QUBITS AND SPACETIME

EDITED BY: Caslav Brukner, Jakub Mielczarek and Daniele Oriti
PUBLISHED IN: Frontiers in Physics and Frontiers in Astronomy and Space Sciences

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-432-9
DOI 10.3389/978-2-88974-432-9

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: frontiersin.org/about/contact

QUBITS AND SPACETIME

Topic Editors:
Caslav Brukner, University of Vienna, Austria
Jakub Mielczarek, Jagiellonian University, Poland
Daniele Oriti, Ludwig-Maximilians-University Munich, Germany

Citation: Brukner, C., Mielczarek, J., Oriti, D., eds. (2022). Qubits and Spacetime. Lausanne: Frontiers Media. doi: 10.3389/978-2-88974-432-9

Table of Contents

04 ***Holographic Space-Time and Quantum Information***
Tom Banks

14 ***On the Testability of the Equivalence Principle as a Gauge Principle Detecting the Gravitational t^3 Phase***
Chiara Marletto and Vlatko Vedral

18 ***On the Possibility of Experimental Detection of the Discreteness of Time***
Marios Christodoulou and Carlo Rovelli

23 ***Simulating Indefinite Causal Order With Rindler Observers***
Aleksandra Dimić, Marko Milivojević, Dragoljub Gočanin, Natália S. Móller and Časlav Brukner

30 ***Complexity and Newton's Laws***
Leonard Susskind

50 ***Replacing the Notion of Spacetime Distance by the Notion of Correlation***
Achim Kempf

58 ***Generalized Holographic Principle, Gauge Invariance and the Emergence of Gravity à la Wilczek***
Andrea Addazi, Pisin Chen, Filippo Fabrocini, Chris Fields, Enrico Greco, Matteo Lulli, Antonino Marcianò and Roman Pasechnik

69 ***Prelude to Simulations of Loop Quantum Gravity on Adiabatic Quantum Computers***
Jakub Mielczarek

79 ***Equivalence of Approaches to Relational Quantum Dynamics in Relativistic Settings***
Philipp A. Höhn, Alexander R. H. Smith and Maximilian P. E. Lock

108 ***Functorial Evolution of Quantum Fields***
Stefano Gogioso, Maria E. Stasinou and Bob Coecke

frontiers in Physics

MINI REVIEW
published: 21 April 2020
doi: 10.3389/fphy.2020.00111

Holographic Space-Time and Quantum Information

*Tom Banks**

The New High Energy Theory Center (NHETC), Department of Physics and Astronomy, Rutgers University, Piscataway, NJ, United States

The formalism of Holographic Space-time (HST) is a translation of the principles of Lorentzian geometry into the language of quantum information. Intervals along time-like trajectories, and their associated causal diamonds, completely characterize a Lorentzian geometry. The Bekenstein-Hawking-Gibbons-'t Hooft-Jacobson-Fischler-Susskind-Bousso Covariant Entropy Principle, equates the logarithm of the dimension of the Hilbert space associated with a diamond to one quarter of the area of the diamond's holographic screen, measured in Planck units. The most convincing argument for this principle is Jacobson's derivation of Einstein's equations as the hydrodynamic expression of this entropy law. In that context, the null energy condition (NEC) is seen to be the analog of the local law of entropy increase. The quantum version of Einstein's relativity principle is a set of constraints on the mutual quantum information shared by causal diamonds along different time-like trajectories. The implementation of this constraint for trajectories in relative motion is the greatest unsolved problem in HST. The other key feature of HST is its claim that, for non-negative cosmological constant or causal diamonds much smaller than the asymptotic radius of curvature for negative c.c., the degrees of freedom localized in the bulk of a diamond are constrained states of variables defined on the holographic screen. This principle gives a simple explanation of otherwise puzzling features of BH entropy formulae, and resolves the firewall problem for black holes in Minkowski space. It motivates a covariant version of the CKN [1] bound on the regime of validity of quantum field theory (QFT) and a detailed picture of the way in which QFT emerges as an approximation to the exact theory.

Keywords: spacetime and information, holographic spacetime, quantum gravity, covariant entropy principle, tensor network

OPEN ACCESS

Edited by:
Daniele Oriti, Ludwig-Maximilians-University Munich, Germany

Reviewed by:
Sayantan Choudhury, Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Germany
Pietro Antonio Grassi, University of Eastern Piedmont, Italy
Etera Livine, École Normale Supérieure de Lyon, France

***Correspondence:**
Tom Banks
banks@physics.rutgers.edu

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: *11 January 2020*
Accepted: *23 March 2020*
Published: *21 April 2020*

Citation:
Banks T (2020) Holographic Space-Time and Quantum Information. Front. Phys. 8:111. doi: 10.3389/fphy.2020.00111

1. INTRODUCTION

Every known human or computer language has the notion of time hard wired into every sentence. One of Einstein's great insights was that this notion is relative. Every information gathering system has its own *proper* time, and part of every physical theory must be a prescription for understanding the relations between the proper times of different systems. His second great insight, that "the speed of light is finite," can be thought of as the definition of what we mean by space and space-time. The region of space accessible to a system grows at a finite rate as a function of the proper time interval. The region of space-time accessible in a given time interval is called a causal diamond. One can view time evolution along a time-like trajectory/set of nested causal diamonds as a foliation of the space-time manifold into space-like leaves. The variety of such trajectories means that this can be

done in many different ways. This led Einstein to formulate his theory of gravitation as a theory of the Lorentzian geometry of the space-time manifold.

Lorentzian geometry can be recast as a theory of timelike trajectories and causal diamonds in a differentiable manifold. A timelike trajectory is a one parameter choice of negative norm tangent vectors, and defines a positive number, the proper time interval, between any two points along the trajectory. The causal diamond of a proper time interval is the set of all points that can be connected to both the past and future tips of the interval by timelike trajectories. The boundary of any finite area[1] diamond can be parametrized by two patches, with metrics

$$ds^2_{\pm} = du^{\pm} A^{\pm}_i(x, u^{\pm}) + g^{\pm}_{ij}(x, u^{\pm})\ dx^i dx^j. \tag{1}$$

The absolute maximum volume of the two Euclidean metrics $g^{\pm}_{ij}$, as a function of the null coordinates $u^{\pm}$ is called "the volume of the holographic screen of the diamond," which we will distort to "the area of the diamond" as a shorthand.

Although Jacobson did not use the language of causal diamonds, his seminal paper [2] showed that Einstein's gravitational equations follow as the hydrodynamics of a law that equates the "entropy" of a diamond to a linear function of its area. The null energy condition (NEC) then follows from increase of entropy and is seen to be a thermodynamic statement, which will have fluctuation corrections. Jacobson's derivation of Einstein's equations uses the frame of reference of a maximally accelerated trajectory to define energy. Such a system has infinite temperature, which is the strongest argument that the entropy in the covariant entropy bound [3] refers to the log of the dimension of the Hilbert space of the diamond [4].

The essence of Jacobson's argument, in the language of causal diamonds, is that the holographic screen [4] of the diamond is, by its definition, a maximum of the area on the boundary. Therefore, if we consider a pencil of null geodesics on the boundary of the diamond, approaching the holographic screen, then the Raychauduri equation can be linearized in the vicinity of the screen, and the increase of area can be written as

$$dA = R_{\mu\nu} k^{\mu} k^{\nu} d\lambda, \tag{2}$$

where λ is the affine parameter along the center of the pencil and k^{μ} is the null tangent vector. By appropriate choice of diamond, k^{μ} can be any null vector in space-time. The pair of future directed null trajectories following the boundary of the diamond past the holographic screen is the limit of a uniformly accelerated Unruh trajectory, with infinite Unruh temperature. Defining energy to be the limit of $k^{\mu} k^{\nu} T_{\mu\nu}$, where T is a covariantly conserved stress tensor, the equation $dE = TdS$, with $S = G_N A/4$, gives us exactly (in 4 dimensions)

$$k^{\mu} k^{\nu} (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - 8\pi G_N T_{\mu\nu}) = 0. \tag{3}$$

This is the content of Einstein's equation without the cosmological constant (c.c.). We've thus derived the gravitational field equations as the hydrodynamic equations of the area law, and simultaneously shown that the c.c. is not a hydrodynamic energy density.

The area law for entropy is the clue for understanding locality/causality in a quantum theory of space-time. Given a time-like trajectory, the causal diamonds of a nested series of proper time intervals partition the interior of the largest diamond into a sequence of quantum subsystems whose maximal entropy is non-decreasing as a function of the length of proper time. Causality is the statement that the smaller subsystems remain unentangled with the rest of the degrees of freedom during the relevant proper time intervals. This implies that time evolution is naturally viewed trajectory by trajectory[2] and that the natural time slices inside a diamond must remain inside the diamond. The Hamiltonian is perforce time dependent. This can be viewed either as the quantum requirement of gradual entanglement of subsystems or, macroscopically, as the geometric requirement that time slices remain within a diamond. As Milne [5] first appreciated, this kind of time slicing induces a redshifting of the "energies" of distant objects[3].

From a more philosophical point of view, what a formalism based on these ideas is saying, is that time is fundamental, but relative (trajectory dependent, many fingered), while space is an emergent concept describing a measure of the amount of quantum information required to describe a certain time interval. The quantum analog of Einstein's principle of relativity then becomes apparent. Consider a pair of causal diamonds along two different trajectories. There is a maximal causal diamond in their intersection. The Covariant Entropy Principle (CEP) assigns this diamond a Hilbert space of fixed dimension, which will always be smaller (geometry) than the dimensions of either intersecting diamond. Each parent diamond is a quantum system with time dependent Hamiltonian and, given a choice of initial pure state, will assign a sequence of density matrices to the subsystem describing the intersection. The Quantum Principle of Relativity (QPR) asserts that the two density matrices assigned by the parent diamonds have the same entanglement spectra. This constrains the choice of both the Hamiltonian and the initial state in each diamond. We'll outline below the utility of this principle for trajectories at relative rest. We have not yet found a model that implements the QPR for pairs of trajectories in relative motion.

While the CEP allows us to localize quantum information on the holographic screens of nested or intersecting diamonds, it does not give us a clear definition of a traditional localized excitation in the bulk of a given diamond. The clue to bulk localization comes from two formulae in black hole physics. The first is the entropy formula for Schwarzschild-de Sitter black holes. The metric is

$$ds^2 = -f(r)dt^2 + dr^2/f(r) + r^2 d\Omega^2, \tag{4}$$

[1] Here we're anticipating the definition of area that we are about to give. In the limit of infinite area, the past and future halves of the diamond's boundary do not have to be joined differentiably. We also, by abuse of language, use the term *area* for the $d-2$ volume of a space-like slice of boundary of a causal diamond in d dimensions.

[2] Jacobson's derivation of Einstein's equations from the first law of (local) thermodynamics uses the energy along a particular maximally accelerated trajectory and thus also points to a trajectory by trajectory view of time evolution.

[3] Milne was of course incorrect in assuming that the observed cosmological redshift could be attributed entirely to this kinematic effect.

where

$$rf(r) = -(r - R_+)(r - R_-)(r + R_+ + R_-), \quad (5)$$

$$R^2 = R_+^2 + R_-^2 + R_+R_-, \quad 2MR^2 = R_+R_-(R_+ + R_-). \quad (6)$$

R is the dS radius, and M is the parameter that becomes the black hole mass in the $R \to \infty$ limit. Everything is written in Planck units. This formula shows that the introduction of any object with a long range Schwarzschild field, at rest on a timelike geodesic of dS space, creates an entropy deficit. The CEP identifies this, when M is small compared to the maximal black hole mass in dS space, as the deficit expected in a thermal ensemble with temperature $T = (2\pi R)^{-1}$, the Gibbons-Hawking temperature [6]. In fact one can demonstrate a similar entropy deficit effect in Minkowski space [7], which suggests that the Minkowski vacuum is an infinite entropy ensemble.

The second hint that localized objects are constrained states of holographic variables comes from the ordinary formula for the entropy of a Minkowski black hole of mass M when an additional small mass m is dropped into it. Despite the fact that the small mass is a very low entropy object, the final equilibrium state is a state of much higher entropy

$$\Delta S = 2\pi R_S m. \quad (7)$$

This indicates that before equilibration, the combined system lived in a much larger Hilbert space than that of the original black hole, but that the initial state had $2\pi R_S m$ frozen degrees of freedom. Inverting this process (by unitarity), we have a derivation of the Hawking temperature for emission of the system of mass m.

If the Hamiltonian that equilibrates the system has a natural time scale R_S and is a "fast scrambler" [8, 9], then the infalling subsystem will remain isolated for a time of order $R_S \ln R_S$, and this is the basis for the resolution of the "firewall paradox" [10–12]. Again, the principle operating here is that a localized state in the causal diamond formed by the horizon of the black hole of mass $m+M$ is a constrained state of the Hilbert space of that black hole. Another important feature we learn from this discussion is that the constraints must have the property that they isolate the degrees of freedom of the small system, from that of the black hole. This immediately suggests that the degrees of freedom should form matrices, with a single trace Hamiltonian, and the constraints implying that off diagonal matrices connecting the "m-block" to the "M-block," vanish.

Constraints can "propagate through a nested sequence of causal diamonds," giving a holographic interpretation of particle trajectories. More properly we'll see that these should be thought of as jets of particles, including many soft gravitons whose number changes with time. Indeed, we'll see that "particle" is a perturbative concept and jets are the fundamental scattering states in models of quantum gravity in Minkowski space. The trajectory of a jet, a quantum system with many states for fixed momentum, is a much more robust semi-classical object than a particle trajectory in quantum field theory.

We'll see that the CEP, the QRP, the identification of particle jets as constraints, and the fast scrambling properties of black hole horizons give us a number of vital clues to the nature of a general theory of quantum gravity. For example, the QRP enables us to tie together jet interactions (some number of jets enter the past boundary of a diamond and a possibly different number exit its future boundary) in different causal diamonds, obtaining a manifestly local, Feynman diagram like, description of transition amplitudes. The same formalism can describe the production and decay of high entropy meta-stable excitations with all of the qualitative properties of black holes.

2. THE HOLOGRAPHIC VARIABLES OF QUANTUM GRAVITY

The CEP implies that a finite area diamond corresponds to a finite dimensional Hilbert space. The fact that the $U(D)$ GellMann matrices, which are closed under both commutation and anti-commutation, form a basis for all complex matrices shows us that this space is the fundamental representation of the super-algebra $SU(P|Q)$ for any integers P, Q such that $P + Q = D$. That is to say, fermionic variables are inevitable in any finite dimensional quantum system. This remark ignores the constraint of spatial locality. A discrete, spatially local system can be defined on the tensor product of finite dimensional Hilbert spaces sitting at the points of some graph, whose links define what we mean by nearest neighbor, next to nearest neighbor, etc. couplings. Fermionic operators on the full Hilbert space will be non-local functions of the bosonic site variables. In some cases [13–15] a local theory of mutually commuting site variables, with a Z_2 gauge invariance, can be rewritten as a local theory of fermions, but this is not always the case.

The fast scrambling property of black holes [8, 9] implies that the correct quantum theory cannot be local on the holographic screen of a diamond[4]. Instead we will suggest that the Hamiltonian should be invariant under a finite dimensional approximation to the group of area preserving maps on the sphere. The theory of fuzzy approximations to Euclidean geometries has a long history. Traditionally it is viewed as the replacement of the algebra of smooth, or continuous, functions on the manifold by a finite dimensional non-abelian matrix algebra. This can be developed in a systematic way for any manifold with a Kahler or symplectic structure. In Banks and Kehayias [16] we proposed a different approach, inspired by Connes' insight about the connection between the Dirac operator and Riemannian geometry. The Dirac operator on any spin manifold is an unbounded operator with spectrum symmetric around 0 and compact inverse on the space of spinor sections orthogonal to its discrete zero modes. Its eigenvalues are invariant under any symmetries of the manifold, and its zero mode spectrum encodes some of the topological properties. The space of spinor bilinears is the space of all differential forms on the manifold, so appropriate products of spinor bilinears are proportional to its volume form and a Hamiltonian given

[4] In AdS space, for black holes larger than the radius of curvature, scrambling *is* ballistic on length scales larger than the AdS radius. This is a consequence of the AdS/CFT correspondence.

by the integral over such products is invariant under area preserving maps.

We fuzzify the geometry by putting a symmetric eigenvalue cutoff $P > 0$ on the Dirac operator. For large P, the eigenvalue degeneracy goes like P^{d-2} where d is the space-time dimension, so if each eigensection is quantized in a finite dimensional Hilbert space of fixed dimension, then we get an area law for the maximal entropy. On the $d-2$ sphere, the counting of spinor spherical harmonics is exactly that of anti-symmetric $d-2$ tensors with indices ranging from 1 to $o(P)$ [16]. We can think of these as little area elements. The matrix $M_i^j \equiv \psi_{i,a^1\ldots a^{d-3}}\psi^{\dagger\, j,a_1\ldots a_{d-3}}$ can be viewed as a $d-3$ sphere "band" on the surface of a $d-2$ sphere and the trace of a product of these matrices is a "line bundle" construction of the $d-2$ sphere from a succession of such bands. In plainer language, it's the picture of the $d-2$ sphere as a succession of "thick" $d-3$ spheres along a polar coordinate

$$ds^2 = d\theta^2 + \sin^2\theta d\Omega_{d-3}^2. \tag{8}$$

The bilinears in spinors are differential forms of varying degrees and the trace is the integral over $d-2$ forms formed as wedge products of these elementary forms. Any action constructed from sums of such single traces is invariant, in the formal continuous limit, under area preserving maps. If we write

$$H_{in}(t) = \frac{1}{t}\mathrm{Tr}\, P(\frac{M_{t\times\, t}}{t^{d-3}}), \tag{9}$$

where t is the proper time in a diamond, and P is a polynomial of finite order whose coefficients are t independent in the large t limit, then the leading term in the energy scales like P^{d-3} as $t \to \infty$. The gaps between an infinite number of low lying states and the ground state are $o(1/t)$ in this limit. The CEP indicates that $P \sim t$ should be proportional to the radius of the sphere, in Planck units. The relation between the short wavelength cutoff on the sphere, and the proper time/area of the holographic screen, is a UV/IR correspondence, generalizing Maldacena's scale radius duality.

The full Hamiltonian of HST is more complicated than this in constrained subspaces, which make the matrices block diagonal. If we have a number of isolated blocks in a causal diamond of size t^5, then for each block of size n_b we add

$$H_{in}^{n_b}(t) = kn_b^{d-3} + \frac{1}{n_b}\mathrm{Tr}\, P(\frac{M_{n^b\times\, n^b}}{n_b^{d-3}}), \tag{10}$$

where k is a constant, which will be determined by the correct normalization of energies in the limits discussed below. The first term is the "asymptotically conserved energy of the jet represented by the block," while the second term represents fragmentation of the jet into subjets constituents. The block COULD also represent an isolated black hole in the diamond and then the second term represents interactions on the black hole horizon. Note that $H_{in}^{n_b}(t)$ is t independent. This is because it represents excitations localized near the trajectory, which have order 1 energies in "Milne" coordinates. There is also, of course, an $H_{out}(t)$ describing interactions of degrees of freedom outside the diamond. We'll see below that this is determined by the QRP.

The commutation relations for these variables that are invariant under $SO(d-1)$ are

$$[\psi_A, \psi^B]_+ = \delta_A^B. \tag{11}$$

Here A, B are $d-2$ dimensional antisymmetrized multi-indices and the right hand side is the appropriately antisymmetrized Kronecker symbol. These, and the Hamiltonian are invariant under the larger $U(t)$ group of unitaries, which can be interpreted as a fuzzy approximation to the group of area preserving maps. It has many $SO(d-1)$ subgroups under which the variables transform as a sum of spherical harmonics up to some maximal angular momentum (simply conjugate one $SO(d-1)$ subgroup by a general element of $U(t)$).

If the ψ variables have another index A, apart from their $SO(d-1)$ spinor label we can try to view them as fuzzy spinors on a higher dimensional manifold, of the form $\mathcal{K}\otimes\mathcal{M}^{1,d-1}$ because the spinor bundle on a product manifold is a tensor product of spinors on the lower dimensional manifolds. More research is needed to find restrictions on the commutation relations as a function of the A label, which approach geometrically sensible rules, where the anti-commutator of two spinor generators involves forms integrated over closed cycles on a manifold, in the limit that the number of A labels gets large.

Given the generators $\psi_{a^1\ldots a^{d-3}}$, we can define mutually commuting Pauli operators by multiplying each fermion by $(-1)^{N-n}$ where N is the total number operator and n the number operator of that particular species. The bilinear

$$\psi_a^A \psi_A^{\dagger\, b},$$

becomes

$$\sigma_{-\ a}^{A}\sigma_{3\ a}^{A}\sigma_{3\ A}^{b}\sigma_{+\ A}^{b}.$$

Here A is a $d-3$ component index, representing an interface between two bands on the $d-2$ sphere. So we can "bosonize" these fermions without introducing any more non-locality than was present in the original Hamiltonian. For models invariant under the fuzzy version of area preserving maps, fermionic variables are natural, invariant, and as local as a bosonic presentation of the same Hamiltonian.

An alternative view of the fermionic variables of HST comes from a proposal for generalized scattering theory for models of quantum gravity in Minkowski space [17–19]. Ordinary scattering theory for quantum field theories with a mass gap is based on the infinite set of asymptotically conserved LSZ currents

$$j_\mu^f = i(f_\pm\partial_\mu\phi - f_\pm\partial_\mu\phi). \tag{12}$$

Here ϕ is an interacting Hermitian Heisenberg field and $f_\pm$ normalizable positive or negative energy solutions of the Klein-Gordon equation with the physical particle mass. Matrix

[5] At this point the reader should be prepared to understand "causal diamond" as "tensor factor of the Hilbert space which interacts only with itself over the time interval $[-t, t]$."

elements of these currents in physical states are assumed to be conserved near the conformal boundary of Minkowski space, and this is true up to exponential corrections, to all orders in perturbation theory. The physical Fock space is the representation space of the algebra of these currents and the Scattering operator intertwines been the past and future bases. This formalism breaks down for massless particles.

However, all particles that are massless for an entire range of couplings are associated with conserved currents. For Goldstone bosons, where we can turn on a mass continuously, violating the current conservation law, the almost conserved current plays the role of the field $\partial_\mu \phi$ in massive scattering theory and it's plausible that the asymptotic Hilbert space is simply the representation space of the conserved current in the limit. Similarly gauge and gravitational fields all have asymptotically conserved currents associated with them [17–19]. The stress tensor in gravitational models plays a special role because the joint spectrum of its asymptotically conserved currents is the momentum null cone, the Fourier dual of the conformal boundary. All other conserved currents can be viewed as generalized functions on this cone. That is, they are "quantized fields" on the momentum null cone. We'll see that the reason for the scare quotes is that these operator valued generalized functions are not the conventional tempered distributions of axiomatic QFT. The null cone is a singular manifold and conventional Wightman fields would not be well-defined there.

More importantly, the behavior of black hole and cosmological quasi-normal modes indicates that the quantum systems living on horizons cannot have the approximate locality in angle that one would expect from even a lattice approximation to a conventional QFT, where the rigorous theorem of Lieb and Robinson proves that information transport over large distances is ballistic. Instead one expects all of the degrees of freedom to be coupled together, without regard to metrical distance. Correspondingly the fields are not expected to satisfy differential equations. The Hamiltonians we have written are not local, and are fast scramblers, because every fermionic variable is coupled to every other one by some term in the Hamiltonian.

The purpose of currents on the conformal boundary is to describe the flow of quantum numbers other than the momentum, at infinity. Helicity or spin must be one of those quantum numbers, so we expect operators $H_i^{\pm}(P)$ carrying helicity out of/into the future/past null boundary and $\tilde{H}_i^{\pm}(P)$ describing flows along the boundary. The two kinds of operators are related by space reflection, and only the tilde-free operators are needed to describe massless particles.

When we retreat from the conformal boundary to a finite area causal diamond, P must become a discrete label[6] and the CEP implies that the Hilbert space on which the generators $H_i^{\pm}(P)$ and $\tilde{H}_i^{\pm}(P)$ act must be finite dimensional, in which case we can always view the same space as generated by fermionic operators as above. If we want the formalism to obey the spin-statistics theorem, then those fermions must carry half integer helicity and must, in the conformal boundary limit, take the form $Q_\alpha^I(P)$, $\tilde{Q}_\alpha^I(P)$, where $\gamma^a P_a Q^J(P) = \gamma^a \tilde{P}_a \tilde{Q}(P) = 0$. Here $\tilde{P}$ is the space reflected null vector. Note that these kinematic arguments do not imply that the model must be supersymmetric. If all fermionic generators come in parity symmetric pairs, then the spin 3/2 particles that must accompany the graviton will be massive.

The algebra of the left or right handed spinor generators is completely determined [20] by Lorentz invariance, cluster decomposition, and the absence of tensor charges in an interacting theory of particles. It is

$$[Q_\alpha^I(P), Q_\beta^J(P')]_+ = \delta^{IJ}\delta(P \cdot P')\gamma^\mu_{\alpha\beta} M_\mu(P, P'). \tag{13}$$

M_μ is the smaller of the two parallel null vectors. Note that the $P = 0$ generators anticommute with all the others. There's a similar equation for the space-reflected generators. The anti-commutation relations between the two sets of generators are not universal, and encode information about the masses of stable particles corresponding to branes wrapped around non-trivial cycles of a compact manifold. As noted above, the detailed mathematics of the connection between finite dimensional super-algebras and the notion of smooth compact manifolds, has not yet been worked out.

3. TIME DEPENDENT HAMILTONIANS AND ERROR CORRECTING CODES

The basic principles of HST imply that causality is implemented by gradually entangling new degrees of freedom in a larger causal diamond with the subset describing a smaller diamond contained in the original one. We can ask where on the holographic screen of the larger diamond, the information about the smaller diamond is stored. As long as the dynamics is invariant under (fuzzy) area preserving maps, this question has no meaning. However, the constraints are a partial breaking of this symmetry. The variables are labeled by spinor harmonic quantum numbers on the sphere, but there are an infinite number of ways of doing this, corresponding to the embeddings of $SO(d-1)$ in the group of area preserving maps. The constraints are interpreted in a way that mirrors a metric geometry on the sphere.

For simplicity, let's work on four dimensions. In a large causal diamond with proper time T, we say that the physical state contains a localized jet on the past or future boundary if of order ET, with $E \ll T$, of the variables ψ_i^J vanish on that state. Given the single trace nature of the interactions, this means that interactions between the variables $\psi_{[ij]}$ and the rest vanish on this state. Here the small letters form an antisymmetric matrix with indices from 1 to E, which can be organized into fuzzily localized spinor sections around some point Ω. We can think of the constraints as the vanishing of variables in an annulus surrounding a spherical cap, whose opening angle is determined by E. More generally there will be multiple isolated subsets of degrees of freedom, which form $E_i \times E_i$ anti-symmetric matrices and are interpreted as belonging to spherical caps localized around different angles Ω_i. It can be shown [21] that $\sum E_i$ is an asymptotically conserved quantum number if the time dependent Hamiltonian has energy differences of order $1/T$. Asymptotically, for large T, there will be a unique choice of rotation subgroup

[6] We'll see later that it is an *emergent* label.

for which all of the jets form spherical caps localized around different points.

