabilistic dimensionality. *J. Math. Psychol.* 50, 390–401. doi: 10.1016/j.jmp.2006.03.003

Dzhafarov, E. N., and Kujala, J. V. (2010). The joint distribution criterion and the distance Tests for selective probabilistic causality. *Front. Psychol.* 1:151. doi: 10.3389/fpsyg.2010.00151

Dzhafarov, E. N., and Kujala, J. V. (2012a). Selectivity in probabilistic causality: where psychology runs into quantum physics. *J. Math. Psychol.* 56, 54–63. doi: 10.1016/j.jmp.2011.12.003

Dzhafarov, E. N., and Kujala, J. V. (2012b). Quantum entanglement and the issue of selective influences in psychology: an overview. *Lect. Notes Comput. Sci.* 7620, 184–195. doi: 10.1007/978-3-642-35659-9/17

Dzhafarov, E. N., and Kujala, J. V. (2013). Order-distance and other metric-like functions on jointly distributed random variables. *Proc. Am. Math. Soc.* 141, 3291–3301. doi: 10.1090/s0002-9939-2013-11575-3

Dzhafarov, E. N., and Kujala, J. V. (2014a). On selective influences, marginal selectivity, and Bell/CHSH inequalities. *Top. Cogn. Sci.* 6, 121–128. doi: 10.1111/tops.12060

Dzhafarov, E. N., and Kujala, J. V. (2014b). Embedding quantum into classical: contextualization vs conditionalization. *PLoS ONE* 9:e92818. doi: 10.1371/journal.pone.0092818

Dzhafarov, E. N., and Kujala, J. V. (2014c). A qualified kolmogorovian account of probabilistic contextuality. *Lect. Notes Comput. Sci.* 8369, 201–212. doi: 10.1007/978-3-642-54943-4/18

Dzhafarov, E. N., and Kujala, J. V. (2014d). Contextuality is about identity of random variables. *Phys. Scr.* T163:014009. doi: 10.1088/0031-8949/2014/T163/014009

Dzhafarov, E. N., and Kujala, J. V. (in press). "Probability, random variables, and selectivity," in *The New Handbook of Mathematical Psychology,* eds W. Batchelder, H. Colonius, E. N. Dzhafarov, and J. Myung (Cambridge, UK: Cambridge University Press).

Dzhafarov, E. N., Kujala, J. V., and Larsson, J.-Å. (in press). Contextuality in three types of quantum-mechanical systems. *Found. Phys.* doi: 10.1007/s10701-015-9882-9

Dzhafarov, E. N., and Schweickert, R. (1995). Decompositions of response times: an almost general theory. *J. Math. Psychol.* 39, 285–314. doi: 10.1006/jmps.1995.1029

Dzhafarov, E. N., Schweickert, R., and Sung, K. (2004). Mental architectures with selectively influenced but stochastically interdependent components. *J. Math. Psychol.* 48, 51–64. doi: 10.1016/j.jmp.2003.12.003

Khrennikov, AY. (2009). *Contextual Approach to Quantum Formalism.* New York, NY: Springer.

Kochen. S., and Specker, E. P. (1967). The problem of hidden variables in quantum mechanics. *J. Math. Mech.* 17, 59–87. doi: 10.1512/iumj.1968.17.17004

Kujala, J. V., and Dzhafarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *J. Math. Psychol.* 52, 128–144. doi: 10.1016/j.jmp.2008.01.008

Landau, L. J. (1988). Empirical two-point correlation functions. *Found. Phys.* 18, 449–460. doi: 10.1007/BF00732549

Masanes, Ll., Acin, A., and Gisin, N. (2006). General properties of nonsignaling theories. *Phys. Rev. A* 73:012112. doi: 10.1103/PhysRevA.73.012112

Popescu, S. and Rohrlich, D. (1994). Quantum nonlocality as an axiom. *Found. Phys.* 24, 379–385. doi: 10.1007/BF02058098

Roberts, S., and Sternberg, S. (1993). "The meaning of additive reaction-time effects: tests of three alternatives," in *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience*, eds D. E. Meyer and S. Kornblum (Cambridge, MA: MIT Press), 611–654.

Schweickert, R., Fisher, D.L., and Sung, K. (2012). *Discovering Cognitive Architecture by Selectively Influencing Mental Processes.* New Jersey, NJ: World Scientific.