To see this, note that we can always choose the $E_i \times (E_i - 1)$ components of a single jet's degrees of freedom to be localized functions around some particular point on the sphere, with localization radius $\sim 1/E_i$, by choosing a particular linear combination of spherical harmonics. We can do the same for all the other jets, keeping them separated in angle, as long as $\sum E_i \ll T$, so that there are plenty of variables available to describe the empty angular regions with no jets. In the limit $T \gg \sum E_i$ with $E_i \to \infty$ we can accommodate an arbitrary number of localized jets. Note that the rotation subgroup we choose should also be used to organize the addition of degrees of freedom each time we increase the proper time by one Planck unit. This corresponds to adding exactly one angular momentum multiplet of the chosen subgroup to the "in" Hilbert space for every Planck size tick of the clock. It's also important to emphasize that the action of the rotations on the decoupled $(T - \sum E_i) \times (T - \sum E_i)$ block of the matrix is, in a sense, trivial since the interactions of these variables are invariant under area preserving maps that leave invariant the spherical caps at which the jets are located. The ratio of the areas of those caps to the area of the sphere goes to zero in the limit. Since the Hamiltonian of this large set of degrees of freedom goes to zero in the limit, they become topological degrees of freedom, sensitive only to the punctures on the sphere. This is the HST description of the infinite dimensional space of arbitrarily soft massless particles that are present in any quantum theory of gravity in Minkowski space.

Thus, both rotation and time translation are asymptotic symmetries, as expected in a theory of gravity. Note that the magnitude of the null momentum is an emergent quantity, proportional to $\sum E_i$. To get a Lorentz invariant scattering operator we must take all E_i to infinity at fixed ratio, keeping $\sum_i E_i \ll T$.

As a consequence, the error correcting code [22–34][7] generated by the expansion of $H_{in}(t)$ to include more degrees of freedom, contains information that allows us to localize information about the constrained variables at angles. Now however, consider the full evolution in the interval from $[-T, T]$. The initial state satisfies constraints corresponding to incoming jets with energies E_i. At some later negative time $-t$ with $t < T$ there are two possibilities. Either we inevitably reach a point where $\sum E_i \sim t$ or some of the constraints and decoupled degrees of freedom are not contained in the Hilbert space on which $H_{in}(t)$ acts. In the former case the fast scrambling nature of the Hamiltonian implies that the constraints will be erased by the time one gets to the end of the time interval $[-t, t]$. The entire Hilbert space will be in equilibrium and we have a causal diamond with energy proportional to $\sum E_i$ filled with an isotropic system on its boundary, in equilibrium with entropy $(\sum E_i)^2$. This system has all of the qualitative properties of a black hole. **Figure 1** shows cartoons of the two possibilities.

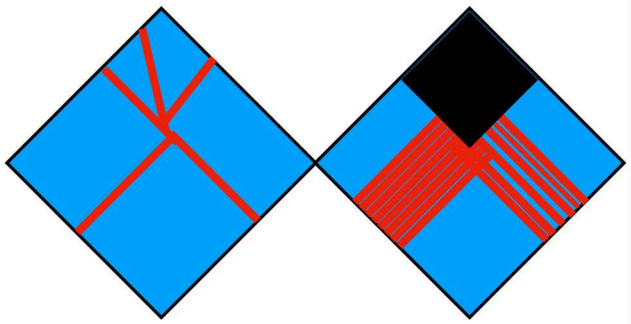

FIGURE 1 | The left figure shows a system with a number of constraints much smaller than the total number of degrees of freedom while the right one is what happens when the constrained subspace has entropy that is an order one fraction of the total. Red lines denote jet degrees of freedom, each of which is surrounded on the past/future boundary of the diamond, by frozen degrees of freedom,indicated by erasure of the boundary. The diamonds in these figures are finite, and the right hand picture does not include black hole evaporation.

There are two different kinds of amplitudes where no black hole production occurs. In the first, the total energy coming into the past boundary of the causal diamond $[-T, T]$ is so small that E^2 is not a large entropy, *and* all of those constraints propagate into smaller diamonds along the same trajectory. Then the future boundary of a diamond of proper time $t > E$ has a small number of constraints, which can be interpreted as jets of particles exiting that boundary. On scales $t = E$ the amplitude looks like a vertex in a Feynman diagram.

Another possibility is that the constraints proportional to the total incoming energy, which might be large, do not all propagate into small diamonds along the trajectory. Here is where the overlap constraints of HST demonstrate the emergence of the concept of *space* in the HST formalism. At time $t \ll T$, constraints that are not imposed on the Hilbert space of the $[-t, t]$ diamond, are imposed on its tensor complement in the $[-T, T]$ Hilbert space, which is acted upon by the Hamiltonian $H_{out}(t)$. The structure of $H_{out}(t)$ is determined by the HST compatibility conditions, the QPR. That is, given an assumed global structure of space-time, which is a dS space with $R \gg T \gg 1$ we can impose boundary conditions on causal diamonds with proper time $[-T, T]$ corresponding to $\sim ET$ constraints, with $E \ll T$, along time-like geodesics "at different spatial points in their common rest frame." These are identical quantum systems, with the same sequence of time dependent Hamiltonians.

Now consider, for a given initial state, the Hilbert spaces of these individual systems over time intervals $[-t + t_i, t_i + t]$. Let us first assume that the dynamics is such that in the large T limit the proportionality constant E in the number of constraints $ET + k$ is conserved. Call it the energy. This is true for every individual trajectory. When $E \sim t$ or greater, these cannot all be constraints on the "in" Hilbert spaces of the small

[7]The connection between quantum error correction and bulk (AdS scale) locality was pointed out in the second paper of this reference, but was anticipated by the tensor network construction of Swingle. The general idea of error correction is to entangle the desired quantum information, with widely distributed degrees of freedom of a much larger system, so that erasing the part entangled with a few q-bits does not degrade the information. The particular use of error correction in the AdS/CFT correspondence exploits/is limited by the locality of the boundary theory. It is not appropriate for discussing horizons whose dynamics is invariant under area preserving maps. HST claims to remedy this.

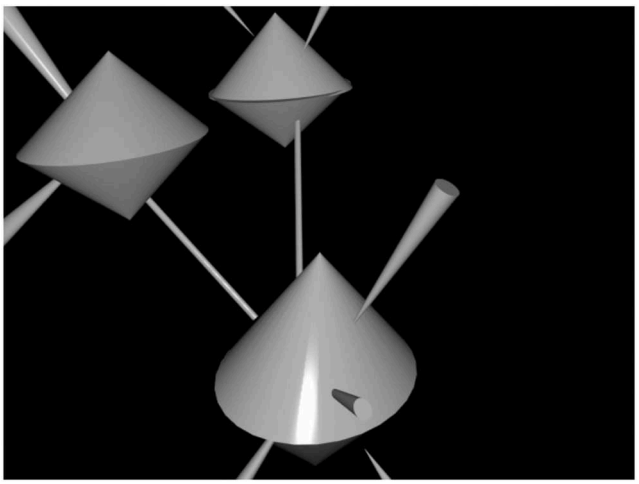

FIGURE 2 | Exchange diagram involving multiple jets.

causal diamonds. Therefore, the energy must be divided between the in and out spaces. The QPR implies that at any time, the "out" Hilbert space of a given diamond can be viewed as a tensor product of small Hilbert spaces corresponding to spatially separated diamonds. If we take all of the t_i equal, the translation of this statement into space-time language is that the energy E is the sum of energies E_i in each of the disjoint diamonds, with $E_i \ll t$ if we want to study a process in which no black holes are created. The QPR implies that the "out" dynamics of any one diamond generates the same entanglement spectrum for the density matrices in each of the external diamonds, that is generated by the "in" dynamics of each of those diamonds. Since we're studying geodesics in Minkowski space, the Hamiltonians are all assumed equal to each other, so the QPR is a constraint on initial states.

Now consider unequal t_i. The final state constraints on the earliest diamond, can become part of the initial constraints on later diamonds, so we can draw a space-time picture of the amplitude that resembles a time ordered Feynman diagram (**Figure 2**). Thus, these models reproduce the clustering structure of field theory amplitudes, which we usually derive from the postulates of QFT. However, the HST formalism can also describe black hole formation and evaporation in a manner consistent with unitarity and causality. So far, we have not found a model that gives Lorentz invariant scattering amplitudes.

We can close this circle by using QPR to finish the proof that the coefficient of T in the number of asymptotic constraints *is* a conserved quantity. If all of the asymptotic energy remains visible along a single time-like trajectory, conservation is a consequence of two facts. The Hamiltonian in a causal diamond of proper time t that is capable of removing constraints that prevent interaction between bulk and boundary DOF has eigenvalue differences of order $1/t$ and can only act to remove $o(1)$ constraints. Furthermore, inside a causal diamond much of the "in" evolution over time t acts on only a small part of the DOF. For amplitudes in which the asymptotic energy divides into clusters localized in space-like separated diamonds, the QPR guarantees that the "out" Hamiltonian along a trajectory that currently sees only part of the energy, has the same effect as an "in" Hamiltonian acting on the individual energies E_i.

3.1. Asymptotic Symmetries of HST

We have just seen that time translation symmetry in HST arises as an asymptotic symmetry. This is to be expected in a theory of gravity, but it's satisfying to see it arising from the quantum dynamics of an explicit model. We've also shown that rotation symmetry arises asymptotically, acting only on the decoupled jet degrees of freedom. It is asymptotic both because the organization of the DOF into spherical harmonics of a fixed rotation subgroup of the fuzzy volume preserving group depends on the asymptotic nesting of causal diamonds, and because rotations only act on the decoupled jets, which become truly independent of the rest of the system only in the limit of infinite proper time.

Spatial translation is more complicated. Part of it is programmed into the construction of the model, by using the same sequence of time dependent Hamiltonians along each time-like geodesic of Minkowski space. But the argument that scattering amplitudes are translation invariant comes from a combination of the QPR applied to asymptotically large diamonds along different geodesics, and the fact that jets decouple from "soft radiation" in the asymptotic limit. Note that the QPR alone is insufficient because the overlap between two spacelike separated diamonds has parametrically smaller entropy than either of the diamonds in the infinite T limit. Thus, the QPR says only that the density matrix on the overlap is maximally uncertain, subject to the constraints. The QPR also says that the angular location of the constraints seen along one trajectory, should look like a spatial translation of the angular locations as seen from the second. Since the bulk of the variables decouple and freeze in the large T limit, this suggests there is an identical Hilbert space, consisting of jets only, along the two relatively translated trajectories, and that the density matrices in that Hilbert space are related by a unitary transformation. Thus, spatial translation is an asymptotic symmetry as well.

4. CONCLUSIONS

The basic principles of the HST formalism are the CEP and the implementation of causality by the unfolding entanglement of degrees of freedom in a nested set of causal diamonds. The unitarity of the entangling map implies that this has the properties of an error correcting code. The fact that the fundamental variables are the fermionic generators in the fundamental representation of $SU(K|L)$[8] follows from the finite dimension of the Hilbert space and the fact that dynamics is invariant under fuzzy area preserving maps of the holographic screen. Area preserving invariance is valid for non-negative c.c. and for proper times sufficiently small compared to the AdS radius for negative c.c., as seen from the behavior of black hole quasi-normal modes. The fact that the fermionic operators must

[8] Equivalently, they are canonical fermions with constraints.

transform like asymptotic spinors, follows from the *presumed* Lorentz or AdS invariance of the boundary amplitudes, and the usual connection between spin and statistics.

For negative c.c. and proper times of order the AdS radius and larger, quasi-normal mode analysis suggests and AdS/CFT dictates, that propagation is ballistic on the screen, on distance scales in the bulk larger than the AdS radius. In order to make this compatible with the CEP for finite area causal diamonds with proper time approaching the AdS radius, we have to make a lattice field theory out of the fermionic variables and invoke the Lieb-Robinson bound. The unfolding entanglement map that implements propagation in proper time is then the inverse of a Tensor Network Renormalization Group map in the sense of Evenbly and Vidal [35].

Quantum information about small regions in the bulk is spread non-locally on the holographic screen, in a manner similar to that found in Error Correcting Codes. Omission of the information in a small area region of a big screen does not destroy the data about the small causal diamond, because the entanglement of the small diamond's variables with those of the large screen is shared uniformly among the large screen variables. The rate of spread over the area (which by the CEP is essentially the number of q-bits on the screen) undergoes a sort of phase transition in asymptotically AdS spaces, when the proper time in the diamond approaches the critical value (of order the AdS radius) at which the area of the diamond goes to infinity. Prior to that regime, the scrambling is "fast" and the information is homogenized on the sphere in a time of order the radius of the sphere times the logarithm of the total number of q-bits. As the proper time approaches the critical value, information scrambling is "fast" only over an area of order the $d-2$ power of the AdS radius. If we imagine a collection of local probes, separated by distances on the holographic screen which are of order the AdS radius, communication between those probes is ballistic. The system thus behaves like a lattice approximation to a quantum field theory.

In Banks and Fischler [36, 37] the authors conjectured that in the regime of the transition the HST Hamiltonian was the inverse of a Tensor Network Renormalization Group (TNRG) transformation [35]. TNRG transformations disentangle the degrees of freedom of a fine grained lattice field theory at its critical point, producing a Hamiltonian on a more coarse grained lattice. It's been shown by numerical analysis of simple one dimensional critical systems that the coarse grained Hamiltonians have a spectrum equal to that of the low lying levels of the radial quantization Hamiltonian of the CFT that describes the critical point. Radial quantization always picks out a particular element of the conformal group as the Hamiltonian, and this is tied to a particular timelike geodesic, making an explicit connection with HST. This is a direct implementation of the scale/radius duality of Maldacena. The TNRG can probably be improved via the technology of Pastawski et al. [24], which constructs tensor networks invariant under discrete subgroups of the conformal group.

The HST formalism adds an extra element to the TNRG formulation of asymptotically AdS dynamics. The TNRG Hamiltonian corresponding to some fixed proper time is conventionally defined to act as the unit operator on the tensor complement of the small Hilbert space corresponding to that causal diamond. In HST language, it is $H_{in}(t)+1$ for the last time slice in that diamond. In HST, we have a Hamiltonian $H_{in}(t)+H_{out}(t)$ where $H_{out}(t)$ acts on the tensor complement. In HST $H_{out}(t)$ is supposed to be determined by the consistency conditions with time evolution along other timelike trajectories. In AdS space, all time-like geodesics are related by elements of the conformal group, so at least some of these consistency conditions are guaranteed asymptotically by the restoration of conformal symmetry implicit in any RG transformation at a fixed point. It's possible that the conditions for finite diamonds and accelerated trajectories add further constraints.

This implementation of HST has implications for the CFT description of diamonds of size much smaller than the AdS radius. The work of Evenbly and Vidal shows that the finite dimensional Hamiltonians of the TNRG can be chosen to have the same spectrum as the low lying part of the exact CFT spectrum. This is a very explicit implementation of Maldacena's scale-radius duality. However, in the HST model, this correspondence breaks down as the proper time in the diamond is taken smaller than the AdS radius. Instead of a lattice field theory we have a highly degenerate Hamiltonian with area preserving map invariance and fast scrambling.

One can argue that this disturbing disconnection must be a property of the AdS/CFT correspondence without any reference to HST. Consider the causal diamond along a particular timelike geodesic in AdS space with proper time interval much smaller than the AdS radius. Now consider the Witten diagrams for a correlation function of a finite number of operators on $R\times S^{d-2}$. The vertices of the diagrams are integrated over the entire AdS space, whose spatial volume on global time slices is all concentrated near the boundary. Thus, the contribution to that correlation function from the causal diamond is very small, but non-zero. The probability that the interactions take place within that diamond is small and is dominated by contributions from the boundary of the diamond. As a consequence, if there is a notion of measurements localized in the diamond, they must register a state that is close to "empty Minkowski space," with deviations concentrated near the boundary of the diamond. This is consistent with the fact that the "energy" of such boundary states, in any coordinate system with spacelike slices localized in the diamond, will be very small. This is an AdS/CFT argument that the "vacuum" of the approximately Minkowski region is a nearly degenerate ensemble rather than a single pure state.

The standard derivation [38–41] of Minkowski amplitudes from CFT correlators illustrates the same principle. All of the work on this subject has concentrated on showing that specially prepared 4-point functions converge to tree level 4 point scattering amplitudes[9]. But now consider a $4+n$ point function with n operators not constrained to focus on a particular "arena" causal diamond. This gives a slightly different amplitude in Witten diagrams, but causes only small changes

[9]There are all sorts of caveats to this statement, particularly to its extension to higher point amplitudes which require us to study states localized in the large compact directions of $AdS_d\times K$, but they've been discussed elsewhere.

to the contribution from the bulk of the arena. The obvious interpretation of these correlators is as a superposition of amplitudes for "n gravitons" and other soft massless particles to be injected into or emitted from the arena. The full Minkowski scattering operator for "2 to 2" scattering includes all of these processes.

Note that in the discussion above there was no particular restriction to which n operators were inserted. We used only the fact that in connected diagrams involving all $4 + n$ particles, most of the vertices were integrated over all of AdS space. Thus, the precise definition of the limiting Hilbert space on which the Minkowski scattering operator acts requires us to find a basis of states that can reproduce all of these amplitudes. This is an unsolved problem in AdS/CFT.

If we accept the CEP for finite area diamonds in AdS space, Page's theorem [42] implies that the empty Minkowski vacuum is a maximally uncertain density matrix. In the large radius limit, the ensemble consists of all states on the boundary of the arena that can be created by Witten diagrams with little weight in the arena causal diamond. It is maximally uncertain because the number of possible Witten diagram states is much larger than the number of states that the CEP allows in the diamond. States that correspond to scattering in the arena must then be constrained states of this ensemble. That is, we've recovered the picture of localized excitations as constrained states of an ensemble of low energy excitations on the horizon.

In summary, HST treats time as fundamental, discrete and relative. Space-time is an emergent phenomenon, measuring the amount of quantum information accessible to an information gathering system on a given timelike trajectory in fixed intervals of proper time. This gives us a quantum definition of causal diamonds. Causal propagation is an error correcting code by which quantum information about events in a small diamond is entangled with the states of a larger diamond containing it. The information is spread rapidly over the holoscreen of the larger diamond, homogeneously for non-negative c.c., or for holoscreen sizes $\ll$ the AdS radius for negative c.c. On proper time scales of order the AdS radius, information is concentrated in the nodes of a tensor network and spreads ballistically over the network. The spatial size of the nodes is of order AdS radius, as is the spacing between them.

The principle, valid for small enough diamonds with any c.c. and any diamond with non-negative c.c., that bulk localized states are constrained states of boundary DOF, with bulk energy proportional to the number of constrained q-bits, explains most of the qualitative features of black hole and cosmological horizons and eliminates the firewall paradox. The HST model gives a very explicit picture of the transition between an ordinary scattering event and black hole formation. Black holes form when the energy, *the number of constraints*, entering into the past boundary of a causal diamond, creates a state so atypical that the fast scrambling Hamiltonian eliminates those constraints before the energy exits the diamond. Thus, the boundary of validity of effective field theory ideas is an entropy bound. The bulk localized entropy must be less that $S^{3/4}$ in four dimensions in order to avoid black hole formation. This is a covariant version of a bound conjectured by Cohen et al. [1]. Their bound was based on trying to understand the failure of field theory to compute the cosmological constant. In HST, the c.c. is an input, but it is correct that the expectation value of the Hamiltonian in dS space does scale like the integral of the c.c. over the spatial volume of the static patch.

The resolution of the firewall paradox for non-negative c.c. is also completely entropic and can be understood without any of the details of the HST formalism. The fact that black holes in these space-times have negative specific heat implies that the state just prior to the event we call "dropping a low entropy system onto a black hole," has a huge entropy deficit relative to the equilibrated black hole of slightly higher mass. If there is any notion at all of a finite dimensional Hilbert space associated with the equilibrated system, then the pre-equilibrium state must be a low entropy constrained state in that Hilbert space. Combining this with the fact that dynamics on the horizon has a natural time scale of order the black hole radius, and the natural conjecture that the frozen degrees of freedom mediate the interactions between the low entropy system and the original black hole horizon[10], we find that there will be a time of order $R_S \ln R_S$ during which the infalling system behaves as if the black hole were not there. This is the "temporary" definition of "the smooth part of the black hole interior." After the scrambling time the last phrase in scare quotes has no meaning, but a new infalling system will create its own interior. Firewalls are a consequence of insisting on an invalid quantum field theory picture of quantum states near the horizon.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported in part by the U.S. Department of Energy under grant DE-SC0010008.

ACKNOWLEDGMENTS

I'd like to thank W. Fischler for years of joint work that went into formulating the HST formalism described here.

[10]In the HST model the constraints are precisely those that imply the off diagonal matrix elements of the matrix connecting the "black hole block" and the "infalling block," vanish, so that the two blocks don't interact.

REFERENCES

1. Cohen AG, Kaplan DB, Nelson AE. Effective field theory, black holes, and the cosmological constant. *Phys Rev Lett.* (1999) **82**:4971. doi: 10.1103/PhysRevLett.82.4971
2. Jacobson T. Thermodynamics of space-time: the Einstein equation of state. *Phys Rev Lett.* (1995) **75**:1260.
3. Fischler W, Susskind L. Holography and cosmology. *arXiv*. hep-th/9806039
4. Bousso R. A covariant entropy conjecture. *J High Energy Phys.* (1999) **9907**:004.

5. Milne EA. *Relativity, Gravitation and World Structure*. Oxford: Oxford University Press (1935).
6. Gibbons GW, Hawking SW. Cosmological event horizons, thermodynamics, particle creation. *Phys Rev D.* (1977) **15**:2738. doi: 10.1103/PhysRevD.15.2738
7. Verlinde EP. Emergent gravity and the dark universe. *Sci Post Phys.* (2017) **2**:016. doi: 10.21468/SciPostPhys.2.3.016
8. Hayden P, Preskill J. Black holes as mirrors: quantum information in random subsystems. *J High Energy Phys.* (2007) **0709**:120. doi: 10.1088/1126-6708/2007/09/120
9. Sekino Y, Susskind L. Fast scramblers. *J High Energy Phys.* (2008) **0810**:065. doi: 10.1088/1126-6708/2008/10/065
10. Braunstein SL, Pirandola S, Zyczkowski K. Better late than never: information retrieval from black holes. *Phys Rev Lett.* (2013) **110**:101301. doi: 10.1103/PhysRevLett.110.101301
11. Almheiri A, Marolf D, Polchinski J, Sully J. Black holes: complementarity or firewalls? *arXiv.* 1207.3123.
12. Banks T, Fischler W, Kundu S, Pedraza JF. Holographic space-time and black holes: mirages as alternate reality. *arXiv.* 1401.3341.
13. Chen YA, Kapustin A, Radicevic D. Exact bosonization in two spatial dimensions and a new class of lattice gauge theories. *Ann Phys.* (2018) **393**:234. doi: 10.1016/j.aop.2018.03.024
14. Chen YA, Kapustin A. Bosonization in three spatial dimensions and a 2-form gauge theory. *Phys Rev B.* (2019) **100**:245127. doi: 10.1103/PhysRevB.100.245127
15. Banks T. Fermi/Pauli duality in arbitrary dimension. *arXiv.* 1908.10453.
16. Banks T, Kehayias J. Fuzzy geometry via the spinor bundle, with applications to holographic space-time and matrix theory. *Phys Rev D.* (2011) **84**:086008. doi: 10.1103/PhysRevD.84.086008
17. Banks T. Current algebra on the conformal boundary and the variables of quantum gravity. *arXiv.* 1511.01147.
18. Banks T. The temperature/entropy connection for horizons, massless particle scattering, and the origin of locality. *Int J Mod Phys D.* (2015) **24**:1544010. doi: 10.1142/S0218271815440101
19. Banks T. The super BMS algebra, scattering and holography. *arXiv.* 1403.3420.
20. Awada MA, Gibbons GW, Shaw WT. Conformal supergravity, twistors and the super BMS group. *Ann Phys.* (1986) **171**:52. doi: 10.1016/S0003-4916(86)80023-9
21. Banks T, Fischler W. Holographic theory of accelerated observers, the S-matrix, and the emergence of effective field theory. *arXiv.* 1301.5924.
22. Swingle B. Entanglement renormalization and holography. *Phys Rev D.* (2012) **86**:065007. doi: 10.1103/PhysRevD.86.065007
23. Almheiri A, Dong X, Harlow D. Bulk locality and quantum error correction in AdS/CFT. *J High Energy Phys.* (2015) **1504**:163. doi: 10.1007/JHEP04(2015)163
24. Pastawski F, Yoshida B, Harlow D, Preskill J. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *J High Energy Phys.* (2015) **1506**:149. doi: 10.1007/JHEP06(2015)149
25. Mintun E, Polchinski J, Rosenhaus V. Bulk-boundary duality, gauge invariance, and quantum error corrections. *Phys Rev Lett.* (2015) **115**:151601. doi: 10.1103/PhysRevLett.115.151601
26. Freivogel B, Jefferson R, Kabir L. Precursors, gauge invariance, and quantum error correction in AdS/CFT. *J High Energy Phys.* (2016) **1604**:119. doi: 10.1007/JHEP04(2016)119
27. Harlow D. The Ryu-Takayanagi formula from quantum error correction. *Commun Math Phys.* (2017) **354**:865. doi: 10.1007/s00220-017-2904-z
28. Kim IH, Kastoryano MJ. Entanglement renormalization, quantum error correction, bulk causality. *J High Energy Phys.* (2017) **1704**:040. doi: 10.1007/JHEP04(2017)040
29. Harlow D. TASI lectures on the emergence of bulk physics in AdS/CFT. *PoS TASI.* (2018) **2017**:002. doi: 10.22323/1.305.0002
30. LŽvay P, Holweck F. Finite geometric toy model of spacetime as an error correcting code. *Phys Rev D.* (2019) **99**:086015. doi: 10.1103/PhysRevD.99.086015
31. Kamal H, Penington G. The Ryu-Takayanagi formula from quantum error correction: an algebraic treatment of the boundary CFT. *arXiv.* 1912.02240.
32. Bao N, Cheng N. Eigenstate thermalization hypothesis and approximate quantum error correction. *J High Energy Phys.* (2019) **1908**:152. doi: 10.1007/JHEP08(2019)152
33. Hirai H. Shrinking of operators in quantum error correction and AdS/CFT. *J High Energy Phys.* (2019) **1912**:128. doi: 10.1007/JHEP12(2019)128
34. Vaid D. Quantum error correction in loop quantum gravity. *arXiv.* 1912.11725.
35. Evenbly G, Vidal G. Tensor network renormalization. *arXiv.* 1412.0732.
36. Banks T, Fischler W. Holographic space-time models of anti-deSitter space-times. *arXiv.* 1607.03510.
37. Banks T, Fischler W. Soft gravitons and the flat space limit of anti-deSitter space. *arXiv.* 1611.05906.
38. Polchinski J. S matrices from AdS space-time. *arXiv.* hep-th/9901076.
39. Susskind L. Holography in the flat space limit. *AIP Conf Proc.* (1999) **493**:98. doi: 10.1063/1.1301570
40. Banks T, Fischler W. Holographic space-time and Newton's law. *arXiv.* 1310.6052.
41. Banks T, Fischler W. Holographic theory of accelerated observers, the S-matrix, and the emergence of effective field theory. *arXiv.* 1301.5924.
42. Page DN. Average entropy of a subsystem. *Phys Rev Lett.* (1993) **71**:1291. doi: 10.1103/PhysRevLett.71.1291

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

BRIEF RESEARCH REPORT
published: 28 May 2020
doi: 10.3389/fphy.2020.00176

On the Testability of the Equivalence Principle as a Gauge Principle Detecting the Gravitational t^3 Phase

Chiara Marletto [1,2,3]* *and Vlatko Vedral* [1,2,3]

[1] Clarendon Laboratory, University of Oxford, Oxford, United Kingdom, [2] Centre for Quantum Technologies, National University of Singapore, Singapore, Singapore, [3] Department of Physics, National University of Singapore, Singapore, Singapore

There have been various claims that the Equivalence Principle, as originally formulated by Einstein, presents several difficulties when extended to the quantum domain, even in the regime of weak gravity. Here we point out that by following the same approach as used for other classical principles, e.g., the principle of conservation of energy, one can, for weak fields, obtain a straightforward quantum formulation of the principle. We draw attention to a recently performed test that confirms the Equivalence Principle in this form and discuss its implications.

Keywords: quantum gravity, quantum phase estimation, equivalence principle, interferometry, quantum information
PACS numbers: 03.67.Mn, 03.65.Ud

OPEN ACCESS

Edited by:
Caslav Brukner,
University of Vienna, Austria

Reviewed by:
Alberto Martin Gago,
Pontifical Catholic University of Peru, Peru
Alexander Randolph Hilliard Smith,
Dartmouth College, United States

***Correspondence:**
Chiara Marletto
chiara.marletto@gmail.com

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: *28 January 2020*
Accepted: *24 April 2020*
Published: *28 May 2020*

Citation:
Marletto C and Vedral V (2020) On the Testability of the Equivalence Principle as a Gauge Principle Detecting the Gravitational t^3 Phase. Front. Phys. 8:176. doi: 10.3389/fphy.2020.00176

The principle of equivalence is a pillar of Einstein's theory of relativity; as such, it was originally formulated within a classical theory, where all observables of a point particle, in particular its position, energy and mass, are sharp in any state of the particle. This is true of other principles, such as the principle of conservation of energy, whose expression and validity in quantum theory are nonetheless widely accepted. Yet, there has been a great degree of controversy about the formulation of the Equivalence Principle for quantum systems: this is because a quantum system can exist in a spatial superposition; and the Equivalence Principle as classically formulated does not cover directly such cases. Consequently, there have been proposals to extend it to quantum systems [1–3]; there also have been claims that the principle is violated by quantum systems (see e.g., the introduction in Anastopoulos and Hu [4] and references herein); some have also claimed that this should be a reason for gravitational state reduction [5]. The main point of discussion here is the fact that the Equivalence Principle implies that different masses should fall at the same rate in the same gravitational field. However, the quantum de Broglie wavelength is a function of the particle's mass and therefore different masses would interfere differently in the same gravitational field; this would seem to violate the prescription, from the Principle of Equivalence, that the behavior of different masses in the same field cannot be distinguished. As we will see below, in our formulation of the quantum Equivalence Principle, this is not a relevant issue. The same we believe is true for other aspects of the controversy, such as those mentioned in Anastopoulos and Hu [4].

Here we would like to extend the Equivalence Principle to the quantum domain via a similar approach to that applied to the energy conservation. Namely, to extend the principle to the quantum domain, we will assume that for any branch of a quantum superposition the principle holds true. Specifically, we will assume that for each branch of a spatial superposition, sharp at location x, the Equivalence Principle holds in one of its currently accepted forms: *the state of motion of a point particle at rest in a uniform gravitational field* $\mathbf{g}$ *is empirically indistinguishable, by local operations at* x*, from the state of motion of a point particle that undergoes an acceleration* $-\mathbf{g}$ *in*

a gravity-free region. To the best of our knowledge, this formulation of the Equivalence Principle is original; however, there are other recent proposals that are similar in spirit [1]. In the limit of weak gravity this is a good approximation; the ultimate formulation of the Equivalence Principle will have to lift this assumption and it should not rely on the idea of a fixed spacetime background. For present purposes, however, it is possible to confine attention to this regime, as even in this regime problems with the quantum formulation of the Equivalence Principle have been claimed to exist. This regime is also very interesting, as it does not involve general-relativistic effects, but it can be used to probe the quantum nature of gravity [6, 7]. We will now derive a number of well-known consequences of the principle formulated in this way and point out a recent experimental confirmation of its validity.

THE PHASES INDUCED BY THE EQUIVALENCE PRINCIPLE

Assuming the Equivalence Principle, in classical mechanics the transformation between a system in a gravitational field and the one in the equivalent accelerated frame is a gauge transformation. Assuming, with the Equivalence Principle, that the inertial and gravitational mass are the same constant m, the Lagrangian for a particle with mass m moving in an accelerated frame with the acceleration rate given by g is:

$$L_F = \frac{1}{2}m(\dot{x} + gt)^2 \,,$$

whereas for a particle moving in the gravitational potential $m_g gx$, the Lagrangian is

$$L_G = \frac{1}{2}m(\dot{x})^2 - mgx = L_F + \frac{d}{dt}\Lambda(x,t)$$

where

$$\Lambda = -mgxt - \frac{1}{6}mg^2t^3$$

is the gauge transformation. For simplicity we only consider a one-dimensional motion, but this results in no loss of generality.

Following the Equivalence Principle, as formulated above, the wavefunctions of the particle described in the two coordinate systems (the freely falling and the g frame), as expressed in the position basis, are related in the following way [8, 9]:

$$|\Psi(x,t)\rangle_G = e^{-\frac{i}{\hbar}\Lambda(x,t)}|\Psi(x,t)\rangle_F \,.$$

This can easily be verified by applying the Equivalence Principle in the above form to the corresponding Schrödinger equations: one, where the particle is freely falling; the other, where the particle is in a uniform gravitational field. The above relation holds for all the solutions of the Schrödinger Equation [9]. Now, as expected, the gauge transformation between the two coordinate systems (the freely falling and the one in the uniform gravitational field) is quantum mechanically reflected in the appearance of the extra phase factor between the two corresponding quantum states. This can be seen as a consequence of the fact that, upon quantization, the classical Lagrangian becomes the phase factor which constitutes the basis of the path-integral formulation of quantum physics. It may at first be surprising that the gauge transformation is not just $-mgxt$ and that there is an additional term proportional to t^3. Mathematically, it is of course clear that this term is needed since the freely falling Lagrangian has a t^2 term which, in order for the Equivalence Principle to hold, needs to be canceled by the additional term in the gauge.

We note in passing that General Relativity would introduce corrections to this phase. One way of thinking about it is that time simply flows at different rates at the two different heights. The difference between the two flows is given by

$$\Delta t = \left(\sqrt{1 - \frac{2GM}{c^2 x}} - \sqrt{1 - \frac{2GM}{c^2(x+\Delta h)}}\right)\tau \qquad (1)$$

$$= \left(\frac{GM}{c^2x^2}\Delta h + \frac{1}{2}\left(\frac{GM}{c^2x^2}\Delta h\right)^2 + ...\right)\tau \qquad (2)$$

The phase difference can now be calculated by multiplying this by $\omega = mc^2/\hbar$ and so

$$\Delta\phi = mg\Delta h\tau/\hbar + mg^2\frac{\Delta^2 h}{c^2}\tau/\hbar + ... \qquad (3)$$

The first term in the gauge transformation was observed in the Colella-Overhauser-Werner (COW) experiment [10], implementing the interference of a single neutron superposed across two different heights, each experiencing different Earth gravitational fields. The potential difference between the paths $mg\Delta h$ leads to the phase difference $\Delta\phi = mg\Delta ht/\hbar$ between the two neutron states. In a variant of the COW experiment, a neutron interferometry was also performed in a uniformly accelerated interferometer, confirming the same results as the original COW experiment and thus, indirectly, the validity of the Equivalence Principle in the above form [11]. The second term in this expression is the GR correction and it is clearly much smaller than the Newtonian one, therefore harder to access experimentally. As far as we can tell, it has never been observed experimentally.

Here we want to focus on the following interesting question: is the t^3 term observable? We are used to hearing that gauge transformations are unobservable, so an immediate response would be "no." However, as we will explain, it is still possible to observe the t^3 as a relative phase between the branches of a quantum superposition of two distinct reference frames.

THE T^3 PHASE IS OBSERVABLE

To observe the t^3 term, we need to perform an experiment that effectively superposes the freely falling and the gravitational field states of motion for a single particle. This is the main point of our paper. The COW experiment can certainly be analyzed from the perspective of both reference frames and the

two treatments ought to give the same predictions due to the Equivalence Principle. But the COW experiment does not involve the superposition of two frames, it involves the neutron being in a superposition of different spatial paths (in Earth's frame).

We discuss here a thought experiment where a single particle is in a superposition of two states, one where it is freely falling in a gravitational field and the other where it is static in the same gravitational field. Conceptually, this experiment could in principle be achieved as follows.

Consider a quantum system of mass m superposed across two different locations, e.g., in the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and placed horizontally in an ideal, uniform gravitational field, as in the COW experiment [10]. Suppose the two branches $|0\rangle$ and $|1\rangle$ are a distance d apart. As confirmed by the COW experiment, the contribution to the phase due to this field is equal on the two arms, therefore amounting to no phase difference. The interferometer can now be tilted to the vertical position (see **Figure 1**), so that the potential on the arm labeled as $|0\rangle$ differs from that on the branch $|1\rangle$ by the amount mgd. At that point, the mass on say the branch $|0\rangle$ is dropped, and let interfere with the branch $|1\rangle$.

The phase difference between the two paths, assuming the Equivalence Principle to hold in every branch of the wave-function, and that $x(t) = d - \frac{1}{2}gt^2$, reads:

$$\frac{1}{\hbar}\int_0^t mgx(\tau)d\tau = \frac{mg}{\hbar}(dt - \frac{1}{6}gt^3)$$

The freely-falling branch of the mass interferes with the stationary branch and the resulting shift of the interference fringes contains the t^3 term, which can therefore in principle be detected. Whether we think of this experiment as a particle in a superposition of two different states of motion, or we think of the particle as an observer and therefore in a superposition of two reference frames, is irrelevant. The result ought to be the same, and the experiment would confirm the Equivalence Principle if the t^3 term were to be observed, at least in this quantum weak-field regime.

In order to observe the phase of course no decoherence should occur. This includes any entanglement of the interfering mass to another system that does not participate in the final interference. This other system could be another simple physical system (such as a colliding atom or the electromagnetic field), or it could be an internal degree of freedom [4]. More subtly, no observer should measure which of the two reference frames the mass is in, as this too would cause decoherence.

We note that an experiment to detect the t^3 phase was proposed and performed very recently, with a Bose-Einstein condensate undergoing interference in a Stern-Gerlach interferometer, subject to a state-dependent force [12]. The experiment we have discussed is conceptually equivalent. Measuring the t^3 in this experiment has confirmed the predictions of the Equivalence Principle expressed as above in the quantum domain.

DISCUSSION

We have deployed the relation between the freely-falling and the gravitational coordinate systems, which is based on the Equivalence Principle in the weak-field limit, to discuss an experiment that could detect the t^3 phase that relates a mass in free fall in a gravitational field, and the same mass being stationary in the same field. We have derived this relation using a formulation of the Equivalence Principle that holds for each branch of a spatial superposition. This expression of the Equivalence Principle harmonizes with the existing ways of extending other principles, originally formulated for classical systems (e.g., energy conservation), to the quantum world. We would also like to point out that the transformation we applied to change the freely falling into the gravitational one is the weak-field limit of the Rindler coordinate transformation [13]. In this limit, it is clearly legitimate to neglect the effects, such as Unruh-Davies [5, 14], because only very weak fields are considered. Moreover, in the Rindler setting, the time coordinate between the inertial and the accelerated observer (with acceleration a) transforms as $t' = \frac{c}{a}\text{arcsinh}\,\frac{a}{c}t$, where c is the speed of light. Now, the second term in the Taylor expansion of this expression is equal to $\frac{a^2}{6c^2}t^3$ which, when multiplied by mc^2 gives us the exact t^3 phase term above. Therefore, the low acceleration limit of Rindler's coordinate transformation is perfectly consistent with our analysis. We conclude by pointing out that even though our

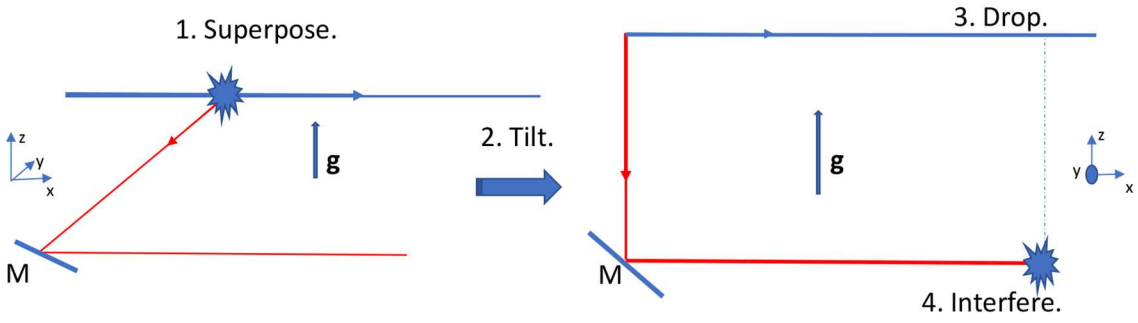

FIGURE 1 | Schematics of the interferometry to measure the t^3 phase. On the left, the particle is superposed across two paths (M is a mirror; **g** the gravitational field). On the right, after tilting the interferometer in the vertical direction, the interferometry is closed by dropping the particle if it is on the upper branch.

discussions are in the Newtonian, weak-field quantum regime, the rest energy of the particle still somehow needs to be taken into account. This is best seen from the perspective of the relativistic action. The difference between the actions in the falling and the stationary coordinate systems can be expressed as: $mc^2(t - T)$, where t and T are the proper times of the two coordinate systems. In the lowest two orders of expansion, this difference reads: $-mgxt - \frac{1}{6}mg^2t^3$. The phase difference between two branches is then, just like in the COW experiment, the difference between the time flows in those two branches (multiplied by $\omega = mc^2/\hbar$).

Our guiding philosophy here has been to take quantum physics seriously and assume that it applies to all systems and all degrees of freedom. This means, in particular, that if any two states of motion of a particle are possible, such as an inertial and an accelerated state of a mass, then their superposition is also possible. Our paper indeed has been exploring the consequences for the relative phase between two such states, which we claim to be observable (and which the experiment in Amit et al. [12] has indeed observed).

One could also think of the measuring system being in a superposition of two different states of motion. For instance, a detector could be in a superposition of being inertial and being accelerated, while measuring another physical system in a sharp state of motion. None of this is a problem to handle quantum mechanically (as far as we can tell), but one has to be careful not to make some unwarranted assumptions. For instance, a particle being in a superposition of different motions while the detector is inertial, will in general yield different results to the detector being in different states of motion while the particle is inertial. If quantum mechanics is assumed to hold universally, none of these situations presents a difficulty: the physical systems involved will have well-defined behaviors that perfectly comply with the quantum postulates.

In summary, given the observability of the t^3 term, it seems to us that at least in the regime of weak-field there should be no qualms about considering the Equivalence Principle to be extended into the quantum domain in the same way as all other classical principles are. However, there are still many interesting other open issues to be investigated both theoretically, e.g., references [1, 2], as well as experimentally.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

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

FUNDING

CM thanks the John Templeton Foundation and the Eutopia Foundation. VV's research was supported by the National Research Foundation and the Ministry of Education in Singapore and administered by Centre for Quantum Technologies, National University of Singapore. This publication was made possible through the support of the ID 61466 grant from the John Templeton Foundation, as part of the The Quantum Information Structure of Spacetime (QISS) Project (qiss.fr). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

REFERENCES

1. Hardy L. Implementation of the quantum equivalence principle. *arXiv.* (2019) 1903.01289.
2. Zych M, Brukner C. Quantum formulation of the Einstein equivalence principle. *Nat Phys.* (2018) **14**:1027–31. doi: 10.1038/s41567-018-0197-6
3. Rosi G, D'Amico G, Cacciapuoti L, Sorrentino F, Prevedelli M, Zych M, et al. Quantum test of the equivalence principle for atoms in coherent superposition of internal energy states. *Nat Commun.* (2017) **8**:15529. doi: 10.1038/ncomms15529
4. Anastopoulos C, Hu BL. Equivalence principle for quantum systems: dephasing and phase shift of free-falling particles. *Class Quant Grav.* (2018) **35**:035011. doi: 10.1088/1361-6382/aaa0e8
5. Howl R, Penrose R, Fuentes I. Exploring the unification of quantum theory and general relativity with a Bose-Einstein condensate. *New J Phys.* (2019) **21**:043047. doi: 10.1088/1367-2630/ab104a
6. Bose S, Mazumdar A, Morley GW, Ulbricht H, Toroš M, Paternostro M, et al. Spin entanglement witness for quantum gravity. *Phys Rev Lett.* (2017) **119**:240401. doi: 10.1103/PhysRevLett.119.240401
7. Marletto C, Vedral V. Gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity. *Phys Rev Lett.* (2017) **119**:240402. doi: 10.1103/PhysRevLett.119.240402
8. Greenberger DM, Overhauser AW. Coherence effects in neutron diffraction and gravity experiments. *Rev Mod Phys.* (1979) **51**:43–78.
9. Nauenberg M. Einstein's equivalence principle in quantum mechanics revisited. *Am J Phys.* (2016) **84**:879. doi: 10.1119/1.4962981
10. Colella R, Overhauser AW, Werner SA. Observation of gravitationally induced quantum interference. *Phys Rev Lett.* (1975) **34**:1472–5.
11. Bonse U, Wrloblewsky B. Measurement of neutron quantum interference in noninertial frames. *Phys Rev Lett.* (1983) **51**:1401–4. doi: 10.1103/PhysRevLett.51.1401
12. Amit O, Margalit Y, Dobkowski O, Zhou Z, Japha Y, Zimmermann M, et al. T^3 Stern-Gerlach Matter-wave interferometer. *Phys Rev Lett.* (2019) **123**:083601. doi: 10.1103/PhysRevLett.123.083601
13. Birrel N, Davies P. *Quantum Field Theory in Curved Spacetime*. Cambridge: Cambridge University Press (1977).
14. Fuentes I, Mann RB. Alice falls into a black hole: entanglement in non-inertial frames. *Phys Rev Lett.* (2005) **95**:120404. doi: 10.1103/PhysRevLett.95.120404

Conflict of Interest: 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.

frontiers in Physics

ORIGINAL RESEARCH
published: 03 July 2020
doi: 10.3389/fphy.2020.00207

On the Possibility of Experimental Detection of the Discreteness of Time

Marios Christodoulou[1]* *and Carlo Rovelli*[2,3,4]

[1] Department of Computer Science, The University of Hong Kong, Hong Kong, China, [2] CPT, Aix Marseille Univ, Université de Toulon, CNRS, Marseille, France, [3] Perimeter Institute for Theoretical Physics, Waterloo, ON, Canada, [4] The Rotman Institute of Philosophy, London, ON, Canada

The Bose-Marletto-Vedral (BMV) experiment tests a quantum gravitational effect predicted by low energy perturbative quantum gravity. It has received attention because it may soon be within observational reach in the lab. We point out that: (i) in relativistic language, the experiment tests an interference effect between proper-time intervals; (ii) the feasibility study by Bose et al. suggests that current technology could allow to probe differences of such proper-time intervals of the order of 10^{-38} seconds, about twenty orders of magnitude beyond the current resolution of the best atomic clocks; (iii) the difference of proper times approaches Planck time (10^{-44} s) if the masses of the particles in the experiment approach the Planck mass (∼micrograms). This implies that the experiment might open a window on the structure of time at the Planck scale. We show that if time differences are discrete at the Planck scale—as research in quantum gravity may suggest—the Planckian discreteness of time would appear as quantum levels of an in principle measurable entanglement entropy.

Keywords: quantum gravity, time discreteness, entanglement, non perturbative effects, quantum gravity phenomenology

OPEN ACCESS

Edited by:
Caslav Brukner,
University of Vienna, Austria

Reviewed by:
Igor Pikovski,
Stockholm University, Sweden
Bhupal Dev,
Washington University in St. Louis, United States

***Correspondence:**
Marios Christodoulou
christod.marios@gmail.com

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: *28 December 2019*
Accepted: *14 May 2020*
Published: *03 July 2020*

Citation:
Christodoulou M and Rovelli C (2020) On the Possibility of Experimental Detection of the Discreteness of Time. Front. Phys. 8:207. doi: 10.3389/fphy.2020.00207

1. INTRODUCTION

Bose et al. [1] and Marletto and Vedral [2, 3] have proposed an ingenious idea to amplify and observe minuscule quantum gravitational effects in a table-top experiment. The idea has received considerable attention [4–14]. In the version proposed in Bose et al. [1] the observable signal is given by Bell–like correlations among the spins of two particles. The correlations are produced by a gravitational interaction. If we assume the identification of gravity with spacetime geometry which is at the basis of general relativity, the observation of these correlations implies that spacetime geometry can be in a quantum superposition (in a non-semiclassical state), and therefore can be taken as evidence for quantum behavior of the geometry [10].

The Bose-Marletto-Vedral (BMV) effect is predicted by low energy perturbative quantum gravity, and hence by any approach to quantum gravity consistent with this low energy expansion, including string theory and loop quantum gravity. It is therefore plausibly real. If detected, it would provide indirect empirical evidence that spacetime geometry does obey quantum mechanics.

On the other hand, the BMV effect is insensitive to the limit $c \to \infty$, hence the gravitational interaction involved can be described in a non–relativistic language. For this reason, it does not test the full relativistic quantum gravitational regime. In fact, the effect can be accounted for purely in terms of the scalar non-radiative modes of the gravitational field, hence it does not test the quantum *dynamics* of gravity. If we do not fold in the relativistic information provided by classical

general relativity, the effect can be interpreted in terms of Newton's action at a distance, leading to objections on the relevance of the effect for the quantum properties of spacetime [4].

Here, we point out that a refinement of the BMV effect *could* open a window on a quantum gravitational effect that would definitely not be accounted for by non-relativistic quantum physics: time discreteness. The reason this is possible is that from the point of view of general relativity the BMV set up is a delicate interference apparatus that picks up a tiny difference $\delta\tau$ in proper time between two quantum branches, due to gravitationally-induced time dilatation. As derived below, the phase difference $\delta\phi$ responsible for the gravity-mediated entanglement, and giving rise to the effect, can be written in a particularly simple form

$$\delta\phi = \frac{m}{m_{Pl}}\,\frac{\delta\tau}{t_{Pl}}, \tag{1}$$

where the experiment is performed with particles of mass m, m_{Pl} is the Planck mass and t_{Pl} the Planck time. This expression shows that the time scales probed are extremely small. With current technology, the BMV effect might be detected in the lab by probing relevant entanglement (generated when $\delta\phi \sim 1$) using mesoscopic particles, with masses of the order of a millionth of a Planck mass ($m \sim 10^{-6} m_P \sim 10^{-14} kg$) [1]. The corresponding time dilation is of the order of a million Planck times, $\delta\tau \sim 10^6 t_P \sim 10^{-38} s$. This is about twenty orders of magnitude above current capabilities of direct time measurements with atomic clocks [15].

Now, it is often pointed out in quantum gravity research that the Planck time t_{Pl} could be a minimal observable time; this follows from relativity plus the fact that many approaches to quantum gravity predict a minimal length [16]. In loop quantum gravity for instance, there is indirect evidence of discretization of time coming from loop quantum cosmology [17, 18] and other arguments [19, 20]. The simplest possibility is to assume that a measurement of a time lapse can only yield multiples of the Planck time. If this holds for the time difference $\delta\tau$, namely if $\delta\tau = n\, t_{Pl}$, with integer n, then

$$\delta\phi = n\frac{m}{m_{Pl}}. \tag{2}$$

That is, a discontinuity in $\delta\tau$ could be detected as a discontinuity in $\delta\phi$. As discussed below, such a discretization of the phase could be detected by the Bell–like correlations among the particles' spins, which would acquire a characteristic quantum band structure.

The extremely small time intervals probed by the current proposal to implement the BMV experiment are still too large to see time discreteness. But if the experiment can be pushed to work with more massive particles, further approaching the Planck mass, $\delta\tau$ will approach the Planck time [see (1)]. While the Planck time t_P is at the—so far—deeply inaccessible scale $t_P \sim 10^{-44} s$, the Planck mass is an easily accessible scale ($\sim$ micrograms). Thus, by directly manipulating quantum superpositions of Planck mass particles, interference as a result of gravitational attraction we can indirectly probe time at the Planck scale. This is the key theoretical observation of this paper.

The analysis that follows is rough and the effect might be questioned by a more detailed investigation. It may turn out that the BMV apparatus does not measure eigenvalues but rather expectation values, or, that the scale of discreteness for differences in duration is actually smaller that Planckian. Nevertheless, a prospect of experimental access to the scale of the Planck time is so interesting to deserve full attention.

2. THE BMV EXPERIMENT

Let us start by describing the version of the BMV experiment of Bose et al. [1] in relativistic language, as in Christodoulou et al. [10]. Two mesoscopic particles (a and b) of mass m and embedded [21] spin $\frac{1}{2}$ are quantum split (say with a Stern-Gerlach-like apparatus) and each is set in a superposition of two distinct states, say with spins $+$ and $-$ in some basis, with different positions in space. This gives rise to four different branches, which we denote $|++\rangle, |+-\rangle, |-+\rangle$, and $|--\rangle$ and a tensor state

$$\begin{aligned} |\psi\rangle &= \frac{|+\rangle_a + |-\rangle_a}{\sqrt{2}} \otimes \frac{|+\rangle_b + |-\rangle_b}{\sqrt{2}} \\ &= \frac{|{+}{+}\rangle + |{+}{-}\rangle + |{-}{+}\rangle + |{-}{-}\rangle}{2}. \end{aligned} \tag{3}$$

After a time t the two components of each particle are recombined. The relative positions of the particles differ in the distinct branches during the time t, giving rise to different gravitational fields, namely different spacetime geometries. Therefore, during the interval t the quantum state of the geometry is in a superposition of four (semiclassical) spacetimes, each corresponding to a classical metric. In particular, the proper time τ along the worldline of one particle is affected by the presence of the other by relativistic time dilation. This effect is obviously very small, but, as we shall see, it may be picked up by interference.