Schweickert, R., and Giorgini, M. (1999). Response time distributions: some simple effects of factors selectively influencing mental processes. *Psychon. Bull. Rev.* 6, 269–288. doi: 10.3758/BF03212330

Schweickert, R., Giorgini, M., and Dzhafarov, E. N. (2000). Selective influence and response time cumulative distribution functions in serial-parallel task networks. *J. Math. Psychol.* 44, 504–535. doi: 10.1006/jmps.1999.1268

Schweickert, R., and Townsend, J. T. (1989). A trichotomy: interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. *J. Math. Psychol.* 33, 328–347. doi: 10.1016/0022-2496(89)90013-8

Sternberg, S. (1969). The discovery of processing stages: Extensions of donders' method. *Acta. Psychologica.* 30, 276–315. doi:10.1016/0001-6918(69)90055-9

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *J. Math. Psychol.* 28, 363–400. doi: 10.1016/0022-2496(84)90007-5

Townsend, J. T. (1990a). The truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychol. Bull.* 108, 551–567. doi: 10.1037/0033-2909.108.3.551

Townsend, J. T. (1990b). Serial vs. parallel processing: sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychol. Sci.* 1, 46–54. doi: 10.1111/j.1467-9280.1990.tb00067.x

Townsend, J. T., and Fific, M. (2004). Parallel & serial processing and individual differences in high-speed scanning in human memory. *Percept. Psychophys.* 66, 953–962. doi: 10.3758/BF03194987

Townsend, J. T., and Nozawa, G. (1995). Spatio-temporal properties of elementary perception: an investigation of parallel, serial, and coactive theories. *J. Math. Psychol.* 39, 321–359. doi: 10.1006/jmps.1995.1033

Townsend, J. T., and Schweickert, R. (1989). Toward the trichotomy method of reaction times: laying the foundation of stochastic mental networks. *J. Math. Psychol.* 33, 309–327. doi: 10.1016/0022-2496(89)90012-6

Yang, H., Fific, M., and Townsend, J. T. (2013). Survivor interaction contrast wiggle predictions of parallel and serial models for an arbitrary number of processes. *J. Math. Psychol.* 58, 21–32. doi: 10.1016/j.jmp.2013.12.001

Corrigendum: Noncontextuality with marginal selectivity in reconstructing mental architectures

Ru Zhang and Ehtibar N. Dzhafarov *

Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

A corrigendum on

Noncontextuality with marginal selectivity in reconstructing mental architectures
by Zhang, R., and Dzhafarov, E. N. (2015). Front. Psychol. 6:735. doi: 10.3389/fpsyg.2015.00735

This corrigendum note points out and corrects two mistakes found in the paper cited in the title. These mistakes do not affect the correctness of the statements proved and expressions derived.

1. In Zhang and Dzhafarov (2015), Lemma 2.6 (p. 7) and Lemma 3.3 (p. 10) are formulated for series-parallel (SP) architectures in which the minimum ($\wedge$) and maximum ($\vee$) operations may be intermixed. The proofs are shown for the min-parallel arrangement of the selectively influenced processes, with the correct statement that the max-parallel arrangement is dealt with analogously. However, by an oversight, the proof for the min-parallel arrangement is shown only for homogeneous $\mathsf{SP}_\wedge$ architectures, those that cannot contain $\vee$ operations. The statements of the lemmas are correct despite this oversight, because the proofs remain valid if the rightmost $\wedge Y$ and $\wedge X$ in all expressions of the form

$$\left(\mathsf{SP}^1(\dots) \wedge \mathsf{SP}^2(\dots) + X\right) \underset{\uparrow}{\wedge} Y \text{ and } \left(\mathsf{SP}^1(\dots) \wedge \mathsf{SP}^2(\dots) \underset{\uparrow}{\wedge} X\right) + Y$$

are replaced with $\vee Y$ and $\vee X$, respectively.

2. Equations (61) and (62) on p. 9 should be disregarded: one of them, (62), contains typos, and both are shown in the wrong place. These transformations are only valid in the context of Theorem 3.5, for sequential architectures, and this is the only place where they are used, in Equations (63) and (64) on p. 11.

(A typo: in Equation (60) on p. 9, $c^{[n]}(t, \infty)$ should be $c_r^{[n]}(t, \infty)$.)

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

REFERENCES

Zhang, R., and Dzhafarov, E. N. (2015). Noncontextuality with marginal selectivity in reconstructing mental architectures. *Front. Psychol.* 6:735. doi: 10.3389/fpsyg.2015.00735

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS

Articles are free to read, for greatest visibility

COLLABORATIVE PEER-REVIEW

Designed to be rigorous – yet also collaborative, fair and constructive

FAST PUBLICATION

Average 85 days from submission to publication (across all journals)

COPYRIGHT TO AUTHORS

No limit to article distribution and re-use

TRANSPARENT

Editors and reviewers acknowledged by name on published articles

SUPPORT

By our Swiss-based editorial team

IMPACT METRICS

Advanced metrics track your article's impact

GLOBAL SPREAD

5'100'000+ monthly article views and downloads

LOOP RESEARCH NETWORK

Our network increases readership for your article