For simplicity, consider the case in which the two particles are kept at a small distance d only in a single branch, say $|--\rangle$, while in the other three branches the time dilation is negligible. According to general relativity, the gravitational time dilation is [10]

$$\delta\tau = \frac{Gm}{dc^2}\, t. \tag{4}$$

where G is the Newton constant and c the speed of light. The phase of the quantum state of a particle of mass m evolves in time as $e^{i\phi} = e^{imc^2\tau/\hbar}$. Therefore, after a (laboratory frame) time t the $|--\rangle$ branch picks up a phase difference

$$\delta\phi = \frac{mc^2}{\hbar}\delta\tau \tag{5}$$

with respect to the other branches. This equation is equivalent to Equation (1).

After the time t the state of the two particles has become

$$|\psi\rangle = \frac{|{+}{+}\rangle + |{+}{-}\rangle + |{-}{+}\rangle + e^{i\delta\phi}\,|{-}{-}\rangle}{2}. \tag{6}$$

This is an entangled state. The amount of entanglement is measured by the entanglement entropy

$$I = \mathrm{Tr}[\rho \ln \rho] \tag{7}$$

where

$$\rho = \mathrm{Tr}_b\, |\psi\rangle\langle\psi| \tag{8}$$

the trace being on the spin states of one of the two particles. A quick calculation gives

$$\rho = \frac{1}{2}\Big(|+\rangle\langle+| + |-\rangle\langle-|\Big) + \frac{e^{-i\delta\phi}+1}{4}\,|+\rangle\langle-| + \frac{e^{i\delta\phi}+1}{4}\,|-\rangle\langle+|\,. \tag{9}$$

This is correctly a hermitian matrix of unit trace. To compute the entropy we need to diagonalize ρ. A straightforward calculation gives the eigenvalues

$$\rho_\pm = \frac{1}{2} \pm \frac{\sqrt{1+\cos\delta\phi}}{2\sqrt{2}} \tag{10}$$

When $\delta\phi = 0$, the eigenvalues are $\rho_+ = 1$ and $\rho_- = 0$, thus giving vanishing entanglement entropy, i.e., there is no interference in the output. When $\delta\phi = \pi$, we have $\rho_+ = 1/2$ and $\rho_- = 1/2$; the state is maximally entangled and $I = \log 2$, i.e., we observe the BMV effect. For a general $\delta\phi$, the entanglement entropy is

$$\begin{aligned} I &= -\rho_+ \ln\rho_+ - \rho_- \ln\rho_- \\ &= -\left(\frac{1}{2} + \frac{\sqrt{1+\cos\delta\phi}}{2\sqrt{2}}\right)\ln\left(\frac{1}{2} + \frac{\sqrt{1+\cos\delta\phi}}{2\sqrt{2}}\right) \\ &\quad -\left(\frac{1}{2} - \frac{\sqrt{1+\cos\delta\phi}}{2\sqrt{2}}\right)\ln\left(\frac{1}{2} - \frac{\sqrt{1+\cos\delta\phi}}{2\sqrt{2}}\right). \end{aligned} \tag{11}$$

See **Figure 1**. In the lab, for a given mass m and distance d, $\delta\phi$ can be controlled by modulating t via

$$\delta\phi = \frac{Gm^2}{d\hbar}\, t. \tag{12}$$

that follows from (4) and (5). The entanglement entropy can be measured by repeated spin measurements on the recombined particles. A specific method would be the following: for a given t, and thus for a given $\delta\phi$ through (12), one of the particles is discarded and state tomography is performed on the second particle. This would give a density matrix ρ' of the form of Equation (9), where the cross terms would be some real numbers. Diagonalizing this matrix, an entanglement entropy I' can be computed. Assuming all noise has been accounted for, I' can be plotted against $\delta\phi(t)$ and compared with I as given by (11) and (12). Any deviation of I' from I could be a signal of a quantum gravity effect, and in particular the presence of plateaus in I' would be a signal of time discreteness.

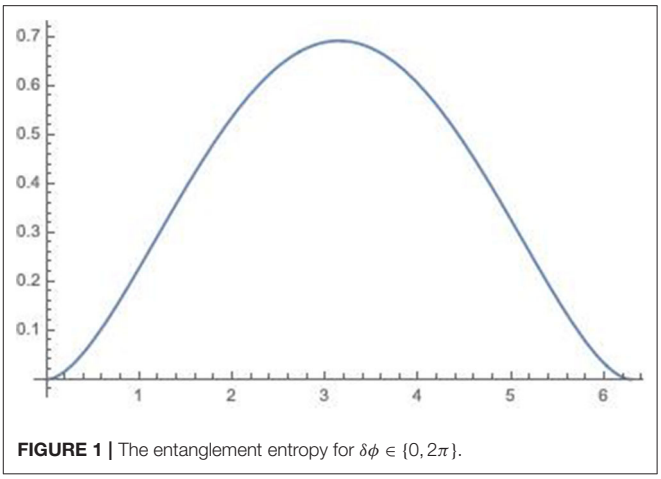

FIGURE 1 | The entanglement entropy for $\delta\phi \in \{0, 2\pi\}$.

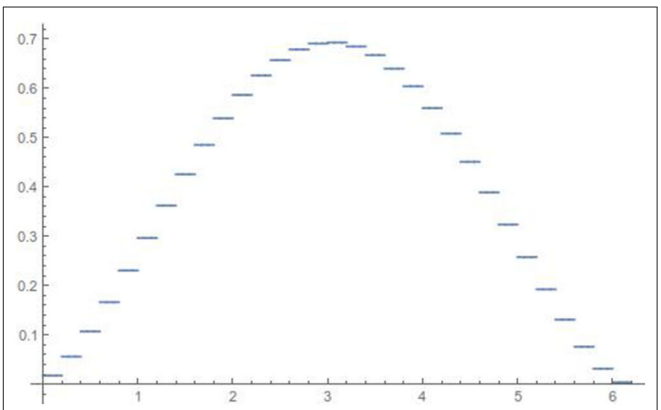

FIGURE 2 | The entanglement entropy for $\delta\phi \in \{0, 2\pi\}$ under the assumption that $\delta t/t_{Pl} \in \mathbb{N}^+$, for particles with mass one fifth of the Planck mass.

Consider now the hypothesis that time is discrete at the Planck scale. We consider here the simplest possible ansatz: that

$$\delta\tau = n\, t_P \tag{13}$$

with a non negative integer n. Writing $m = \alpha\, m_P$ with α a dimensionless positive real parameter, we have that the only values of ϕ that are actually realized are

$$\delta\phi = \alpha\, n, \tag{14}$$

that is, the phase ends up taking only discrete quantized values, when t is varied continuously. It follows that the entropy is not anymore given by a continuous curve as in **Figure 1**, but has characteristic quantum steps. As long as $\alpha \ll 1$ the steps are too fine to be resolved, but if α approaches unit the steps become visible, as in **Figure 2**, where $\alpha = .2$.

For particles with masses larger that the Planck mass interference is likely to disappear altogether, as is common in interference experiments when the wave frequency is much higher than the relevant scale of the apparatus. In this case wave theory goes to the Eikonal approximation. Wave mechanics goes to classical mechanics. The Compton frequency

$$\nu_c = \frac{mc^2}{\hbar} = \frac{m}{m_{Pl}} \frac{\nu_P}{2\pi} \tag{15}$$

of objects with mass larger than the Planck mass is formally larger than the Planck frequency $\nu_{Pl} = \frac{2\pi}{t_{Pl}}$ and probably meaningless.

Notice that in this case an apparatus capable of detecting $\delta\phi \sim 1$ is going to be affected by genuine dynamical effects since we can also write

$$\delta\phi = \frac{m^2}{m_{Pl}^2} \frac{ct}{d} \tag{16}$$

and if the left hand side and the first fraction are of order unit, so must be the second, with the consequence that the duration t of the interaction must be of the same order than the light travel time d/c between the particles. This would take us outside the static approximation used in the analysis (see also [22]).

3. DISCUSSION

The current hope is to realize the BMV experiment in the lab with masses $m \sim 10^{-6} m_P$ in the next few years [1]. Our key observation is that with masses at this scale, the BMV experiment is testing time differences of the order of $\delta\tau \sim 10^{-38} s \sim 10^{6} t_{Pl}$. This is already an extraordinarily small time. For comparison, the most accurate direct measurements of time at our disposal make use of the frequencies corresponding to energy differences in atomic states, atomic clocks, with an accuracy corresponding at best to a period of the order $\sim 10^{-19} s$ [15]. It is this extraordinary sensitivity to small time intervals in the BMV interference that makes this experiment so interesting.

A relativistic language is not needed to derive the correlations that the BMV experiment is expected to detect. In the non-relativistic language no small time intervals are in play: instead of $\delta\phi = mc^2\delta\tau/\hbar$, the phase reads $\delta\phi = t\,\delta E/\hbar$, and the c^2 makes all the difference, where the relevant time t at play is that of the laboratory frame.

But, if time discreteness is detected, the non-relativistic language becomes insufficient to describe the relevant physics. Time discreteness, according to current tentative theories, is a genuine relativistic effect arising from quantum gravity. On the other hand, this study does not go beyond the observation that the scales probed are possibly close to the relevant scale for such (naïve) models. Here we do not justify or discuss in depth a possible discretization of time and its implications. The effect discussed here however is a relativistic effect, and we do not think it could be relevant for systems in Newtonian gravity. In contrast for instance to Muller et al. [23], here the relativistic interpretation is used to discuss signatures of possible new physics associated with the discretization of time at the Planck-scale. These effects, if real, require the relativistic interpretation to be described. For completeness, although we are not aware of any such indication, let it be noted that as this is a yet unexplored time regime, there remains the logical possibility that there is time discreteness in this regime independently from quantum gravity.

As mentioned in the introduction, the analysis given here assumes the discreteness of $\delta\tau$ in Planck time multiples. It is possible, but it is not certain that this is implied by quantum gravity. Two reasons that could question this assumption are the following. First, the spectrum of τ could be less trivial and, as a consequence, *differences* of proper time could be much smaller. For instance, if the spacing between eigenvalues decreases when the eigenvalues are large, their differences may become small. Second, a more careful analysis might show that the interference depends on averages, or expectation values of time durations, and these may be continuous even if direct duration measurements are quantized. This issue in particular requires a more detailed quantum treatment of the phenomenon that will be developed elsewhere.

Even with these caveats, the possibility that quantum interference effects could depend on time differences of the order of Planck time, a scale so far considered totally out of reach, definitely deserves attention.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

We thank Pierre Martin-Dussaud and Andrea di Biagio for valuable discussions. The authors acknowledge support from the kind donors to the SM Center for Space, Time and the Quantum. This publication was made possible through the support of the ID 61466 grant from the John Templeton Foundation, as part of the The Quantum Information Structure of Spacetime (QISS) Project (qiss.fr). The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation.

REFERENCES

1. Bose S, Mazumdar A, Morley GW, Ulbricht H, Toroš M, Paternostro M, et al. Spin entanglement witness for quantum gravity. *Phys Rev Lett.* (2017) **119**:240401. doi: 10.1103/PhysRevLett.119.240401
2. Marletto C, Vedral V. Witness gravity's quantum side in the lab. *Nature.* (2017) **547**:156–8. doi: 10.1038/547156a
3. Marletto C, Vedral V. Gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity. *Phys Rev Lett.* (2017) **119**:240402. doi: 10.1103/PhysRevLett.119.240402
4. Anastopoulos C, Hu BL. Comment on "a spin entanglement witness for quantum gravity" and on "gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity". *arXiv [Preprint]*. arXiv:1804.11315.
5. Hall MJ, Reginatto M. On two recent proposals for witnessing nonclassical gravity. *J Phys A.* (2018) **51**:085303. doi: 10.1088/1751-8121/aaa734
6. Belenchia A, Wald R, Giacomini F, Castro-Ruiz E, Brukner Č, Aspelmeyer M. Quantum superposition of massive objects and the quantization of gravity. *ArXiv: 1807.07015*. (2018). doi: 10.1103/PhysRevD.98.126009
7. Giampaolo SM, Macrì T. Entanglement, holonomic constraints, and the quantization of fundamental interactions. *arXiv:1806.08383*. (2018). doi: 10.1038/s41598-019-47844-8
8. Marletto C, Vedral V. When can gravity path-entangle two spatially superposed masses?. *arXiv:1803.09124* (2018). doi: 10.1103/PhysRevD.98.046001
9. Albers M, Kiefer C, Reginatto M. Measurement analysis and quantum gravity. *Phys Rev D.* (2008) **78**:064051. doi: 10.1103/PhysRevD.78.064051
10. Christodoulou M, Rovelli C. On the possibility of laboratory evidence for quantum superposition of geometries. *Phys Lett B.* (2018) **792**:64–8. doi: 10.1016/j.physletb.2019.03.015
11. Reginatto M, Hall MJW. Entanglement of quantum fields via classical gravity. *arXiv:1809.04989* (2018).
12. Carney D, Stamp PCE, Taylor JM. Tabletop experiments for quantum gravity: a user's manual. *arXiv:1807.11494* (2018). doi: 10.1088/1361-6382/aaf9ca
13. Kosterina N, Wang R, Eriksson A, Gutierrez-Farewik EM. Force enhancement and force depression in a modified muscle model used for muscle activation prediction. *J Electromyogr Kinesiol.* (2013) **23**:759–65. doi: 10.1016/j.jelekin.2013.02.008
14. Balushi AA, Cong W, Mann RB. Optomechanical quantum Cavendish experiment. *Phys Rev A.* (2018) **98**:043811. doi: 10.1103/PhysRevA.98.043811
15. Marti GE, Hutson RB, Goban A, Campbell SL, Poli N, Ye J. Imaging Optical Frequencies with 100 μhz Precision and 1.1 μ m Resolution. *Phys Rev Lett.* (2018) **120**:103201. doi: 10.1103/PhysRevLett.120.103201
16. Garay LJ. Quantum gravity and minimum length. *Int J Mod Phys.* (1994) **A10**:145–66. doi: 10.1142/S0217751X95000085
17. Bojowald M. Loop quantum cosmology. IV: Discrete time evolution. *Class Quant Grav.* (2001) **18**:1071–88. doi: 10.1088/0264-9381/18/6/308
18. Ashtekar A, Singh P. Loop quantum cosmology: a status report. *Class Quant Grav.* (2011) **28**:213001. doi: 10.1088/0264-9381/28/21/213001
19. Rovelli C, Vidotto F. Compact phase space, cosmological constant, and discrete time. *Phys Rev D.* (2015) **91**:084037. doi: 10.1103/PhysRevD.91.084037
20. Gambini R, Pullin J. Canonical quantization of general relativity in discrete space-times. *Phys Rev Lett.* (2003) **90**:21301. doi: 10.1103/PhysRevLett.90.021301
21. Wan C, Scala M, Morley GW, Rahman ATMA, Ulbricht H, Bateman J, et al. Free nano-object ramsey interferometry for large quantum superpositions. *Phys Rev Lett.* (2016) **117**:143003. doi: 10.1103/PhysRevLett.117.143003
22. Mari A, De Palma G, Giovannetti V. Experiments testing macroscopic quantum superpositions must be slow. *Sci Rep.* (2016) **6**:22777. doi: 10.1038/srep22777
23. Muller H, Peters A, Chu S. A precision measurement of the gravitational redshift by the interference of matter waves. *Nature.* (2010) **463**:926–9. doi: 10.1038/nature08776

Conflict of Interest: 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.

frontiers in Physics

ORIGINAL RESEARCH
published: 26 October 2020
doi: 10.3389/fphy.2020.525333

Simulating Indefinite Causal Order With Rindler Observers

Aleksandra Dimić[1]*, *Marko Milivojević*[1], *Dragoljub Gočanin*[1], *Natália S. Móller*[2,3] and *Časlav Brukner*[4,5]

[1] Faculty of Physics, University of Belgrade, Belgrade, Serbia, [2] Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [3] Physics Department and Research Center Optimas, Technische Universität Kaiserslautern, Kaiserslautern, Germany, [4] Faculty of Physics, Vienna Center for Quantum Science and Technology (VCQ), University of Vienna, Vienna, Austria, [5] Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, Vienna, Austria

Realization of indefinite causal order (ICO), a theoretical possibility that even causal relations between physical events can be subjected to quantum superposition, apart from its general significance for the fundamental physics research, would also enable quantum information processing that outperforms protocols in which the underlying causal structure is definite. In this paper, we start with a proposition that an observer in a state of quantum superposition of being at two different relative distances from the event horizon of a black hole, effectively resides in ICO space-time generated by the black hole. By invoking the fact that the near-horizon geometry of a Schwarzschild black hole is that of a Rindler space-time, we propose a way to simulate an observer in ICO space-time by a Rindler observer in a state of superposition of having two different proper accelerations. By extension, a pair of Rindler observers with entangled proper accelerations simulates a pair of entangled ICO observers. Moreover, these Rindler-systems might have a plausible experimental realization by means of optomechanical resonators.

Keywords: Rindler observer, quantum switch, indefinite causal order, equivalence principle, causality

OPEN ACCESS

Edited by:
Rosa Poggiani,
University of Pisa, Italy

Reviewed by:
Jonathan S Ben-Benjamin,
Texas A&M University, United States
Paolo Grigolini,
University of North Texas, United States
Rainer Kaltenbaek,
University of Ljubljana, Slovenia

***Correspondence:**
Aleksandra Dimić
aleksandra.dimic@ff.bg.ac.rs

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: *08 January 2020*
Accepted: *18 September 2020*
Published: *26 October 2020*

Citation:
Dimić A, Milivojević M, Gočanin D, Móller NS and Brukner Č (2020) Simulating Indefinite Causal Order With Rindler Observers. Front. Phys. 8:525333. doi: 10.3389/fphy.2020.525333

1. INTRODUCTION

The principle of causality is an implicit assumption of every physical theory and it is universally supported by our experience of nature. From an operational point of view, causality can be understood as a system of signaling or communication relations between physical systems; an information flow whose properties are intimately related to the nature of space and time. One may even say that *the very essence of the classical structure of space and time is to impose a physical constraint on information processing.*

In the old Newtonian picture of the World, space and time are two generically different entities, universal for all observers. There is a single, three-dimensional flat Euclidean space and a single global time that enable us to unambiguously distinguish between past, present and future. Together, they constitute an absolute, independent background structure relative to which every physical event takes place. Signals can propagate in space with unlimited speed (action at a distance) and, consequently, each event can be caused by any other in its present or past. The special theory of relativity (SR) changed this paradigm: space and time became united into the $(3+1)$-dimensional *space-time continuum*—Minkowski space - in which signals cannot travel faster than the speed of light, enforcing them to stay within the local light cone. Nevertheless, the structure of Minkowski space adhered to the character of an independent, fixed background on which dynamical matter fields propagate.

The radical change came with Einstein's general theory of relativity (GR). The gravitational field came to be understood as the curved space-time itself, gravity being encoded in a metric tensor coupled to dynamical matter fields. There is no fixed, independent metric structure, no absolute background stage relative to which locations of physical events are to be defined, there are just dynamical fields, the metric being one of them, and physical events can only be located *relative to each other*. The possibility of communication between different observers, i.e. the causal order, is entirely determined by the dynamical configuration of light cones, and so, although dynamical, space-time, as a landscape of physical events, has definite causal order (DCO). Thinking about Quantum Mechanics (QM) of gravitating objects, the question arises whether there is a way to relax the restrictions of classical space-time structure and enable processes that do not obey definite causal relations, i.e., can there be a quantum superposition of different causal orders—an *indefinite causal order* (ICO)?

It is generally expected that unification of QM and gravitational physics will provide us with some deeper insights concerning the nature of space and time and their relationship with matter. However, the standard methods of quantization of matter fields employed in Quantum Field Theory (QFT) do not seem to work for Einstein's gravity; it holds a status of a non-renormalizable effective theory with undetermined high-energy degrees of freedom. In order to surpass the traditional concepts of GR and QFT, various ways of "quantizing" gravity were proposed so far, such as String Theory, Quantum Loop Gravity, Noncommutative Geometry, Supergravity, etc. However, to date, there has been no conclusive empirical evidence that would support or disprove any of the proposed "high-energy theories." This state of affairs motivates us to reconsider in which sense and to what extent can the seemingly contradictory principles of QM and GR be reconciled, while adhering to the tenets of both theories [1, 2].

There are two main incentives for this paper. The first came from the work of Oreshkov, Costa and Brukner [3], where it was found that it is possible to formulate quantum mechanics without any reference to a global causal structure, i.e. without predefined space-time. The resulting framework - the *process matrix formalism* - allows for processes incompatible with any definite causal order between operations performed on quantum systems. These abstract indefinite causal structures are shown to be advantageous for quantum computing [4, 5] and quantum communication [6–8]. One particular example that has an experimental demonstration is the so called "quantum switch" [4, 9–14], where the main idea is to use an auxiliary quantum system that can coherently control the order in which certain operations are applied. In the case of the so-called gravitational quantum switch (GQS) [15] the role of the control system is played by a gravitating object prepared in a state of quantum superposition of being at two different spatial locations. The second incentive comes from the intriguing idea of *quantum reference frames* (QRF) [16, 17] where one regards reference frames not as abstract systems of coordinates, but as actual physical objects subjected to the laws of quantum mechanics and describes the world from their perspective.

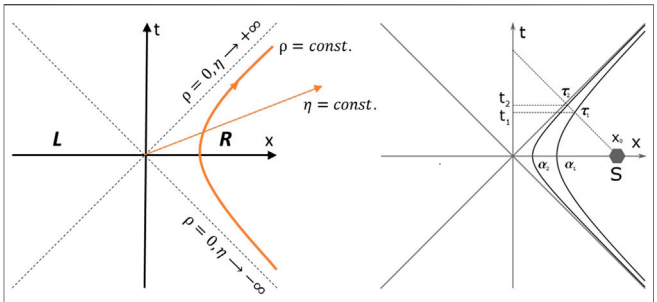

FIGURE 1 | *Rindler observers.* **(Left panel)**: *Coordinate lines in the right (R) Rindler wedge of the Minkowski space.* Coordinate lines of constant ρ are hyperbolas and they correspond to worldlines of Rindler observers. Regions L and R, called the left and the right Rindler wedge, respectively, are causally disconnected due to the presence of the Rindler horizon $t = \pm x$. **(Right panel)**: *Photon's worldline intersects the Rindlers.* A pair of Rindler observers in the R-wedge with different proper accelerations α_1 and α_2, $\alpha_1 < \alpha_2$. A photon sent from the source *S*, located at $t_S = 0$ and $x_S = x_0 > 0$, intersects the worldlines of the Rindler observers at their respective proper times τ_1 and τ_2. One can arrange things so that $\tau_1 = \tau_2$.

In this paper, we propose a way to simulate ICO processes by considering the fact that the near-horizon geometry of a Schwarzschild black hole (BH) is that of a Rindler space-time. Namely, a Rindler observer in a state of superposition of having two different proper accelerations corresponds to a near-horizon Schwarzschild observer in a state of superposition of being at two different locations along a single radial direction. From the viewpoint of such an observer the geometry of space-time is indefinite. This correspondence can be extended to a pair of Rindler observers with entangled proper accelerations simulating a pair of entangled ICO observers, as we illustrate by means of a simple example. Although they represent idealizations, these Rindler-systems could become a valuable resource for studying ICO processes in laboratory conditions, with plausible experimental realization in the form of opto-mechanical oscillators [18–21].

2. RINDLER OBSERVERS

In order to set the stage, consider the (1 + 1)-dimensional Minkowski space $\mathscr{M}_2$ and a central light cone defined by $t = \pm x$ (we set $c = 1$). In these globally inertial coordinates (t, x) the Minkowski metric is given by $ds^2_{\mathscr{M}_2} = -dt^2 + dx^2$. If we introduce *Rindler coordinates* (η, ρ) defined by $t = \rho\sinh(\eta)$ and $x = \rho\cosh(\eta)$, the metric becomes $ds^2_{\mathscr{M}_2} = -\rho^2 d\eta^2 + d\rho^2$. However, these new coordinates do not cover the whole Minkowski space, only the patch given by $x \geq 0$ and $|t| \leq x$. This region is called the *right Rindler wedge* or simply the R-wedge [see **Figure 1** (left panel)]. The family of coordinate lines of constant ρ are the branches of hyperbolas $x^2 - t^2 = \rho^2$ embedded in the R-wedge (the other set of branches belongs to the L-wedge, defined by $x \leq 0$ and $|t| \leq -x$) asymptotically approaching the Rindler horizon $t = \pm x$ ($\rho = 0$ and $\eta \to \pm\infty$). They correspond to the worldlines of physical systems that have constant proper acceleration of magnitude $\alpha(\rho) = 1/\rho$—*Rindler observers*. The proper time of a Rindler observer, with a given $\rho = const.$, is $d\tau = \rho d\eta$. Therefore, we can

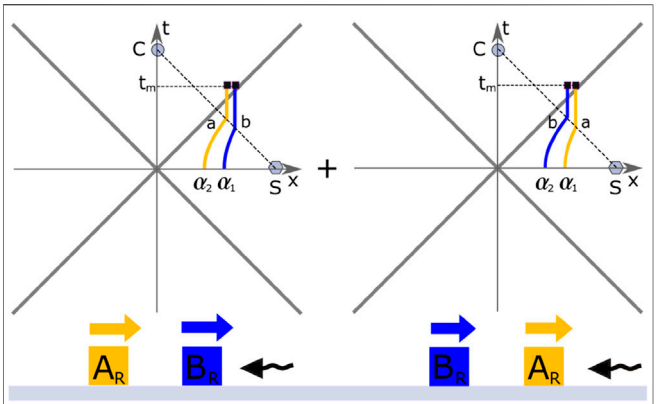

FIGURE 2 | *Entangled Rindler observers.* Worldline of a photon emitted by the source *S* intersects worldlines of the two Rindler observers, A_R and B_R, that have entangled proper accelerations. On the left panel, A_R has greater proper acceleration (α_2) than B_R (α_1), and on the right panel the accelerations are "switched". By choosing a suitable pair of values for the proper accelerations, these meetings (events *a* and *b*) occur at the same proper time τ^*. Rindler observers act on the polarization of the photon according to their "color", amber (*A*) or blue (*B*), that distinguishes them. Observer A_R performs a unitary transformation U_A at τ^* and observer B_R performs a unitary transformation U_B at τ^*. When Rindler observers meet the photon, they come to rest and remain that way until some particular moment t_m at which a projective measurement is performed in order to disentangle the state of the photon from that of the Rindler observers. The photon is observed at *C*.

define a worldline of a Rindler observer in the R-wedge with proper acceleration α by a pair of parametric equations:

$$t(\tau) = \frac{1}{\alpha}\sinh(\alpha\tau),\ x(\tau) = \frac{1}{\alpha}\cosh(\alpha\tau). \tag{1}$$

Rindler observer with greater proper acceleration has a more curved worldline (closer to the Rindler horizon). Note also that, due to the presence of the Rindler horizon, Rindler observers in the R-wedge are causally disconnected from the ones in the L-wedge, meaning that they are unable to communicate with each other.

Consider now a pair of Rindler observers in the R-wedge, with different proper accelerations α_1 and α_2. Let the worldline of the second one be more curved. That is, let $\alpha_1 < \alpha_2$. A photon sent to the left from the source *S*, located at $t_s = 0$ and $x_s = x_0 > 0$, intersects worldlines of the Rindler observers at proper times τ_1 and τ_2, respectively [see **Figure 1** (right panel)]. At $t = 0$ both observers are closer to the origin than *S*, implying that $\alpha_2 x_0 > \alpha_1 x_0 > 1$. This configuration has an interesting feature that will turn out to be important. Namely, given the values of x_0 and α_1, there exists *a unique* value of α_2, defined as the non trivial solution ($\alpha_2 \neq \alpha_1$) of the equation

$$\alpha_2 x_0 = (\alpha_1 x_0)^{\frac{\alpha_2}{\alpha_1}}, \tag{2}$$

for which $\tau_1 = \tau_2$ (for details, see **Supplementary Material**).

3. INDEFINITE CAUSAL ORDER VIA RINDLER OBSERVERS

Let us assume that we have a pair of Rindler observers in the R-wedge, Rindler-Amber (A_R) and Rindler-Blue (B_R). *Amber* and *Blue* are the colors by which we distinguish the two observers, see **Figure 2**. Note, however, that these "observers" need not be actual macroscopic measuring devices of any sort, nor sentient beings; they could be microscopic physical systems with some internal degrees of freedom (like spin). On the other hand, they have definite worldlines since they are confined within accelerating laboratories (imagine well-enough localized classical "boxes" each carrying an atom). We assume that these internal degrees of freedom are such that they do not get affected by the accelerated motion of the Rindler laboratory, which we also assume to be completely isolated. Source *S* emits a photon whose worldline intersects the worldlines of A_R and B_R. The photon starts in some polarization state $|\Psi\rangle$ and the Rindler observers can perform instantaneous unitary transformations on it.

When their worldlines intersect, A_R performs a unitary transformation U_A on the photon's polarization state. This constitutes event *a*. In general, U_A is a function of A_R's proper time. We can abstractly think of A_R as a Rindler "clock" whose worldline and the ticking rate are defined by the proper acceleration α_A. The state of A_R will therefore be denoted by $|\tau_{\alpha_A}; A\rangle$, without getting into details of what A_R's actual degrees of freedom are and what kind of Hamiltonian governs the dynamics thereof. And the same protocol applies to B_R. Its meeting with the photon and application of a unitary transformation U_B constitutes event *b*. By choosing a suitable values of the proper accelerations α_A and α_B, we can arrange that meetings of the Rindler observers with the photon (events *a* and *b*) occur *at the same proper times* $\tau_a = \tau_b = \tau^*$ (see the discussion at the end of Section 2).

Here we want to stress that physical events are not regarded as pure geometrical points that constitute space-time manifold (modulo diffeomorphisms) with some definite set of causal relations defined by the metric. Rather they are defined operationally, through application of a specific unitary transformation, or more generally a specific completely positive trace-preserving (CPTP) map. Taking quantum mechanics into account, we consider the possibility that the same physical event can be in a superposition of occurring at different space-time locations. This would enable the realization of indefinite causal order between pairs of events, such as *a* and *b* in the above discussion. One example of this situation is the already mentioned quantum switch [9–14].

On the left panel of **Figure 2**, the proper acceleration of A_R (α_2) is greater than the proper acceleration of B_R (α_1) and on the right panel, the values are interchanged, A_R has the smaller proper acceleration (α_1) and B_R has the greater proper acceleration (α_2). In the reference frame of the inertial observer sitting at $x = 0$ the initial state (at $t = 0$) of the whole system (Rindlers ⊗ photon) in the former case is the separable state $|\tau_{\alpha_2}(0), A\rangle|\tau_{\alpha_1}(0), B\rangle|\Psi\rangle$; the photon first meets B_R and then A_R. In the latter case, when the initial state of the system is $|\tau_{\alpha_1}(0), A\rangle|\tau_{\alpha_2}(0), B\rangle|\Psi\rangle$, the photon first meets A_R and then B_R. If the Rindler observers are prepared in the *entangled state* that is a superposition of the two previous ones, at $t = 0$ we have

$$\frac{1}{\sqrt{2}}\left(|\tau_{\alpha_2}(0), A\rangle|\tau_{\alpha_1}(0), B\rangle + |\tau_{\alpha_1}(0), A\rangle|\tau_{\alpha_2}(0), B\rangle\right)|\Psi\rangle. \tag{3}$$

It is important to realize that the event a has to be *one and the same* in both "branches" of the superposition. The meeting of the photon with A_R in the case when A_R has greater proper acceleration than B_R and the meeting of the photon with A_R when A_R has smaller proper acceleration than B_R, have to be locally indistinguishable events in every respect. That is why we demand that the event occurs at the same proper time, τ^*, in both situations. In principle, the state of the photon could be affected by the kinematic state of a Rinldler laboratory, which might give rise to entanglement between proper accelerations of the Rindler laboratories and photon'|'s polarization state. To avoid this possibility, we put them to rest just before they meet the photon, thus making them inertial from that point on. Conditional deacceleration of the laboratories along the Rindler trajectories can be performed sufficiently fast, yet gradually, not to produce the Unruh radiation.

For $t < t_1$ (where t_1 is the time coordinate of the intersection of the photon's worldline with the less curved Rindler worldline) the state is

$$\frac{1}{\sqrt{2}}\left(\left|\tau_{\alpha_2}(t), A\right\rangle\left|\tau_{\alpha_1}(t), B\right\rangle + \left|\tau_{\alpha_1}(t), A\right\rangle\left|\tau_{\alpha_2}(t), B\right\rangle\right)|\Psi\rangle. \tag{4}$$

By the time the photon went through the laboratories, unitary transformations, $U_A(\tau^*)$ and $U_B(\tau^*)$, have been applied on it. At some instant $t > t_2$ (where t_2 is the time coordinate of the intersection of the photon'|'s worldline with the more curved Rindler worldline) the state of the whole system is given by

$$\frac{1}{\sqrt{2}}\left(\left|\tau^* + t - t_2, A\right\rangle\left|\tau^* + t - t_1, B\right\rangle U_A(\tau^*)U_B(\tau^*) + \left|\tau^* + t - t_1, A\right\rangle\left|\tau^* + t - t_2, B\right\rangle U_B(\tau^*)U_A(\tau^*)\right)|\Psi\rangle, \tag{5}$$

where $t - t_1$ and $t - t_2$ are the time intervals during which the respective Rindler laboratories were at rest.

Finally, we need to disentangle the state of the photon from the state of the Rindler observers. To this end, at some moment t_m, a projective measurement (postselection on the internal state of the Rindlers) is performed in the superposition basis $\{|m_i\rangle, |m_i^{\perp}\rangle \,|\, i = 1, 2\}$, separately for each laboratory. The basis states are given by

$$|m_i\rangle = \frac{1}{\sqrt{2}}\left(\left|\tau^* + t_m - t_i, A\right\rangle + \left|\tau^* + t_m - t_i, B\right\rangle\right),$$
$$|m_i^{\perp}\rangle = \frac{1}{\sqrt{2}}\left(\left|\tau^* + t_m - t_i, A\right\rangle - \left|\tau^* + t_m - t_i, B\right\rangle\right). \tag{6}$$

Postselection on any pair of possible measurement results leads to the final state of the photon

$$\frac{1}{\sqrt{2}}\left(U_A(\tau^*)U_B(\tau^*) \pm U_B(\tau^*)U_A(\tau^*)\right)\left|\Psi\right\rangle \tag{7}$$

Subsequently, the photon may be observed at C. However, being in the state **7**, there is no way to distinguish, given the photon alone, which of the two events (a and b) lies in the causal future of the other, and the information about the causal order is lost.

4. GRAVITATIONAL SCENARIO

Imagine now that we have a system that involves a Schwarzschild BH and an observer (outside the horizon) in a state of superposition of being at two different relative proper distances from the horizon. The observer is well-enough localized and has a negligible effect on the gravitational field. Also, we do not assume the existence of a fixed background geometry with reference to which we could define positions; only the *relative* distance between the BH (its horizon) and the observer has physical meaning.

In Schwarzchild coordinates (t, r, θ, ϕ), the metric of the BH exterior is

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_2^2, \tag{8}$$

with $f(r) = 1 - \frac{R_S}{r}$ and $\Omega_2^2 = d\theta^2 + \sin^2\theta d\phi^2$. We are only interested in a single radial direction, so we can ignore the angular part of the metric. The proper radial distance between the stationary observer at r_{lab} and the event horizon at $R_S = 2MG$ (M being the mass of the BH) is

$$\rho = \int_{R_s}^{r_{lab}} \frac{dr}{\sqrt{f(r)}}. \tag{9}$$

Therefore, our bipartite system can be interpreted as a situation where we have an observer in the state $\frac{1}{\sqrt{2}}\left(|\rho_1\rangle + |\rho_2\rangle\right)$ with indefinite proper distance from the horizon in definite Schwarzchild geometry. On the other hand, from the viewpoint of the observer (quantum reference frame), the gravitational field appears to be indefinite, as if the BH is in the state of superposition of being at two different places relative to the observer.

The idea that a gravitating object in a state of quantum superposition of being at two different locations somehow "induces" a quantum superposition of different geometries, dates back to Feynman [22] and it has been successfully promoted recently [15, 23–26]. Although this seems as a natural way to combine GR and the linearity of QM, it remains unclear in which sense can a gravitational field (i.e., space-time geometry) be in a state of quantum superposition, see for example [27]. Here, we propose a way of looking at this situation based on the relational character of quantum superposition [28].

Einstein's equivalence principle states that a gravitational field is *locally equivalent* to an accelerating reference frame in flat space-time. As a consequence, for every well-enough localized stationary (r = const.) observer in Schwarzschild geometry there is an equivalent uniformly accelerating observer in Minkowski space. Moreover, the near-horizon geometry of a Schwarzschild BH is that of a Rindler space-time, and therefore a Schwarzschild observer whose proper distance from the horizon is ρ corresponds to the Rindler observer whose proper acceleration is $1/\rho$ (see **Supplementary Material** for details).

By extending this reasoning, we propose a "quantum" version of Einstein's equivalence principle by stating that ICO space-time is locally equivalent to a non-inertial reference frame with

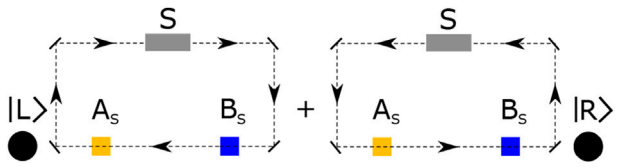

FIGURE 3 | *Gravitational quantum switch*. The system involves a photon source S, two observers, A_S and B_S, and a BH in a state of superposition of being at two different positions relative to them, $|L\rangle$ or $|R\rangle$. The direction in which the photon is emitted depends on the position of the BH and therefore the BH plays the role of a quantum control for the whole process. As a result, the photon is in a superposition of traveling in two opposite directions.

superposed proper accelerations. In particular, we can relate a Schwarzschild observer in the state $\frac{1}{\sqrt{2}}(|\rho_1\rangle+|\rho_2\rangle)$ of being at two different relative distances from the horizon, to the Rindler observer in the state of superposition of having proper accelerations $1/\rho_1$ and $1/\rho_2$, respectively. The similar principle has been invoked and derived within quantum reference frames formalism [16] in the Newtonian limit.

Consider now a pair of observers in the near-horizon region of a Schwarzschild BH, Schwarzschild-Amber (A_S) and Schwarzschild-Blue (B_S). Let the observers have entangled proper distances from the horizon (along a single radial ray) and fixed relative distance between each other. From the reference frame of the BH, space-time has a definite geometry and the state of this tripartite system is $\frac{1}{\sqrt{2}}|0\rangle_{BH}(|\rho_1\rangle_A|\rho_2\rangle_B + |\rho_2\rangle_A|\rho_1\rangle_B)$, where $|0\rangle_{BH}$ is the position state of the black hole. Observers A_S and B_S correspond to a pair of Rindler observers, A_R and B_R, both in the R-wedge, with *entangled proper accelerations*. On the other hand, from the point of view of the pair of observers, we have a BH in a state of superposition of being "at two different sides" of the observers, symmetrically. These states of the BH are denoted by $|L\rangle$ and $|R\rangle$, see **Figure 3**. The middle point between A_S and B_S is well defined in relative terms. From the reference frame associated to this point [29] the joint state of the system is $\frac{1}{\sqrt{2}}(|L\rangle+|R\rangle)|-\vec{\rho}\rangle_A|\vec{\rho}\rangle_B$, where $\vec{\rho}=\frac{\rho_2-\rho_1}{2}$ is the half-distance between the two observers.

We can perform a photon experiment, similar to the one described in Section 3, that involves the observers A_S and B_S and a source S that emits a photon in the direction that depends on the position of the BH relative to the observers. The position of the BH plays the role of a quantum control for the whole process (gravitational quantum switch). Due to the gravitational time dilation, we can arrange things so that the photon passes through both laboratories at the same moment of their local proper time (see **Supplementary Material** for details). This is analogous to the case of the Rindler quantum switch from Section 2. When the photon gets inside the laboratory, a unitary transformation, U_A or U_B, depending on the laboratory, is applied instantaneously on its polarization state. The meeting of the photon and the laboratory A_S and the application of the unitary U_A is the *event a*, and likewise, the meeting of the photon and the laboratory B_S and the application of the unitary U_B is the *event b*. After performing a projective measurement in the superposition basis $1/\sqrt{2}(|L\rangle \pm |R\rangle)$ of the BH (to disentangle its state from the photon'|'s state, as described in [15]), the final state of the photon implies the two events do not possess definite causal order.

5. CONCLUSION

In this paper, we proposed a way to characterize ICO space-time as a space-time associated with the reference frame of a quantum observer - quantum observer perceives ICO space-time. As an illustration, we considered a bipartite system that involves a Schwarzschild BH and an observer outside the horizon, in a state of quantum superposition of being at two different relative distances. By invoking the fact that near-horizon geometry is that of a Rindler space-time, we can relate this ICO observer to the Rindler observer in a state of superposition of having two different proper accelerations. By extension, a pair of ICO observers with entangled proper distances from the BH horizon corresponds to a pair of Rindler observers with entangled proper accelerations. As an example, we analyzed Rindler quantum switch and the related gravitational quantum switch.

Furthermore, a Bell's inequality for temporal order of events was found in [15]. The same kind of inequality can be derived by using two pairs of Rindler observers, one in the left and the other in the right Rindler wedge. In the corresponding gravitational scenario we would have to take into account the Kruskal extension of the Schwarzschild solution. In this case, the gravitational quantum switch would involve two pairs of observers residing in conformally flat space-times connected by Einstein-Rosen bridge. We postpone this interesting analysis for future work.

On a more practical side, there is a growing effort in demonstrating quantum features of nano-to-mesoscale optomechanical systems. This may provide a challenging, yet feasible experimental realizations for the proposed Rindler systems [18]. Recently, mesoscopic mechanical resonators were considered as quantum non-inertial reference frames [19, 20] and entanglement of two massive mechanical oscillators is achieved [21]. It has been proposed to utilize quantum optical fields in order to prepare and measure the quantum states of mechanical resonators, conceivably opening the possibility to quantum-mechanically control the acceleration of such quantum non-inertial reference frames [18].

In an actual experiment, potential decoherance effects can compromise the predicted result [30]. Moreover, QFT effects could also be taken into account. In this context, our Rindler observers could be viewed as Unruh-DeWitt detectors, where an increase of the thermal noise, due to the Unruh effect, may affect the evolution of the system, such that it can no longer be considered as a coherent superposition, but rather a (convex) classical mixture. However, since we can choose proper accelerations of the Rindler observers to be arbitrarily small by putting the photon source at suitable position, the Unruh effect can always be made negligible. Just to put some numbers, if we set $x_0 = 1$ m, the accelerations would be of order 10^{17} m/s^2, which corresponds to the Unruh temperature of order 10^{-4} K, and this is far too small for the Unruh effect to be detectable. Correspondingly, for a solar mass black hole with $R_s = 3$ km, we have a Schwarzschild observer at 1m proper distance from the horizon, which is a good near-horizon

approximation. Furthermore, depending on the parameters of the objects involved, e.g., masses of the laboratories, these Rindler systems could be used to test hypotheses such as the Ghirardi-Rimini-Weber (GRW) model of objective collapse [31]. Namely, a failure to maintain a coherent macroscopic superposition even after screening off the system from decoherence effects, might be taken as an indication of a spontaneous GRW-type collapses. However, as we noted in the text, the systems involved in our setting need not be macroscopic systems (they can be microscopic ones, such as atoms). Moreover, the issues concerning macroscopic systems may still be avoided by performing the experiment "sufficiently fast" (before the alleged GRW-type of collapse should take place). This kind of assessment was provided in [15] for gravitational quantum switch. Finally, we should also mention that there is a relativistic version of the GRW model [32]. Our relativistic Rindler systems could perhaps be used for studying and testing such models.

Finally, and perhaps most importantly, we would like to establish a more rigorous framework that would allow us to formally define an ICO space-time related to a general quantum reference frame. This could be an important step towards a better understanding of the quantum nature of space-time.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article.

AUTHOR CONTRIBUTIONS

AD, MM, DG, and ČB conceived the research. AD, MM, and DG provided main theoretical idea and equally contributed to writing the paper. NM participated in discussion concerning entangled laboratories. ČB supervised the project.

ACKNOWLEDGMENTS

This is a significantly revised version of the paper that can already be found on the arXiv [33]. The authors thank Dejan Simić, Marko Vojinović, Nikola Paunković, and Ämin Baumeler for helpful comments. AD, MM, and DG acknowledge support from the bilateral project SRB 02/2018, *Causality in Quantum Mechanics and Quantum Gravity*, between Austria and Serbia. AD and MM acknowledge support from the project No. ON171035 and D.G. from the project No. ON171031 of Serbian Ministry of Education and Science. Additionally, AD acknowledges support from scholarship awarded from The Austrian Agency for International Cooperation in Education and Research (OeAD-GmbH). AD and DG acknowledge grant FQXi-MGA-1806 that supported their stay in Vienna. AD and DG would also like to thank University of Vienna and IQOQI for hospitality during their stay. NM thanks Marcelo Terra Cunha, Nelson Yokomizo, Raphael Drumond, Leonardo Neves, and Davi Barros for useful discussions. NM acknowledge Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for financial support and VI Paraty Quantum Information School for the introduction into the topic. ČB acknowledges the support from the research platform TURIS, from the European Commission *via* Testing the Large-Scale Limit of Quantum Mechanics (TEQ) (No. 766900) project, and from the Austrian-Serbian bilateral scientific cooperation no. 451-03-02141/2017-09/02, and by the Austrian Science Fund (FWF) through the SFB project BeyondC and a grant from the Foundational Questions Institute (FQXi) Fund. This publication was made possible through the support of the ID 61466 grant from the John Templeton Foundation, as part of the The Quantum Information Structure of Space-time (QISS) Project (qiss.fr). The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2020.525333/full#supplementary-material

REFERENCES

1. Hardy L. Probability theories with dynamic causal structure: a new framework for quantum gravity. arXiv preprint gr-qc/0509120 (2005)
2. Hardy L. Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure. *J Phys A: Math Theor* (2007) **40**:3081–9. doi:10.1088/1751-8113/40/12/s12
3. Oreshkov O, Costa F, Brukner Č. Quantum correlations with no causal order. *Nat Commun.* (2012) **3**:1092. doi:10.1038/ncomms2076
4. Chiribella G, D'Ariano GM, Perinotti P, Valiron B. Quantum computations without definite causal structure. *Phys Rev A* (2013) **88**:022318. doi:10.1103/physreva.88.022318
5. Araújo M, Costa F, Brukner Č. Computational advantage from quantum-controlled ordering of gates. *Phys Rev Lett* (2014) **113**:250402. doi:10.1103/physrevlett.113.250402
6. Feix A, Araújo M, Brukner Č. Quantum superposition of the order of parties as a communication resource. *Phys Rev A* (2015) **92**:052326. doi:10.1103/physreva.92.052326
7. Guérin PA, Feix A, Araújo M, Brukner Č. Exponential communication complexity advantage from quantum superposition of the direction of communication. *Phys Rev Lett* (2016) **117**:100502. doi:10.1103/physrevlett.117.100502
8. Ebler D, Salek S, Chiribella G. Enhanced communication with the assistance of indefinite causal order. *Phys Rev Lett* (2018) **120**:1205020. doi:10.1103/physrevlett.120.120502
9. Chiribella G, Giovannetti V, Maccone L, Perinotti P. Teleportation transfers only speakable quantum information. *Phys Rev A* (2012) **86**:040301. doi:10.1103/physreva.86.010304
10. Friis N, Dunjko V, Dür W, Briegel HJ. Implementing quantum control for unknown subroutines. *Phys Rev A* (2014) **89**:030303. doi:10.1103/physreva.89.030303
11. Procopio LM, Moqanaki A, Araújo M, Costa F, Alonso Calafell I, Dowd EG, et al. Experimental superposition of orders of quantum gates. *Nat Commun* (2015) **6**:7913. doi:10.1038/ncomms8913

12. Rambo TM, Altepeter JB, Kumar P, D'Ariano GM. Functional quantum computing: an optical approach. *Phys Rev A* (2016) **93**:052321. doi:10.1103/physreva.93.052321
13. Rubino G, Rozema LA, Feix A, Araújo M, Zeuner J. M, Procopio LM, et al. Experimental verification of an indefinite causal order. *Sci Adv* (2017) **3**: e1602589. doi:10.1126/sciadv.1602589
14. Goswami K, Giarmatzi C, Kewming M, Costa F, Branciard C, Romero J, et al. Indefinite causal order in a quantum switch. *Phys Rev Let* (2018) **121**:090503. doi:10.1103/physrevlett.121.090503
15. Zych M, Costa F, Pikovski I, Brukner Č. Bell's theorem for temporal order. *Nat Commun* (2019) **10**:3772. doi:10.1038/s41467-019-11579-x
16. Giacomini F, Castro-Ruiz E, Brukner Č. Quantum mechanics and the covariance of physical laws in quantum reference frames. *Nat Commun* (2019) **10**:494. doi:10.1038/s41467-018-08155-0
17. Guérin PA, Brukner Č. Observer-dependent locality of quantum events. *New J Phys* (2018) **20**:103031. doi:10.1088/1367-2630/aae742
18. Kaltenbaek R, Aspelmeyer M, Barker PF, Bassi A, Bateman J, Bongs K, et al. Precise and robust optical beam steering for space optical instrumentation. *EPJ Quant Technol* (2015) **3**:5. doi:10.1007/s12567-019-00271-x
19. Katz BN, Blencowe MP, Schwab KC. Mesoscopic mechanical resonators as quantum noninertial reference frames. *Phys Rev A* (2015) **92**:042104. doi:10.1103/physreva.92.042104
20. Abdi M, Degenfeld-Schonburg P, Sameti M, Navarrete-Benlloch C, Hartmann MJ. Dissipative optomechanical preparation of macroscopic quantum superposition states. *Phys Rev Lett* (2016) **116**:233604. doi:10.1103/physrevlett.116.233604
21. Ockeloen-Korppi CF, Damskägg E, Pirkkalainen J-M, Asjad M, Clerk AA, Massel F, et al. Stabilized entanglement of massive mechanical oscillators. *Nature* (2018) **556**:478. doi:10.1038/s41586-018-0038-x
22. Feynman R. Chapel Hill conference proceedings, 1957. Available at: http://www.edition-open-sources.org/sources/5/3/.
23. Anastopoulos C, Hu BL. Probing a gravitational cat state. *Class Quant Grav* (2015) **32**:165022. doi:10.1088/0264-9381/32/16/165022
24. Bose S, Mazumdar A, Morley GW, Ulbricht H, Toroš M, Paternostro M, et al. Spin entanglement witness for quantum gravity. *Phys Rev Lett.* (2017) **119**: 240401. doi:10.1103/physrevlett.119.240401
25. Marletto C, Vedral V. Gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity. *Phys Rev Lett* (2017) **119**:240402. doi:10.1103/PhysRevLett.119.240402
26. Christodoulou M, Rovelli C. On the possibility of laboratory evidence for quantum superposition of geometries. *Phys Lett B* (2019) **792**:64–68. doi:10.1016/j.physletb.2019.03.015
27. Penrose R. On the gravitization of quantum mechanics 1: quantum state reduction. *Found Phys* (2014) **44**:557–575. doi:10.1007/s10701-013-9770-0
28. Zych M, Costa F, & Ralph TC. Relativity of quantum superpositions. arXiv preprint arXiv:1809.04999 (2018)
29. Vanrietvelde A, Hoehn PA, Giacomini F, Castro-Ruiz E. A change of perspective: switching quantum reference frames via a perspective-neutral framework. *Quant* (2020) **4**:225. doi:10.22331/q-2020-01-27-225
30. Joos E, Zeh HD, Kiefer C, Giulini DJ, Kupsch J, Stamatescu IO. *Decoherence and the appearance of a classical world in quantum theory*. Berlin, Germany: Springer Science Business Media (2013)
31. Ghirardi GC, Rimini A, Weber T. Unified dynamics for microscopic and macroscopic systems. *Phys Rev D.* (1986) **34**(2):470. doi:10.1103/physrevd.34.470
32. Tumulka R. A relativistic version of the Ghirardi–Rimini–Weber model. *J. Stat. Phys* (2006) **125**:821–840.
33. Dimić A, Milivojević M, Gočanin D, Brukner Č. Simulating spacetime with indefinite causal order *via* Rindler observers. arXiv preprint arXiv:1712.02689 (2017).

Conflict of Interest: 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.

frontiers in Physics

ORIGINAL RESEARCH
published: 19 November 2020
doi: 10.3389/fphy.2020.00262

Complexity and Newton's Laws

*Leonard Susskind**

Stanford Institute for Theoretical Physics and Department of Physics, Stanford University, Stanford, CA, United States

In a recent note [1], I argued that the holographic origin of ordinary gravitational attraction is the quantum mechanical tendency for operators to grow under time evolution. In a follow-up [2] the claim was tested in the context of the SYK theory and its bulk dual—the theory of near-extremal black holes. In this paper I give an improved version of the size-momentum correspondence of [2], and show that Newton's laws of motion are a consequence. Operator size is closely related to complexity. Therefore, one may say that gravitational attraction is a manifestation of the tendency for complexity to increase. The improved version of the size-momentum correspondence can be justified by the arguments of Lin et al. [3] constructing symmetry generators for the approximate symmetries of the SYK model.

Keywords: gravitation, quantum complexity, holographic principle, SYK model, AdS spacetime

1. PRELIMINARY REMARKS

What is it that takes place in the holographic representation of a theory when an object in the bulk is gravitationally attracted to a massive body? Consider a holographic theory representing a region of empty space. By operating with a simple boundary operator ψ, a particle can be introduced into the bulk. As the particle moves away from the boundary the operator ψ evolves with time,

$$\psi(t) = e^{-iHt}\psi e^{iHt}, \tag{1.1}$$

and becomes increasingly complex. If expanded in simple boundary operators the average number of such operators will increase and one says the size of the operator grows [4]. A closely related fact is that the complexity of $\psi(t)$ grows. We might expect that the complexity is a good holographic indicator of how far from the boundary the particle is located. However, there is more to the particle than just its location; we may want to know how its momentum or velocity is encoded in the evolving operator $\psi(t)$. The size or complexity is not enough to determine both its distance from the boundary and its momentum.

Let's say that the particle is moving away from the boundary so that the size is increasing. It seems plausible that velocity is related to the rate of change of size. This is oversimplified but it roughly captures the idea that size, and its rate of change, holographically encode the motion of the particle.

Now suppose there is a heavy mass at the center of the bulk region. The gravitational pull of the heavy mass will accelerate the particle away from the boundary. We may expect that the growth of ψ—both its size and complexity—will be accelerated relative to the empty case. Thus, it is plausible that the holographic representation of gravitational attraction has something to do with the tendency for operators to grow and become more complex [1]. Gravity accelerates that tendency.

In [2], the SYK model and its bulk dual, which in many ways resembles the theory of near-extremal Reissner-Nordstrom (NERN) black holes, provided a testing ground for this hypothesis.

OPEN ACCESS

Edited by:
Jakub Mielczarek,
Jagiellonian University, Poland

Reviewed by:
Tadashi Takayanagi,
Kyoto University, Japan
Jan De Boer,
University of Amsterdam, Netherlands

***Correspondence:**
Leonard Susskind
sonnysusskind@gmail.com

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: *29 August 2019*
Accepted: *11 June 2020*
Published: *19 November 2020*

Citation:
Susskind L (2020) Complexity and Newton's Laws. Front. Phys. 8:262. doi: 10.3389/fphy.2020.00262

In this paper I will continue the line of reasoning of [2]. A connection between the evolution of complexity and Newton's second and third laws of motion, as well as Newton's law of attraction, will be derived:

- Newton's second law is summarized by the familiar equation,

$$F = ma \tag{1.2}$$

or its generalization,

$$F = \frac{dP}{dt}. \tag{1.3}$$

- Newton's third law—the law of action and reaction—says that the force exerted by A on B is equal and opposite to the force exerted by B on A.
- Newton's law of attraction,

$$F = \frac{mMG}{r^2} \tag{1.4}$$

My arguments are a heuristic mix of quantum information and gravitation and involve some guesswork, but a more formal basis has been found by Lin et al. [3]. In section 7, I'll briefly explain the connection insofar as I understand it.

Note on Size and Complexity

The size of an operator is roughly the average number of elementary constituent degrees of freedom that appear in the expansion of the operator. While in general it is a somewhat imprecisely defined notion, in the SYK context in which we will be working it can be very precisely defined in terms of the average number of fermions in the operator [4]. Some care is needed in order to define size at finite temperature. The concept of temperature-dependent size that I will use in this paper is due to Qi and Streicher [5]. Size and complexity are logically different concepts but for reasons that will become clear, over the time period relevant for this paper the two are essentially indistinguishable. In order to minimize notation, and to avoid confusing size with entropy, I will use the symbol $\mathcal{C}$ to represent both. The quantitative equivalence of size and complexity continues for times of order the scrambling time, but by then the connection between size and the motion of an infalling particle breaks down as the particle reaches the stretched horizon.

If the considerations of this paper are to be useful, it will be necessary to generalize the concept of size to more general cases, in particular to higher dimensional gauge-gravity dualities. At the present time I do not know of any precise definition of size in strongly coupled CFTs. This is a serious hole in our knowledge that I hope will be filled.

Numerical Coefficients

Many of the equations in this paper are correct up to numerical factors relating SYK quantities to NERN quantities. These factors are in-principle computable using numerical SYK techniques, and depend on the locality parameter q. I will use the symbol $\approx$ to indicate that an equation is correct up to such numerical factors.

2. NEAR-EXTREMAL BLACK HOLES

The bulk dual of the SYK model is usually taken to be a version of the $(1 + 1)$-dimensional Jackiw-Teitelboim dilaton-gravity system. But that description (of a system with no local degrees of freedom) does not do justice to the spectrum of excitations of the SYK system. In many ways SYK is similar to the long throat of a near-extremal charged black hole whose geometry is approximately $AdS_2 \times S^2$. Unlike pure JT gravity SYK contains matter that can propagate in the throat as it would in the NERN geometry, and the properties of quantum-complexity are not well described by the simple dilaton-gravity system [6]. For these reasons I prefer the language of NERN black holes although no exact SYK/NERN correspondence is known.

To keep the paper self-contained, in this section I will review near-extremal black holes, and then in section 3, the dictionary relating SYK and near-extremal black holes will be explained. I will closely follow the discussion of NERN black holes in [2].

The metric of the $(3 + 1)$-dimensional Reissner-Nordstrom black hole is,

$$\begin{aligned} ds^2 &= -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega^2 \\ f(r) &= \left(1 - \frac{r_+}{r}\right)\left(1 - \frac{r_-}{r}\right). \end{aligned} \tag{2.1}$$

The inner $(-)$ and outer $(+)$ horizons are located at,

$$r_\pm \equiv GM \pm \sqrt{G^2M^2 - GQ^2}.$$

Define

$$(r_+ - r_-) = \delta r. \tag{2.2}$$

The temperature is given by,

$$T = \frac{1}{\beta} = \frac{1}{4\pi}\left(\frac{r_+ - r_-}{r_+^2}\right). \tag{2.3}$$

or

$$\boxed{T = \frac{1}{\beta} = \frac{\delta r}{4\pi r_+^2}} \tag{2.4}$$

The extremal limit is defined by $Q^2 = GM^2$ at which point the horizon radii are equal, $r_+ = r_-$. Our interest will be in near-extremal Reissner-Nordstrom (NERN) black holes, for which

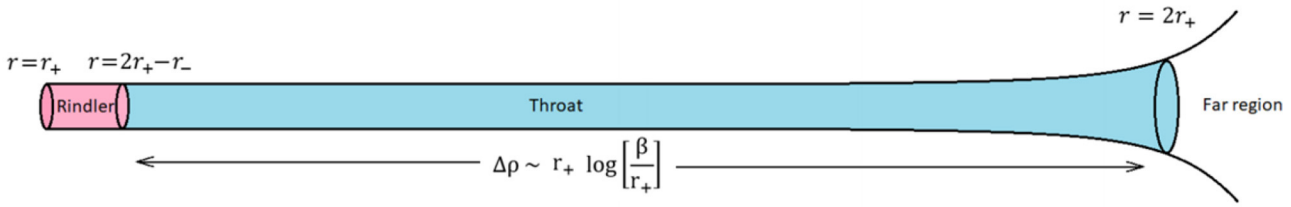

FIGURE 1 | The three regions outside a near-extremal charged black hole. Unlike for uncharged black holes, there is now a "throat" separating the Rindler and far regions.

$$\delta r << r_+.$$

In the NERN limit the temperature is small ($\beta \gg r_+$) and the near-horizon region develops a "throat" whose length is much longer than r_+. The throat is an almost-homogeneous cylinder-like region in which the gravitational field is uniform over a long distance.

2.1. The Geometry of the Throat

The exterior geometry consists of three regions shown in **Figure 1**.

- The *Rindler region* is closest to the horizon where the geometry closely resembles the Schwarzschild black hole with the same entropy. It is defined by,

$$r_+ < r \lesssim 2r_+ - r_- \tag{2.5}$$

  The Rindler region has proper length $\sim r_+$ which means that it's about as long as it is wide.

  The gravitational field (i.e., the proper acceleration $\alpha = \partial_r\sqrt{f(r)}$ required to remain static at fixed r) grows rapidly near the horizon. While the quantity $(1 - \frac{r_+}{r})$ varies significantly in the Rindler region, $(1 - \frac{r_-}{r})$ is essentially constant.
- Proceeding outward, the next region is the *throat* defined by

$$2r_+ - r_- \lesssim r \lesssim 2r_+ \tag{2.6}$$

  The throat is long and of almost constant width. The geometry in the throat region is approximately $AdS_2 \times S^2$, and the gravitational field is almost constant. The throat ends at $r = 2r_+$, which we will soon see is the location of a potential barrier which separates the throat from the far region. The throat is a feature of charged black holes and is absent from the Schwarzschild black hole.

  For most purposes the geometry in the throat can be approximated by the extremal geometry with $r_+ = r_-$.

  The proper length of the throat is,

$$\Delta\rho \int_{2r_+-r_-}^{2r_+} \frac{dr}{\sqrt{f}}$$

  giving

$$\boxed{\Delta\rho = r_+ \log\left(\frac{2\pi\beta}{r_+}\right).} \tag{2.7}$$

  We will assume that $\log\left(\frac{2\pi\beta}{r_+}\right) >> 1$ which means that the throat is much longer than it is wide.
- Next is the *far region* where

$$(1 - \frac{r_-}{r}) \sim (1 - \frac{r_+}{r}) \sim 1.$$

  The far region lies beyond $r = 2r_+$. The far region will not be of much interest to us. We will cut it off and replace it by a boundary condition at $r = 2r_+$.

2.2. The Black Hole Boundary

The black hole is effectively sealed off from the far region by a potential barrier. Low energy quanta in the throat are reflected back as they try to cross from the throat to the far region, or from the far region to the throat. The barrier height for a NERN black hole is much higher than the temperature and provides a natural boundary of the black hole region. It may be thought of as the holographic boundary in a quantum description. It is also the so-called Schwarzian boundary that appears in current literature on SYK theory [7–9]. The boundary will play an important role in this paper.

The S-wave potential barrier has the form

$$V(r) = \frac{\partial_r(f^2)}{4r}$$

and for a NERN black hole it is given by,

$$\boxed{V(r) = \frac{r_+(r - r_+)^3}{r^6}.} \tag{2.8}$$

The width of the barrier in proper distance units is of order r_+ and for near extremal RN it is much narrower than the length of

the throat. It therefore forms a fairly sharp boundary separating the black hole from the rest of space.

At the top of the barrier the potential is,

$$\boxed{V_{top} = \left(\frac{1}{8r_+}\right)^2 \approx \mathcal{J}^2} \tag{2.9}$$

where $\mathcal{J}$ is the scale of energy in the SYK theory (see section 3). The units of V are energy-squared rather than energy. For a particle to get over the barrier (without tunneling) its energy must be at least $\sqrt{V_{top}}$. This is much higher than the thermal scale and for that reason the barrier is very effective at decoupling the black hole, including its thermal atmosphere, from the far region. Another relevant point is that a particle that starts at rest at the top of the potential has energy of order $\frac{1}{r_+} \approx \mathcal{J}$.

The top of the potential barrier serves as an effective boundary of the black hole. It occurs at,

$$r = 2r_+ \tag{2.10}$$

We may eliminate reference to the entire region beyond the boundary and replace it by a suitable boundary condition[1] on the time-like surface at which $r = 2r_+$. This is accomplished by the introduction of a boundary term in the gravitational action.

We define a radial proper-length coordinate ρ measured from the black hole boundary[2],

$$\rho = \int_r^{r_b} \frac{dr'}{\sqrt{f(r')}} \tag{2.11}$$

In the throat r and ρ are related by,

$$\frac{r - r_+}{r_+} = e^{-\rho/r_+} \tag{2.12}$$

At the boundary $\rho = 0$, and at the beginning of the Rindler region $\rho = r_+ \log(\beta/r_+)$. Note that ρ has a large variation over the throat region which makes it a more suitable radial coordinate than r which hardly varies at all.

The black hole boundary, defined as the place where $r = 2r_+$, is not a rigid immovable object. Fluctuations or dynamical back reaction can change the metric so that the distance from the horizon to the boundary varies. This can be taken into account by allowing the boundary to move from its equilibrium position at $\rho = 0$.

[1]In the SYK literature the corresponding boundary condition is placed on the point where the dilaton achieves a certain value. In the correspondence between the dilaton theory and the NERN black hole the dilaton is simply the area of the local 2-sphere at a given radial location.

[2]Frequently a radial proper coordinate is defined as the distance to the horizon. Note that in this paper ρ measures distance to the black hole boundary at $r = 2r_+$, not to the horizon.

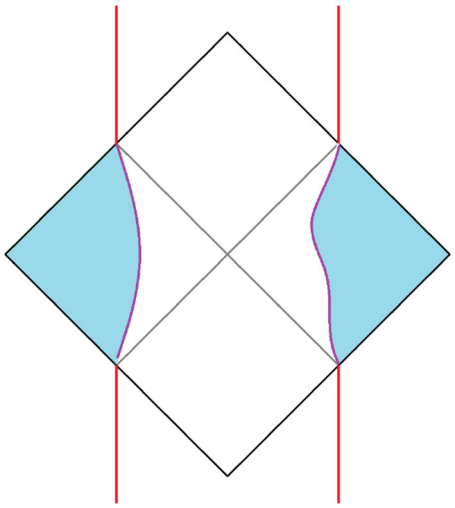

FIGURE 2 | Penrose diagram for a NERN black hole. The curved lines represent the trajectory of the black hole boundary at $r = 2r_+$. On the left side the boundary is shown in its equilibrium location while on the right it is moving in reaction to some matter.

In **Figure 2**, a Penrose diagram for a two-sided NERN black hole is shown along with the trajectories of the boundary and the regions beyond the boundary. The left-side boundary is shown in its static equilibrium position but on the right side the dynamical nature of the boundary is illustrated.

The equation of motion of the boundary is generated by the Hawking-Gibbons-York boundary term (Schwarzian action in SYK literature) needed to supplement the Einstein-Maxwell action in the presence of a boundary. For small slow perturbations the boundary motion is non-relativistic with a large mass of order S/r_+ (S is the black hole entropy). The mass of the boundary is of order the mass of the black hole itself[3]. Using the SYK-NERN dictionary in section 3 we see that the boundary mass is,

$$\boxed{M_B \approx \mathcal{J}N.} \tag{2.13}$$

2.3. Particle Motion in the Throat

Consider a particle dropped at $t = 0$ from $\rho = 0$, i.e., from the top of the potential as in **Figure 3**. The energy of the particle is $\sim 1/r_+$, which corresponds to an energy $\mathcal{J}$ in the SYK theory [2].

Under the influence of a uniform gravitational field it accelerates[4] toward the horizon. **Supplementary Material** works out the equation of motion for the particle, and one finds that the force is constant throughout the throat. The momentum increases linearly with time.

So far a small but important effect has been ignored. There is a back reaction that occurs when the particle falls off the potential.

[3]The idea of the boundary as a very massive particle was suggested by Kitaev, who developed this idea in [10]. It was further developed in [11].

[4] In the sense that its momentum grows. Being relativistic the velocity is close to 1.

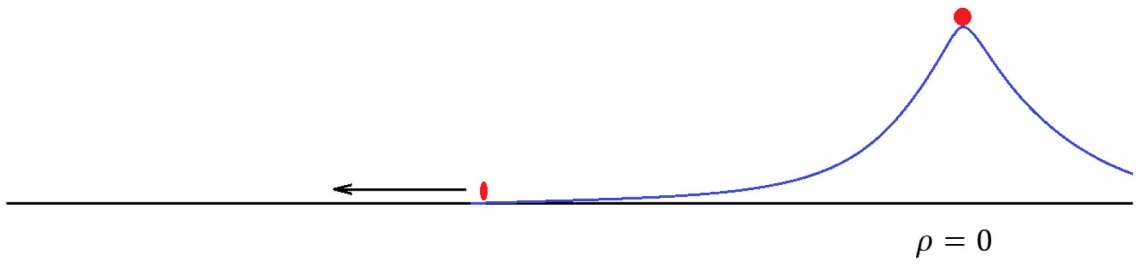

FIGURE 3 | A particle is introduced at the top of the potential, and subsequently rolls down the potential.

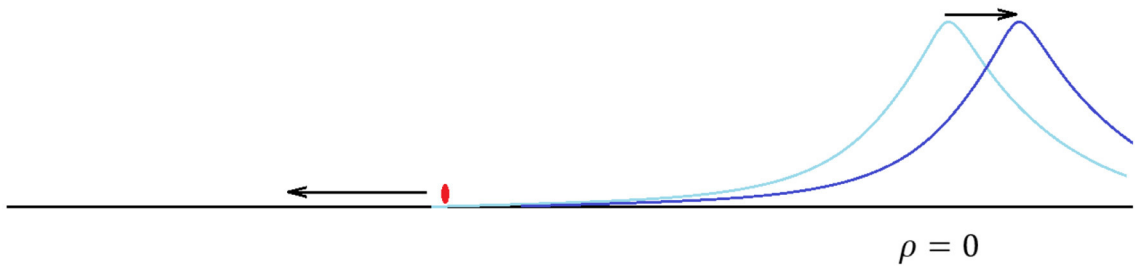

FIGURE 4 | The boundary recoils when the particle is accelerated. At all times the particle and the boundary have equal and opposite momentum.

The potential exerts a force on the particle, which in turn exerts an equal and opposite force on the boundary. The result is that the boundary recoils with a small velocity. [With some effort this can be seen in the Schwarzian analysis [9]]. This recoil, illustrated in **Figure 4**, will be important later.

Once the particle falls off the potential it quickly becomes relativistic. In the throat region its trajectory is given by

$$\begin{aligned} dt &= \frac{1}{\sqrt{f}} d\rho \\ &= \frac{r_+}{(r - r_+)} d\rho \\ &= e^{\rho/r_+} d\rho \end{aligned} \tag{2.14}$$

Thus, the particle trajectory satisfies,

$$t = r_+(e^{\rho/r_+} - 1) \tag{2.15}$$

or

$$\boxed{\rho(t) = r_+ \log\left(\frac{t - r_+}{r_+}\right)} \tag{2.16}$$

The total time to fall from $\rho = 0$ to the beginning of the Rindler region is β. During that time the distance traveled is

$$\boxed{\Delta\rho = r_+ \log\left(\frac{2\pi\beta}{r_+}\right).} \tag{2.17}$$

2.4. Schwarzschild r in Terms of ρ

Let's consider the relation between the Schwarzschild coordinate r and the proper coordinate ρ. To a very good approximation, in the throat we can assume $r_+ = r_-$ and that r is constant. The emblackening factor

$$\left(\frac{r - r_+}{r}\right)\left(\frac{r - r_-}{r}\right)$$

may be replaced by its extremal value

$$f(r) \approx \left(\frac{r - r_+}{r_+}\right)^2 \tag{2.18}$$

Recall that ρ is the proper distance measured from the boundary at $r = 2r_+$,

$$\begin{aligned} d\rho &= \frac{dr}{\sqrt{f(r)}} \\ &= r_+ \frac{dr}{r - r_+} \\ \rho &= r_+ \log\left(\frac{r_+}{r - r_+}\right) \end{aligned} \tag{2.19}$$

or,

$$\boxed{\frac{r - r_+}{r_+} = e^{-\rho/r_+}} \tag{2.20}$$

2.5. Surface Gravity and $\tilde{\beta}$

The so-called surface gravity κ will play an important role in what follows. At the horizon the surface gravity is related to the temperature of the black hole by,

$$T = \frac{1}{2\pi}\,\kappa_{horizon}. \tag{2.21}$$

More generally it is defined at any radial position r by

$$\tilde{\kappa}(r) = \frac{1}{2}\frac{df}{dr} = \frac{r_+(r-r_-)+r_-(r-r_+)}{2r^3} \tag{2.22}$$

which in the throat is approximated by,

$$\tilde{\kappa}(r) = \frac{r-r_+}{r_+^2} \tag{2.23}$$

The purpose of the tilde notation is to indicate a local quantity, i.e., one that may vary throughout the throat. Corresponding variables without the tilde indicate the value of the quantity at the horizon. We may also define $\tilde{T}$ and $\tilde{\beta}$ by,

$$\begin{aligned}\tilde{T} &= \frac{1}{2\pi}\tilde{\kappa} = \frac{1}{2\pi}\frac{r-r_+}{r_+^2}\\ \tilde{\beta} &= \frac{1}{\tilde{T}} = 2\pi\frac{r_+^2}{r-r_+}\end{aligned} \tag{2.24}$$

(Except at the horizon the quantity $\tilde{T}$ is not a real temperature. It is a useful quantity defined by 2.23 and 2.24 whose importance will become clear.)

In the throat let's express $\tilde{\beta}$ in terms of ρ. Using 2.16, 2.20, and 2.24,

$$\tilde{T}(\rho) = \frac{1}{2\pi r_+}e^{-\rho/r_+}$$

and,

$$\boxed{\tilde{\beta}(\rho) = 2\pi\ r_+ e^{\rho/r_+}.} \tag{2.25}$$

At $\rho = 0$, $\tilde{\beta}$ is given by

$$\tilde{\beta} = 2\pi\ r_+ \approx \mathcal{J}^{-1} \qquad (\rho = 0) \tag{2.26}$$

At the Rindler end of the throat where $\rho = r_+\log(\beta/r_+)$, $\tilde{\beta}$ is given by

$$\tilde{\beta} = \beta \qquad (\rho = \beta) \tag{2.27}$$

By following the trajectory of the infalling particle 2.15, and using 2.25 we find that $\tilde{\beta}$ grows according to,

$$\boxed{\tilde{\beta}(t) = 2\pi(t + r_+)} \tag{2.28}$$

As the Rindler region is approached $\tilde{\beta}$ stops increasing and remains at β until the horizon is reached.

3. SYK/NERN DICTIONARY

We can only go so far in understanding the quantum mechanics of NERN black holes without having a concrete holographic system to analyze. That brings us to the well-studied SYK model. In this section the SYK/NERN dictionary is spelled out.

3.1. Qualitative Considerations

We'll begin with qualitative aspects of the SYK/NERN dictionary and then attempt to determine more precise numerical coefficients in the next subsection. The two-sided arrows in this subsection indicate qualitative correspondences.

- The overall energy scale of the SYK model is called $\mathcal{J}$. Its inverse $\frac{1}{\mathcal{J}}$ is a length scale which corresponds to the Schwarzschild radius r_+. In the SYK model acting with a fermion operator ψ adds an energy $\approx \mathcal{J}$. On the NERN side dropping a particle from the top of the barrier adds energy $\approx 1/r_+$. Thus it makes sense to identify the process of dropping a particle from the black hole boundary, with acting with a single fermion operator.

$$\boxed{1/r_+ \approx \mathcal{J}} \tag{3.1}$$

- A single boundary fermion operator in SYK has size 1 corresponding the assumption of [2] that the initial size of the operator that creates the particle at the top of the barrier is also 1.

$$\boxed{\text{size of 1 fermion} \leftrightarrow \text{size of initial particle.}} \tag{3.2}$$

- Up to a numerical factor $\approx$ 1, the zero temperature extremal entropy of SYK is the number of fermion degrees of freedom N.

$$\boxed{S_0 \approx N} \tag{3.3}$$

- The 4-dimensional Newton constant can be obtained from the entropy formula,

$$S_0 = \pi r_+^2/G$$

Using 3.1 and 3.3 gives,

$$\boxed{G \approx \frac{1}{\mathcal{J}^2 N}} \tag{3.4}$$

- The SYK theory does not have sub-AdS locality (locality on scales smaller than r_+). It is comparable to a string theory in which the string scale is of order r_+ or $1/\mathcal{J}$.
- The black hole mass is r_+/G. This translates to,

$$\boxed{M \approx N\mathcal{J}.} \tag{3.5}$$

- Many of the detailed coefficients that appear in the subsequent formulas are dependent on q, the SYK-locality parameter that determines the number of fermion operators in each term in the Hamiltonian. For the most part I will treat q as a constant of order unity and not try to track the q-dependent details.

The literature on the bulk dual of SYK theory [8, 9, 11] has its own conventions and notations which are not the standard ones used for NERN black holes. Here I'll add to the dictionary the translation between the two.

- The dynamical boundary of SYK (described by the Schwarzian action) corresponds to the NERN black hole boundary, i.e., the top of the barrier where the throat meets the far region. The action governing the motion of the boundary is the Gibbons-Hawking-York boundary action.

$$\boxed{\text{GHY} \leftrightarrow \text{Schwarzian}} \tag{3.6}$$

- The dilaton field ϕ in [8, 9, 11] is related to the area of the transverse geometry at a given radial position,

$$\phi = 4\pi r^2. \tag{3.7}$$

- The time coordinate used in the SYK literature is called u. It is the proper time measured at the boundary. We may identify it with the proper time at the top of the potential barrier at $r = 2r_+$.

  The time coordinate t used in this paper is the asymptotic Schwarzschild time coordinate for the NERN black hole. The relation between u and t is,

$$f(r)|_{2r_+}\, dt^2 = du^2. \tag{3.8}$$

  For NERN black holes $f(r)|_{2r_+} = 1/4$, from which it follows that,

$$t = 2u. \tag{3.9}$$

3.2. Quantitative Considerations

In some cases the numerical coefficients appearing in the various correspondences have been studied and allow more quantitative correspondences. I'll give some examples here, but I won't keep track of these coefficients in subsequent sections.

The specific heats of the SYK model and the NERN black hole can both be computed. On the NERN side the calculation is analytic and yields,

$$c = \frac{dM}{dT} = \frac{4\pi^2}{G} r_+^2 T \tag{3.10}$$

For SYK the calculation was done in [7]. The result is,

$$c = 4\pi^2 \alpha_s(q) \frac{N}{\mathcal{J}} T \tag{3.11}$$

where $\alpha_s(q)$ is a numerically computed function of the SYK locality parameter q. For $q = 4$ $\alpha_s = 0.007$ and for large q it decreases $\sim 1/q^2$.

Setting 3.10 and 3.11 equal, we find the relation,

$$\boxed{\alpha_s \frac{N}{\mathcal{J}} = \frac{r_+^3}{G}.} \tag{3.12}$$

Let λ and p be dimensionless coefficients defined by,

$$G = \frac{\lambda}{\mathcal{J}^2 N} \tag{3.13}$$

and

$$r_+ = \frac{p}{\mathcal{J}}. \tag{3.14}$$

Plugging 3.13 and 3.14 into 3.12 gives one relation between p and λ,

$$\boxed{\alpha_s = \frac{p^3}{\lambda}.} \tag{3.15}$$

Another relation can be found by considering the entropy of SYK and the NERN black hole. On the NERN side we use the Bekenstein-Hawking formula which gives,

$$S = \frac{\pi r_+^2}{G}. \tag{3.16}$$

On the SYK side reference [7] Stanford and Maldacena computed the near extremal entropy:

$$S = d(q)N. \tag{3.17}$$

where $d(q)$ is another numerically computed function of q which varies from $d(4) = .23$ to $d(\infty) = 0.35$.

Combing 3.16 and 3.17 with 3.13 and 3.14 gives another equation for p and λ,

$$\frac{\pi p^2}{\lambda} = d. \tag{3.18}$$

The two relations 3.15 and 3.18 yield the following expressions for λ and p,

$$\begin{aligned} \lambda &= \frac{\pi^3 \alpha_s^2}{d^3} \\ p &= \frac{\pi \alpha_s}{d} \end{aligned} \tag{3.19}$$

Thus, we find the following correspondences,

$$r_+ = \left(\frac{\pi \alpha_s}{d}\right) \frac{1}{\mathcal{J}} \tag{3.20}$$

$$G = \left(\frac{\pi^3 \alpha_s^2}{d^3}\right) \frac{1}{N\mathcal{J}^2}. \tag{3.21}$$

For $q = 4$ the numerical values of α_s and d are,

$$\begin{aligned} \alpha_s &= 0.007 \\ d &= 0.23 \end{aligned} \tag{3.22}$$

giving,

$$\boxed{r_+ = \frac{.10}{\mathcal{J}}} \tag{3.23}$$

and

$$\boxed{G = \frac{.12}{N\mathcal{J}^2}} \tag{3.24}$$

Now let's return to the problem of a light particle dropped from the top of the potential 2.9 and estimate its energy ϵ. The height of the barrier is

$$\sqrt{V_{top}} = 1/8r_+ \approx \mathcal{J}.$$

We may compare this energy with the energy added to the SYK ground state by applying a single fermion operator ψ (In other words it is the energy associated with a size 1 perturbation). This energy is expected to be of order $\mathcal{J}$ and to have some smooth q dependence. It is given by,

$$\epsilon(q)\mathcal{J} = \left\langle \frac{1}{Z(\beta)} \mathrm{Tr} H(2\psi e^{-\beta H}\psi - e^{-\beta H}) \right\rangle \tag{3.25}$$

where the average $\langle \dots \rangle$ indicates disorder average. (The factor of 2 in the first term is present because of the SYK convention that $\psi^2 = 1/2$).

4. GROWTH OF SIZE

Consider applying a single fermion operator at time $t = 0$. The operator evolves in time according to,

$$\psi(t) = e^{-iHt}\psi e^{iHt}. \tag{4.1}$$

and becomes a superposition of many-fermion operators [4, 5]. The average number of Fermions at time t is the size. The evolution is described by Feynman-like diagrams which, up to the scrambling time, grow exponentially [4, 5]. At each stage the average number of fermions increases by common factor. The process resembles an exponentially expanding tree as shown in **Figure 5**.

It is similar to the evolution of a quantum circuit and it is natural to define a circuit depth. In general the circuit depth may not unfold uniformly with time. For example, if for some reason the computer runs at a variable time-dependent rate, the size will grow exponentially with depth but not necessarily with time. The time associated with a unit change in circuit depth is defined to be Δt and it may be time-dependent. This type of time-dependence occurs in the evolution of size at low temperature [5].

We can express this in terms of a rate of growth R,

$$\begin{aligned} R(t) &\equiv \frac{d \log \mathcal{C}(t)}{dt} \\ &= \frac{1}{\Delta t} \end{aligned} \tag{4.2}$$

The exponential growth as a function of circuit depth (for time less than the scrambling time) is the reason that size and complexity are proportional to each other. One may think of the size at a given depth as the number of "leaves" of the tree, and the complexity as the integrated number of vertices up to that point. Because the tree grows exponentially, the number of leaves and the number of vertices are proportional, and with some normalization (of complexity) the size and complexity can be set equal.

4.1. Infinite Temperature

Roberts et al. [4] have calculated the time dependence of size at infinite temperature and find,

$$\mathcal{C}(t) \sim e^{2\mathcal{J}t}. \tag{4.3}$$

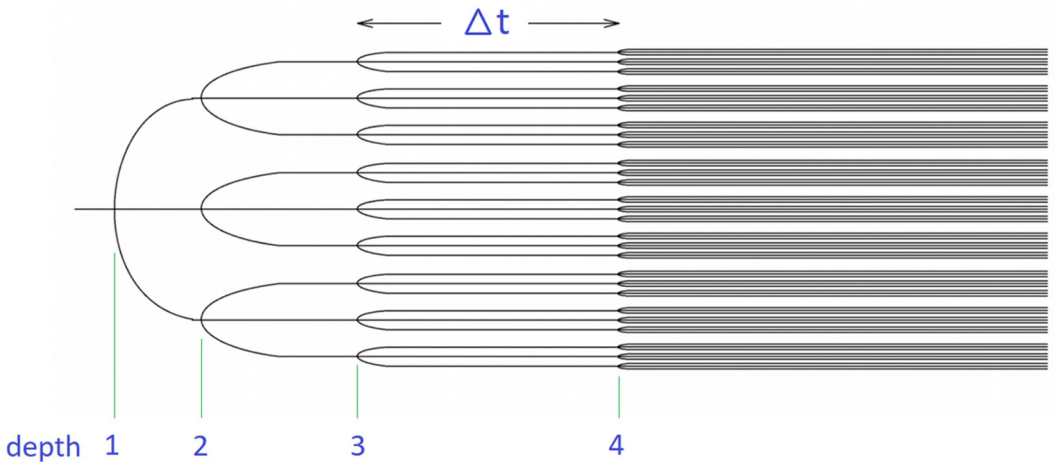

FIGURE 5 | Tree-like operator growth. The size at any circuit-depth is the final number of fermions while the complexity is the number of vertices in the diagram. In this figure the size is 81 and the complexity is 40. The complexity at the next step would be 40 + 81 = 121. The time scale for a unit change in depth is Δt. In general Δt may itself be time dependent.

Roberts, Stanford, and Streicher give a more detailed formula,

$$\mathcal{C}(t) = 1 + 2\sinh^2(\mathcal{J}t) \tag{4.4}$$

Apart from a brief transient the size grows exponentially. Dropping the 1 which is unimportant, the rate $R(t)$ is

$$\begin{aligned} R(t) &= \frac{1}{\mathcal{C}(t)}\frac{d\mathcal{C}(t)}{dt} \\ &= \mathcal{J}\frac{\cosh \mathcal{J}t}{\sinh \mathcal{J}t} \end{aligned} \tag{4.5}$$

which after a short time $\mathcal{J}^{-1}$ tends to

$$R \to \mathcal{J} \tag{4.6}$$

We may restate this in terms of Δt,

$$\Delta t \approx \mathcal{J}^{-1} \qquad (T = \infty). \tag{4.7}$$

4.2. Low Temperature, T << J

At very low temperatures the pattern is quantitatively different. According to Qi and Streicher the size for low T is given by,

$$\mathcal{C}(t) = 1 + 2\frac{\mathcal{J}^2\beta^2}{\pi^2}\sinh^2\left(\frac{\pi t}{\beta}\right) \tag{4.8}$$

Early on the rate is comparable to the infinite T case,

$$\boxed{\frac{1}{\mathcal{C}}\frac{d\mathcal{C}}{dt} \approx \mathcal{J} \qquad (\mathcal{J}t \sim 1)} \tag{4.9}$$

but after a time $\beta/2\pi$ (at which the infalling particle has reached the Rindler region) the rate has slowed to

$$\boxed{\frac{1}{\mathcal{C}}\frac{d\mathcal{C}}{dt} = \frac{2\pi}{\beta} \qquad (\beta/2\pi < t < t_*)} \tag{4.10}$$

Our interest will lie in the throat region during time period between $t = 0$ and $t = 2\pi\beta$, where the rate is time-dependent, varying from $\approx \mathcal{J}$ to $2\pi/\beta$. In fact the rate is not so much time-dependent as it is position dependent. To understand the rate in more detail [5] we consider a particle falling from the black hole boundary. The particle falls along a trajectory $\rho(t)$. The time dependence of the growth rate is really ρ-dependence: the rate depends on t only through the position ρ.

Let $\kappa(\rho)$ be the surface gravity at position ρ,

$$\kappa(\rho) \equiv \frac{1}{2}\partial_r f(r) \tag{4.11}$$

and let $\tilde{\beta}$ be,

$$\tilde{\beta}(\rho) = 2\pi/\kappa(\rho). \tag{4.12}$$

At the horizon the surface gravity is related to the temperature of the black hole,

$$T = \frac{1}{2\pi}\,\kappa_{horizon}. \tag{4.13}$$

and $\tilde{\beta}_{horizon}$ to the inverse temperature,

$$\tilde{\beta}_{horizon} = \beta \tag{4.14}$$

The obvious guess for the interpolation between 4.9 and 4.10 is,

$$\frac{1}{\mathcal{C}}\frac{d\mathcal{C}}{dt} = \frac{2\pi}{\tilde{\beta}}.$$

This is correct in the Rindler region but in the throat it is off by a factor of 2. Consistency between the Qi-Streicher formula and 2.28 requires,

$$\boxed{\frac{1}{\mathcal{C}}\frac{d\mathcal{C}}{dt} \sim \frac{4\pi}{\tilde{\beta}}} \tag{4.15}$$

or in terms of Δt,

$$\Delta t = \frac{\tilde{\beta}(\rho)}{4\pi}. \tag{4.16}$$

5. MOMENTUM-SIZE CORRESPONDENCE

5.1. Formulation

In [2], it was proposed that the holographic dual to the momentum of an infalling particle is related to the size (or complexity) of the operator that created the particle. By itself this is not dimensionally consistent. One needs a quantity with units of length to multiply the momentum in order to get a dimensionless size. For a Schwarzschild black hole there is only one length scale, the Schwarzschild radius, which is proportional to $\beta/2\pi$. Thus,

$$\mathcal{C} \approx \frac{\beta}{2\pi}P, \tag{5.1}$$

(the factor of proportionality being q-dependent). However, in the NERN case this cannot be the right relation. Pick a point ρ_0 a fixed distance from the boundary. If the temperature is sufficiently low the geometry between $\rho = 0$ and $\rho = \rho_0$ is extremely insensitive to β and the growth up to that point should also be insensitive to β. But equation 5.1 implies that $\mathcal{C}(\rho_0)$ blows up as $T \to 0$.

The formula used in [2] was originally suggested by Ying Zhao. It is obtained by replacing equation 5.1 by a local version,

$$\boxed{\frac{\tilde{\beta}}{4\pi}P \approx \mathcal{C}.} \tag{5.2}$$

From 5.2 one sees that complexity (or size) is not in one to one relationship with either position (ρ) or momentum (P) but it is a combination of both variables. For fixed position the complexity is proportional to momentum, but for fixed momentum the complexity increases the deeper the particle is into the throat. I will not repeat the argument here but just remark that in [2] it was shown that 5.2 gives an accurate account of the evolution of size, reproducing a non-trivial result of [12]. As we'll now see, it is also agrees with the calculations of [5].

5.2. Qi-Streicher Formula

Qi and Streicher [5] have made a first-principles calculation of the growth of a single fermion operator ψ at finite temperature $1/\beta$ in the SYK theory. As time evolves the complexity of $\psi(t)$ grows until the scrambling time t_*. Between $t = 0$ and $t = t_*$ Qi and Streicher find[5],

$$\boxed{\mathcal{C}(t) = 1 + 2\frac{\mathcal{J}^2\beta^2}{\pi^2}\sinh^2\left(\frac{\pi t}{\beta}\right)} \tag{5.3}$$

Let us compare 2.28,

$$\tilde{\beta}(t) = 2\pi(t + r_+)$$

with the SYK calculation of Qi-Streicher. We first note from 4.15 that for $t > r_+$,

$$\frac{2\pi}{\tilde{\beta}} \sim \frac{d\log\mathcal{C}(t)}{dt}. \tag{5.4}$$

The first term in the Qi-Streicher formula 5.3 is unimportant. We may write,

$$\mathcal{C}(t) = 2\frac{\mathcal{J}^2\beta^2}{\pi^2}\sinh^2\pi t/\beta.$$

and

$$\frac{d\log\mathcal{C}}{dt} = \frac{2\pi}{\beta}\tanh^{-1}\pi t/\beta. \tag{5.5}$$

Using 5.4 we find,

$$\boxed{\tilde{\beta} \sim \beta\tanh(\pi t/\beta)} \tag{5.6}$$

For $r_+ < t < \beta/2\pi$ this gives $\tilde{\beta} \sim 2\pi t$ in agreement with 2.28.

Actually this is accurate for almost the entire passage through the throat. The ratio

$$\frac{\beta\tanh(\pi t/\beta)}{\pi t}$$

is close to 1 as long as $\pi t/\beta < 1$. (Note $\left(\frac{\tanh .3}{.3}\right) = 0.97$) In terms of ρ this means until,

[5]Qi and Streicher calculate the size but for reasons I have explained size and complexity are interchangeable for our purposes.

$$\rho = r_+ \log \beta/r_+ - r_+ \log(\pi)$$
$$= \Delta\rho - r_+ \log(\pi) \tag{5.7}$$

where $\Delta\rho$ is the length of the throat (see **Figure 1**). In other words there is very good agreement between the Qi-Streicher formula, and the rate 4.15 conjectured in [2], over the entire throat, right up to the start of the Rindler region. The agreement continues to be qualitatively good into the Rindler region. The discrepancy by the time the particle has reached a Planck distance from the horizon is less than a factor of 2.

There is a striking similarity between 5.3 and the infinite temperature formula 4.4 but quantitatively they are quite different. From 4.4 we see that at $T = \infty$ the size quickly tends to the exponential form $e^{\mathcal{J}t}$. The quadratic growth only persists for a very short time of order $1/\mathcal{J}$. This shows the lack of a throat region.

By contrast, in the low T limit the quadratic growth last for a time of order β which is much greater than $1/\mathcal{J}$, demonstrating the existence of the long throat.

6. NEWTON'S EQUATIONS FOR COMPLEXITY

6.1. Complexity and Momentum

Now we come to the main point, the relation between the evolution of complexity and Newton's equations of motion. Let us compare 4.15,

$$d\mathcal{C} \approx \mathcal{C}\left(\frac{4\pi\, dt}{\tilde{\beta}(t)}\right)$$

and 5.2,

$$\frac{\tilde{\beta}}{4\pi} P \approx \mathcal{C}.$$

Eliminating $\tilde{\beta}$ we find a relation[6]

$$\boxed{P \approx \frac{d\mathcal{C}}{dt}} \tag{6.1}$$

between a dynamical quantity P, and an information-theoretic quantity, complexity:

The momentum of an infalling particle created by ψ is proportional to the rate at which the complexity of the precursor $\psi(t)$ grows.

The numerical constant relating the two sides of 6.1 is connected with the coefficient ϵ in the additional energy of applying a fermion operator the SYK low temperature state ground state.

[6]This relation was derived by Lin et al. by different arguments. See section 7 and Lin et al. [3]. As in other formulas there is an implicit q-dependent proportionality factor.

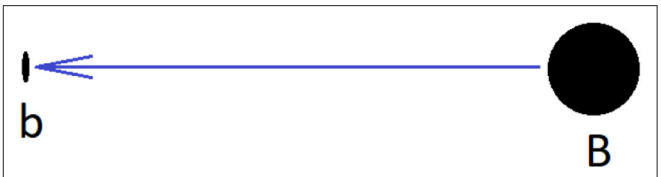

FIGURE 6 | Toy model involving a big and little ball. The big ball represents the boundary and little ball represents the particle. The big ball remains non-relativistic while the little ball quickly become relativistic.

Equation (6.1) resembles the ordinary non-relativistic relation between momentum and velocity. One might be tempted to think that $\frac{d\mathcal{C}}{dt}$ is proportional to the spatial velocity of the infalling particle, but the simple proportionality of momentum and velocity is only valid for non-relativistic motion. The infalling particle however quickly becomes relativistic.

Nevertheless let's proceed to time-differentiate [6.1],

$$\frac{dP}{dt} \approx \frac{d^2\mathcal{C}}{dt^2}. \tag{6.2}$$

We next use the fact that the rate of change of momentum is the applied force,

$$\boxed{F \approx \frac{d^2\mathcal{C}}{dt^2}.} \tag{6.3}$$

In **Supplementary Material**, the force F on an infalling particle in the gravitational field of a NERN black hole is calculated using the standard Lagrangian formulation of particle mechanics. It is explicitly shown to agree with $\frac{d^2\mathcal{C}}{dt^2}$ as calculated from the Qi-Streicher formula—the formula being a pure SYK relation whose derivation does not explicitly involve particle mechanics. This and the interpretation of 6.3 as Newton's equation of motion (despite the comment just before Equation 6.2) are the principle results of this paper.

6.2. Toy Model

Equation (6.3) looks temptingly like Newton's equation $F = ma$ for a non-relativistic particle in a uniform gravitational field but for the reason stated above, it does not make sense to identify that particle with the relativistic infalling particle. To understand what is going on consider a toy model. Two balls, **B** and **b** are shown in **Figure 6**.

One—the big-ball **B**—is very heavy with mass M_B and the other—little-ball **b**—is very light with mass m_b. Initially the two are attached and the combined system is at rest. At $t = 0$ the two balls are ejected from one another along the X axis with equal and opposite momentum. We also assume the balls repel each other with a constant force. The result is that **b** will quickly

become relativistic while **B** remains non-relativistic. Throughout the motion the momenta of the balls are equal and opposite.

It is evident from Newton's third law that both balls satisfy the equations,

$$dP/dt = F \tag{6.4}$$

but only **B** satisfies the non-relativistic Newton equation.

$$F = M_B \frac{d^2X}{dt^2}. \tag{6.5}$$

The connection between the toy model and the NERN system is clear: **b** is the light particle that was dropped from the black hole boundary, and **B** is the boundary itself with mass M_B.

It is also worth noting that the heavy ball **B** serves as a quantum frame of [13]. As Maldacena has noted, this is similar to the way that the condensate of a superfluid or superconductor serves as a frame of reference for a phase variable.

These considerations, along with equation 6.3, lead to the conclusion that it is *the nonrelativistic velocity of the heavy boundary,* not the particle, which is proportional to the rate of change of the complexity of $\psi(t)$, and that it satisfies the Newtonian Equation (6.3).

Since P is conjugate to ρ, and the boundary is non-relativistic, we can write,

$$\begin{aligned} P &= M_B \frac{d\rho_B}{dt} \\ &= \frac{d\mathcal{C}}{dt}. \end{aligned} \tag{6.6}$$

where ρ_B is the location of the boundary. If further follows that,

$$\mathcal{C} = M_B(\rho_B - \rho_0) \tag{6.7}$$

where ρ_0 is constant. The obvious choice is for ρ_0 to be the horizon location in which case $\mathcal{C}$ is proportional to the distance of the boundary from the horizon. In section 7, where the two-sided case is discussed, the distance defining complexity is naturally taken to be the distance separating the two boundaries.

6.3. Comparison With CV

There are a number of ways of estimating the boundary mass M_B. One way is to directly analyze the Schwarzian boundary term in the action. I will do something different making direct use of the complexity-volume (CV) correspondence [14, 15]; volume now referring to the length of the throat times its area. For this subsection I will not bother keeping track of numerical factors.

The standard volume-complexity (CV) relation is,

$$\mathcal{C} = \frac{V}{Gl_{AdS}} \tag{6.8}$$

The volume is the area of the throat times the length ρ,

$$V = A\rho \tag{6.9}$$

where A is the horizon area. Also observe that A/G is proportional to the entropy of the black hole and the AdS radius is proportional to r_+. One finds

$$\mathcal{C} \approx \left(\frac{S}{r_+}\right)\rho \tag{6.10}$$

or using the SYK/NERN dictionary,

$$\mathcal{C} \approx \mathcal{J}N\rho \tag{6.11}$$

From 6.3 we may write,

$$F \approx \mathcal{J}N\frac{d^2\rho}{dt^2}. \tag{6.12}$$

It follows that the mass of the boundary is,

$$\boxed{M_B \approx \mathcal{J}N.} \tag{6.13}$$

This is to be compared with the energy of the infalling particle which is $\mathcal{J}$. The big-ball, little-ball analogy is quite apt. Another point worth noting is that M_B is of the same order as the mass of the NERN black hole.

$$M_{BH} = \frac{r_+}{G} \sim \mathcal{J}N. \tag{6.14}$$

If we now combine 6.12 and 6.13 with Equation (A.14) from the **Supplementary Material** we arrive at Newton's equation,

$$\boxed{\frac{m_b M_B G}{r^2} = M_B \frac{d^2\rho}{dt^2}.} \tag{6.15}$$

for the motion of the boundary[7].

The derivation in **Supplementary Material** of the left side of 6.15 was based on the bulk equation of motion for a particle in a gravitational field. One may wonder whether it can be derived from the holographic SYK quantum mechanics. The answer is that up to factors of order unity, it can. Using the SYK/NERN dictionary in section 3 we can write $\frac{m_b M_B G}{r^2}$ in terms of SYK variables (for $q = 4$),

[7]It should be kept in mind that the r that appears in the inverse square law is not generally the distance of the test particle to the gravitating mass. According to Gauss' law it is the radius of the 2-sphere surrounding the mass at the test point. Only in flat space is it the distance to gravitating mass.

$$(m_b)(M_B)(G)\left(\frac{1}{r_+^2}\right) = (2\mathcal{J})(M_B)\left(\frac{.12}{N\mathcal{J}^2}\right)\left(\frac{\mathcal{J}^2}{.01}\right) \tag{6.16}$$

On the other hand, the right side is just $d^2\mathcal{C}/dt^2$ which can be evaluated from the Qi-Streicher formula. In the throat region one finds the QS formula gives

$$d^2\mathcal{C}/dt^2 = 4J^2. \tag{6.17}$$

Equating the right side of 6.16 to the right side of 6.17 determines the value of M_B,

$$M_B \approx 0.2N\mathcal{J}, \tag{6.18}$$

consistent with 6.13.

There is also information in the Qi-Streicher formula about the relativistic motion of the light particle. For example consider the time that it takes, moving relativistically, for the particle to travel the distance $\Delta\rho = r_+ \log \beta/r_+$ from the boundary to the Rindler region. From 2.16 one sees that the time is β. Once the particle is in the Rindler region the size begins to grow exponentially with time. The Qi-Streicher formula 5.3 shows that this is indeed the case.

7. FORMAL CONSIDERATIONS

7.1. Symmetries of AdS_2

The basis for the derivation of Newton's equations in section 6 was the relation between momentum and the time derivative of complexity (Equation 6.1), which itself was based on the momentum-size correspondence of [2]. The momentum-size correspondence fit some non-trivial facts about scrambling by NERN black holes [12], but it was never derived from first principles. If we had an alternate route to 6.1 we could turn the argument around and derive the momentum-size correspondence. Lin et al. [3] described such a route which I will briefly explain as far as I understand it[8].

We begin by considering the approximate symmetries of matter in the background of a fixed, almost infinite, AdS_2 throat. The Penrose diagram for the throat is shown in **Figure 7**.

The symmetry of infinite AdS_2 is the non-compact group $SL(2,R)$. If β is finite the symmetry is approximate. Deep in the throat the geometry is indistinguishable from AdS_2 but the left and right boundaries break the symmetry. As long as matter is far from the boundaries the symmetry will be respected.

$SL(2,R)$ has three generators called E_0, P_0, B_0, satisfying the algebra,

$$\begin{aligned} [B_0, E_0] &= iP_0 \\ [B_0, P_0] &= iE_0 \\ [P_0, E_0] &= iB_0 \end{aligned} \tag{7.1}$$

[8]I am grateful to Henry Lin and Ying Zhao for explaining the argument to me.

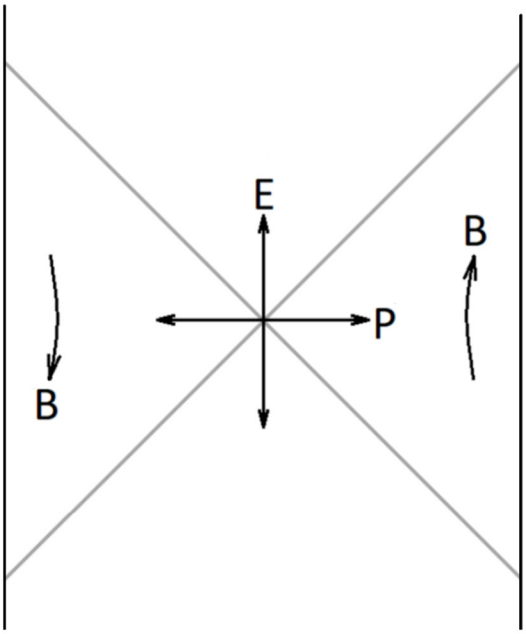

FIGURE 7 | Penrose diagram for a two-sided non-dynamical background in the limit of low temperature and infinite throat length. Also shown are the matter generators E, B, P that generate $SL(2,R)$ motions of the matter fields. The generators have been normalized so that the commutation relations are $[B,E] = iP$, $[B,P] = iE$, $[P,E] = i\mathcal{J}^2 B$.

It is convenient to rescale P and E in order to give them units of energy. Thus define,

$$\begin{aligned} E &= \mathcal{J}E_0 \\ P &= \mathcal{J}P_0 \\ B &= B_0. \end{aligned} \tag{7.2}$$

The commutation relations become,

$$\boxed{[B,E] = iP} \tag{7.3}$$

$$\boxed{[B,P] = iE} \tag{7.4}$$

$$\boxed{[P,E] = i\mathcal{J}^2 B} \tag{7.5}$$

Let's consider the generators one by one. The action of E is to shift the Penrose diagram rigidly in the vertical direction. We can introduce a time variable τ that is constant on horizontal slices,

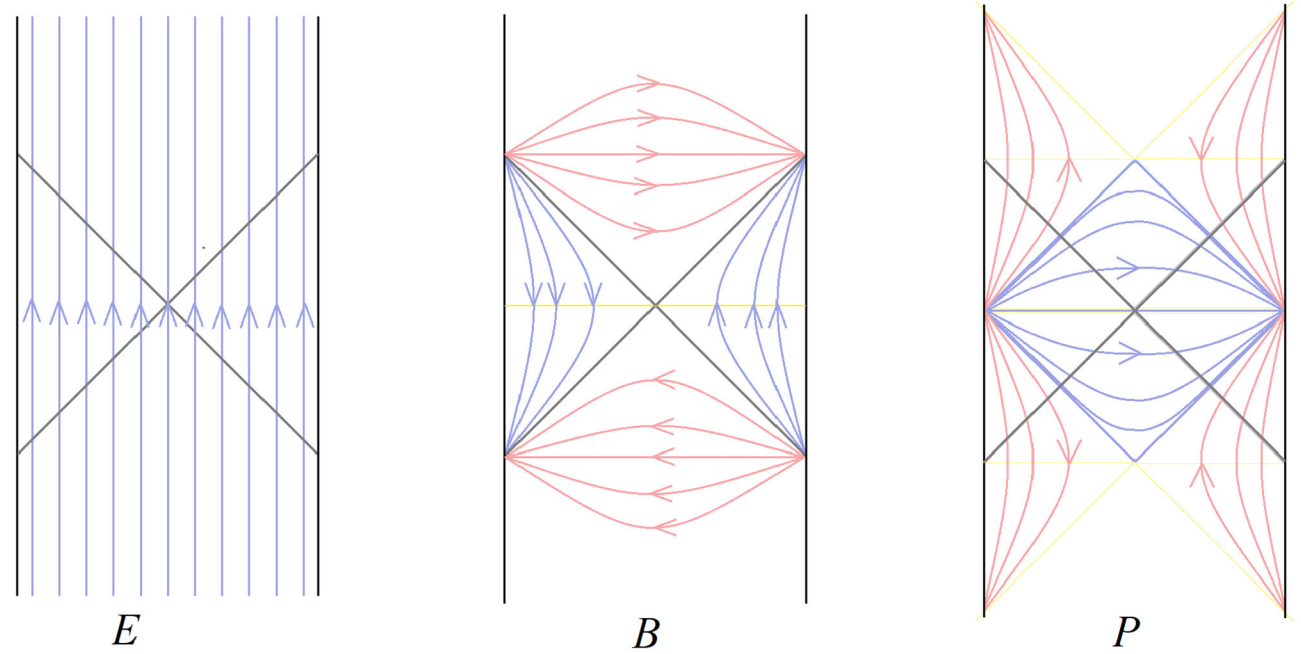

FIGURE 8 | Orbits of the three generators.

and which at the center of the diagram registers proper time. E may be represented by the differential operator,

$$E = i\frac{\partial}{\partial \tau}. \tag{7.6}$$

The generator P shifts the diagram along spacelike directions. It has fixed points at the asymptotic boundaries on the $t = 0$ slice. It may be thought of as the translation generator with respect to the proper coordinate ρ defined in 2.11,

$$P = -i\frac{\partial}{\partial \rho} \tag{7.7}$$

Finally B is the boost generator that has the bifurcate horizon as a fixed point. It is conjugate to the Rindler hyperbolic angle ω.

$$B = -i\frac{\partial}{\partial \omega}. \tag{7.8}$$

The Rindler time is related to t by,

$$\omega = \frac{2\pi t}{\beta} \tag{7.9}$$

so that B can be written,

$$B = -i\frac{\beta}{2\pi}\frac{\partial}{\partial t} \tag{7.10}$$

The orbits of the three generators are shown in **Figure 8**.

The two-sided Penrose diagrams 7 and 8 represents two uncoupled but entangled SYK systems with Hamiltonians H_R and H_L. The generator B is given in terms of the two Hamiltonians by

$$\boxed{B = \frac{\beta}{2\pi}(H_R - H_L)} \tag{7.11}$$

7.2. Left-Right Interaction

One might think that the global energy E should be identified with $\beta\mathcal{J}[H_L + H_R]$. However, there is no symmetry of AdS_2 generated by $(H_L + H_R)$. Without going into details, Maldacena and Qi [16] argue that the generator E requires the introduction of another term, H_{int} that couples the left and right sides,

$$\boxed{E = \beta\mathcal{J}(H_L + H_R + H_{int}).} \tag{7.12}$$

Using

$$i[B, E] = P$$

and

$$B = i\beta\frac{d}{dt}$$

we can write

$$P = i\beta\mathcal{J}[B, H_{int}]$$
$$= \beta^2\mathcal{J}\frac{dH_{int}}{dt} \tag{7.13}$$

In [5], an operator representing size was constructed in terms of the two-sided degrees of freedom ψ_{iL} and ψ_{iR}. Using our convention of calling size $\mathcal{C}$,

$$\mathcal{C} = \frac{i}{\delta_\beta}\sum_i \psi_{iL}\psi_{iR} \tag{7.14}$$

where δ_β is a dimensionless normalization factor which normalizes the size of a single fermion to unity. This same operator appears in the interaction term H_{int} in [16],

$$H_{int} = i\mu\sum_i \psi_{iL}\psi_{iR}.$$
$$= \mu\delta_\beta\,\mathcal{C} \tag{7.15}$$

Combining 7.15 with 7.13 we find,

$$\boxed{P = \mu\delta_\beta\beta^2\mathcal{J}\,\frac{d\mathcal{C}}{dt}} \tag{7.16}$$

Thus, apart from the factor $\mu\delta_\beta\beta^2\mathcal{J}$ the matter momentum P is indeed proportional to the time derivative of the size. However, consistency with 6.1 requires a relation between the parameters μ, δ_β, β, and $\mathcal{J}$,

$$\mu\delta_\beta\beta^2\mathcal{J} \approx 1. \tag{7.17}$$

Again, the meaning of $\approx$ in 7.17 is: *equals up to a numerical constant which may depend on q*. This is a significant constraint since the parameters μ and δ_β have an intricate mixed dependence [16] on q and the dimensionless quantity $\beta\mathcal{J}$.

7.3. Determining the Prefactor

It is known that the quantity μ is not independent of the other three parameters and that there is a relation between them. Zhao[9] has suggested that the coefficient $\mu\delta_\beta\beta^2\mathcal{J}$ can be determined by comparing the calculation of $P(t)$ using the equation of motion in **Supplementary Material**, with the Qi-Streicher formula 5.3. From the **Supplementary Material** the force on the infalling particle is constant during passage through the throat and given by $F \approx \mathcal{J}^2$. It follows that,

$$P(t) \approx \mathcal{J}^2 t. \tag{7.18}$$

[9]Unpublished communication.

Differentiating the Qi-Streicher formula also gives,

$$\frac{d\mathcal{C}}{dt} = 4\mathcal{J}^2 t. \tag{7.19}$$

(In **Supplementary Material**, a more complete comparison between the particle orbit and the Qi-Streicher formula is carried out for the entire range of ρ from the boundary at $r = 2r_+$ to the horizon at $r = r_+$.)

It follows that the coefficient μ must satisfy,

$$\boxed{\mu\delta_\beta\beta^2\mathcal{J} \approx 1} \tag{7.20}$$

so that 6.1 is satisfied. Equation (7.20) is non-trivial. On dimensional grounds the q can appear in any combination with the product $\beta\mathcal{J}$, but 7.20 allows only a multiplicative dependence by a function of q alone.

That the product in 7.20 only depends on q is non-trivial and is confirmed in the analysis of [16] where it appears in a somewhat hidden form in Equations [4.25], [4.29], and [4.50].

The formal considerations of this section did not involve the momentum-size correspondence 5.2 postulated in [1, 2] but they would allow us to work backward from 6.1 and derive it.

We are almost where we want to be, but not quite because we have assumed the throat is infinite. If we make the throat finite by allowing T to be small but not zero, the symmetry of the matter system will be broken by the interaction of the matter with the boundary. In a sense that's not surprising since the matter will interact with the dynamical boundary (through the potential barrier) so that the momentum of the matter will not, by itself, be conserved.

There is a formal way to restore the symmetry as a gauge symmetry [3, 8, 11]. Although the finite throat does not have $SL(2, R)$ symmetry it can be embedded in AdS_2 as illustrated in **Figure 9**.

The curved boundary separating the blue regions from rest of the diagram represents the Schwarzian boundary. The Penrose diagram can be conveniently parameterized by dimensionless coordinates $-\infty < T < \infty$ and $0 < X < \pi$. The embedding is not unique due to the $SL(2, R)$ invariance of AdS_2. This invariance allows us to move the geometry in various ways. In other words the representation of the finite throat in AdS_2 is redundant; the symmetry is a gauge symmetry. As such its generators should be set to zero. Denoting the gauge generators by tilde-symbols,

$$\boxed{\tilde{E} = \tilde{B} = \tilde{P} = 0} \tag{7.21}$$

But the tilde generators are no longer the matter charges; they now include the charges of the boundary. In particular the spatial charge $\tilde{P}$ is,

$$\tilde{P} = P + P_{boundary}. \tag{7.22}$$

Therefore, the gauge condition

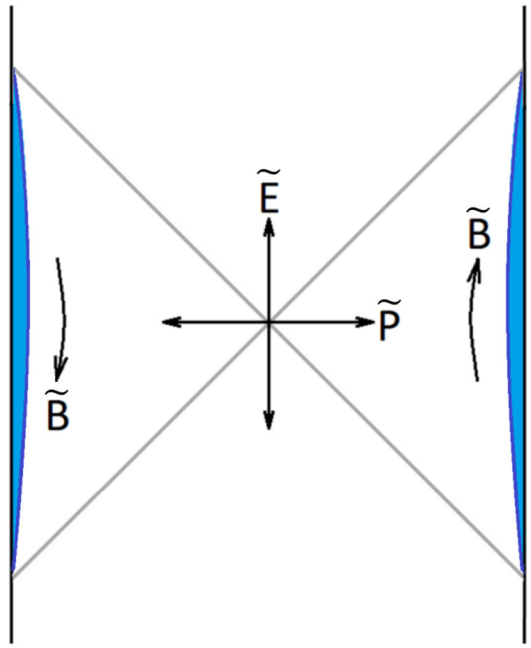

FIGURE 9 | Embedding a finite throated geometry in AdS_2. Also shown are the three $SL(2,R)$ gauge generators. The blue regions are part of the embedding space but not part of the actual finite temperature spacetime. The inner boundaries of the blue region are the dynamical boundaries governed by the Schwarzian action.

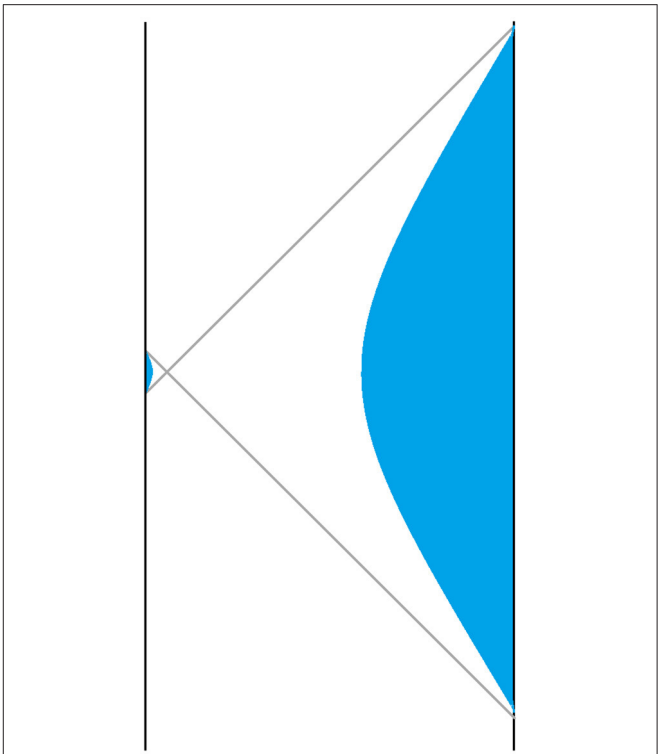

FIGURE 10 | Fixing a gauge.

$$\tilde{P} = 0 \tag{7.23}$$

is the Newtons third law of action and reaction, which tells us that the boundary recoils when the matter particle is emitted into the throat. Keeping track of the action=reaction condition seems to be the main point of the gauge symmetry. The un-tilded operators are the physical matter generators and their negatives are the generators that act on the boundary degrees of freedom.

7.4. Fixing a Gauge

There are a number of ways of insuring gauge invariance. One way is to construct manifestly gauge invariant objects and work with them. Lin et al. [3] does this. The other way is to completely fix the gauge so that there is no residual gauge freedom. I'll illustrate such a gauge-fixing here.

The embedding is not unique due to the $SL(2,R)$ invariance of AdS_2. This invariance allows us to move the entire geometry—matter and boundary—in various ways by applying the three gauge generators.

The action of $\tilde{P}$ moves the bifurcate horizon as well as the excised (blue) regions. Such a transformation can shift the NERN geometry from **Figure 9** to **Figure 10**.

We can use the gauge symmetries them to fix a convenient gauge:

- The left black hole has a bifucate horizon. Using the $\tilde{E}$ symmetry we can shift it to the $t = 0$ slice.
- Next we can use $\tilde{P}$ to shift the position of the right boundary so that it passes through the spatial midpoint of the diagram on the $t = 0$ slice. More generally we can choose a point X_0 in along the $t = 0$ surface and have the boundary pass through it. This defines a one parameter family of gauges parameterized by X_0.
- Finally we can fix the boost symmetry by assuming a particle is dropped from the right boundary at $t = 0$.

That completely fixes the gauge. The resulting Penrose diagram is shown in 11.

Notice that in the limit that the temperature goes to zero that the bifurcate horizon moves all the way to the left boundary. The right Rindler patch becomes the Poincare patch, and the boosts become Poincare time translations. Again there is a one parameter family parameterized by X_0. The boost operator $\tilde{B}$ may now be used to boost the $t = 0$ surface forward in time to as illustrated by the green line in **Figure 11**.

The transformations generated by $\tilde{P}$ are shifts of the X_0 parameter that move the right boundary. The momentum of the infalling particle that we called $P(t)$ is the proper momentum on that slice.

Dropping the particle from the right-side boundary causes the boundary to recoil and move outward. That is indicated by the small separation shown as light blue. The effect is to change the right-side horizon (not shown) so that its bifurcate point is no longer on the $t = 0$ surface but is slightly below it. The bifurcate point on the left horizon is unchanged.

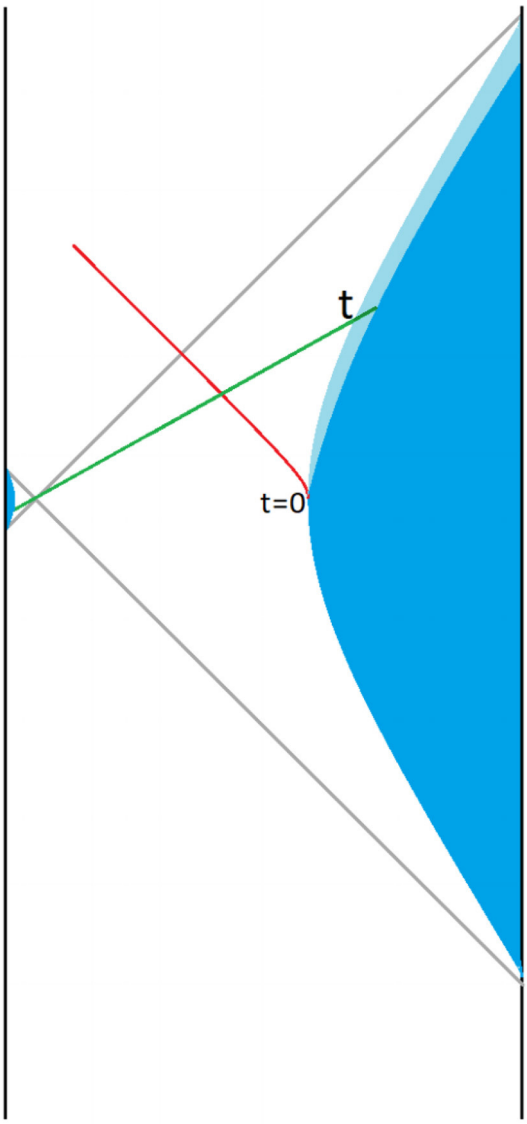

FIGURE 11 | The gauge fixed Penrose diagram with the right boundary intersecting the $t = 0$ surface at a fixed location half way between boundaries. The red curve is the world line of a particle dropped at $t = 0$ from the right boundary. The green surface is boosted from the $t = 0$ surface. The boost time t is the time variable that corresponds to the earlier discussion.

The time-slice shown as green is anchored on the boundaries at "boost time" t. The holographic quantum system—two copies of SYK—has a quantum state associated with the time slice and if the particle had not been thrown in, the state would be independent of the time t. But the insertion of ψ_R at $t = 0$ breaks the boost symmetry and the state evolves with t. Since ψ_R is a purely right-side operator it evolves according to,

$$\begin{aligned}\psi(t) &= e^{-i(H_R-H_L)t}\psi e^{i(H_R-H_L)t} \\ &= e^{-iH_Rt}\psi e^{iH_Rt}. \end{aligned} \tag{7.24}$$

Under this evolution $\psi_R(t)$ grows in the way I described earlier.

The complexity of the evolving state can be determined from CV duality. Apart from some constant factors it is just the length of the geodesic connecting the left and right boundaries at time t. If the particle had not been thrown in, the boost symmetry would imply that the length/complexity would be constant, but the small kick causes the length/complexity to grow after the particle is dropped in.

Lin-Maldacena-Zhao argue that the generators can be decomposed into bulk matter, and gravitational (boundary) contributions. The bulk matter contribution to $\tilde{P}$ is the momentum P. In the case in which a particle has been dropped into the geometry, P is the particle's momentum. The gravitational part on the other hand is the momentum of the heavy non-relativistic boundary, which by the gauge condition is $-P$. (In the case at hand only the right boundary recoils. The momentum of the left boundary stays zero.) The fact that the sum of the particle and boundary momentum must be zero is Newton's third law of action and reaction.

The low energy $SL(2R)$ symmetry of SYK dictates a particular form for the action governing the motion of the boundary. Known as the Schwarzian action, it is equivalent to the Gibbons-Hawking-York extrinsic curvature that has to be added to the Einstein Maxwell action in the presence of boundaries. It's rather complicated but in the non-relativistic limit when the boundary moves slowly, the kinetic term in the Schwarzian action must reduce to the action for a non-relativistic particle[10] of mass $M_B = N\mathcal{J}$, or in NERN terms, $M_B = S/r_+$.

$$I \approx \frac{1}{2}M_B\,\dot{\rho}^2. \tag{7.25}$$

This agrees with the analysis in the previous section and provides a formal justification for it.

In addition there is a coupling between the matter and the boundary which has the form of a repulsive potential energy. As long as the particle is in the throat region the potential is linear in the distance between the infalling particle and the boundary. As shown in the **Supplementary Material** this leads to a constant Newtonian force which accelerates both the particle and the boundary in opposite directions, so as to keep the total momentum zero. The result is that the particle is effectively attracted toward the horizon, and as it falls the complexity grows according to the pattern described in earlier sections and in **Supplementary Material**.

8. FALLING THROUGH EMPTY ADS$_2$

Susskind [1] and Brown et al. [2], and the present paper up to this point, deal with the gravitational attraction of a black hole. If the tendency for complexity to increase is the general holographic mechanism behind gravitation it is important to demonstrate it outside the black hole context. For example we would like to know when a particle falls toward an ordinary cold mass with

[10] I am grateful to Herny Lin for a helpful discussion of this point.

little or no entropy, does the holographic complexity grow? What happens when a comet falls in a long elliptical orbit toward the sun and then goes off into interstellar space. Does the complexity increase and decrease periodically?

We could try modeling questions like this in AdS/CFT, but the tools I've used in this paper are special to SYK. Fortunately there is a simple case in which the question can be addressed. Anti de Sitter space has a gravitational field even in the AdS vacuum. The negative vacuum energy of AdS gravitates and attracts matter to the center. One does not need an additional mass.

The metric of AdS is,

$$\begin{aligned} ds^2 &= -f(r)dt^2 + \frac{1}{f(r)}dr^2 + r^2 d\Omega^2 \\ f(r) &= \left(1 + \frac{r^2}{l_{AdS}^2}\right) \end{aligned} \tag{8.1}$$

Particles dropped from a distance experience an attractive radial gravitational force which behaves similarly to a harmonic oscillator force. A particle will move in a periodic orbit oscillating about the origin. There is no black hole, no horizon, no entropy.

Two dimensional AdS is not an exception, but engineering empty AdS_2 is subtle in the SYK system. Maldacena and Qi [16] arrange it by perturbing a two-sided black hole with a Left-Right interaction. The resulting space is called a traversable wormhole; in fact it is a cutoff version of AdS_2. The geometry does not extend out to $r = \infty$, but instead is cut off at some large radial distance by a Schwarzian boundary, or to be precise, two Schwarzian boundaries[11]: one for the left side and one for the right side, as in **Figure 12**. The geometry is AdS_2 except that the blue regions near the boundary have been excised.

In **Figure 13**, by applying a right-side fermion operator a particle can be dropped in from the right boundary. The initial state has the form

$$\psi_R|0\rangle$$

and subsequently evolves to

$$\psi_R(t)|0\rangle.$$

In this case there is no black hole and the particle endlessly oscillates back and forth between the two boundaries.

The force on the particle is gravitational. From the bulk GR viewpoint it is produced by the vacuum energy in the region between the boundaries. The state without the particle (**Figure 12**) is the ground state of the Hamiltonian and the complexity—in this case represented by the distance between the two boundaries—is constant in time.

When the particle is injected at $t = 0$ by applying the right-side fermion operator ψ_R the additional complexity of the state is initially very small. As the particle accelerates toward the center of AdS its momentum increases. The right boundary recoils so

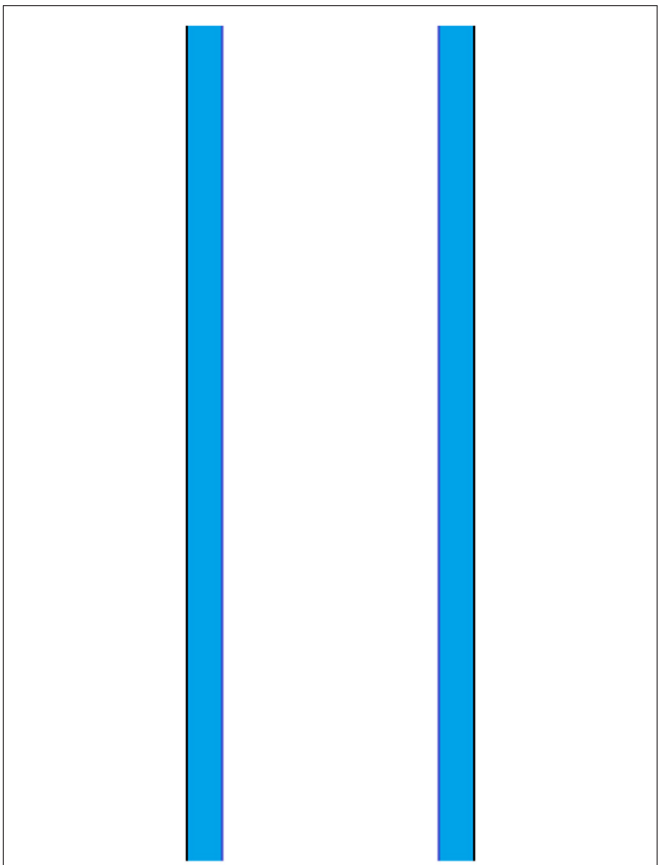

FIGURE 12 | Traversable wormhole with two boundaries.

that the distance between the boundaries increases. According to CV duality, the complexity also increases.

Because the boundary is very heavy it moves non-relativistically which means its momentum and velocity are proportional to one another, and once again,

$$P \approx \frac{dC}{dt} \tag{8.2}$$

for both the boundary and for the particle.

The radial momentum reaches a maximum when the particle reaches the center of the diagram. It then switches sign. At the same time the complexity starts to decrease[12]. By the time the particle reaches the left boundary the complexity has decreased to its original value. The state at that point is

$$\psi_L|0\rangle.$$

The particle then gravitates back to the center and subsequently returns to the right boundary. The oscillating behavior of complexity may seem odd, but in fact it is generic for integrable

[11] AdS two is unique in having two disconnected boundaries.

[12] This conclusion is based on the ability of the gravitational dressing to switch from the right to the left side. Such switching would be impossible without left-right coupling, but there is no obstruction to it when the Maldacena-Qi interaction is included.

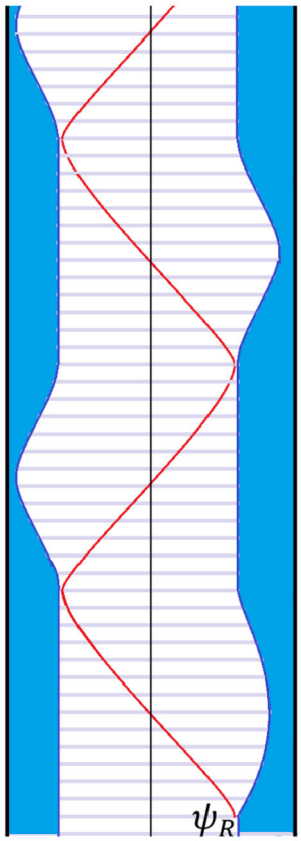

FIGURE 13 | A particle has been added to the right side of the traversable wormhole by acting with ψ. The subsequent motion is oscillatory with periodic variation in the distance between the boundaries, thus indicating periodic variation of complexity. The figure has been foliated with constant time slices to help guide the eye. The oscillations of the boundary are very small and have been greatly exaggerated.

systems. It is also characteristic of holographic systems below the black hole threshold [17].

To reiterate, the connection between gravitational attraction and complexity is not dependent on the presence of a black hole, or on the presence of a system with a large entropy. However without a black hole the system is integrable and the complexity oscillates. It should be pointed out that the complexity never get's very large during the oscillating behavior. At the maximum when the particle is at the center of the geometry the complexity is $\sim \beta^2 J^2$ which is much less than N, i.e., the complexity at scrambling.

It would be interesting to confirm this behavior in the SYK theory using the methods of Qi and Streicher.

9. CONCLUDING REMARKS

In this article, I have assembled further evidence that the holographic avatar of gravitational attraction is the growth of operator-size during the run-up to the scrambling time. During this period, size and complexity are indistinguishable, and one can say that gravitational attraction is an example of the tendency for complexity to increase. The presence of a massive object creates a kind of complexity-force, driving the system toward greater complexity in the same way that an ordinary force accelerates a particle toward lower potential energy. This conclusion was based on three things: the CV correspondence between complexity and volume; a duality between momentum and the time-derivative of complexity,

$$P \approx \frac{d\mathcal{C}}{dt};$$

and the Qi-Streicher calculation of the time dependence of size at low temperature.

To test the duality, on the left side we used the standard relativistic classical theory of particle motion (in a gravitational field) to compute $P(t)$. On the other side the Qi-Streicher calculation of $\mathcal{C}(t)$ (a pure quantum calculation that makes no reference to particle motion) allows us to compute $\frac{d\mathcal{C}}{dt}$. The two sides agree.

One can object to such a connection (between momentum and complexity) on the grounds that it relates two fundamentally different kinds of quantities. Momentum is a linear quantum observable. Complexity is a nonlinear property of states; linear superpositions of states with the same complexity may have very different complexity. Thus, equating momentum and the time-derivative of complexity is inappropriately mixing concepts.

Similar things have been seen before. The Bekenstein formula and more recently, the Ryu-Takyanagi formula, equate area—a quantum observable—to entropy. This also seems inadmissible for similar reasons. A number of authors have written about this tension [see for example [18, 19] and references therein] and the resolution seems to be that quantities like entropy may behave like observables over a relatively small subspace of states—a so called code subspace. Thus, for states near the ground state of AdS, area and entanglement entropy can coincide, but the relation does not hold for most states.

The same things should be true for complexity: in the small subspace of states encountered while a particle is falling toward the horizon of a black hole complexity and its derivative can behave like an observable, but beyond the scrambling time or when superpositions of classical states are considered the relation between complexity and observables must break down[13].

On another point, E. Verlinde has also emphasized the need for a holographic explanation of gravitational attraction and has proposed an entropic mechanism [20]. He argues that lowering an object toward a horizon increases the thermodynamic entropy an entropic force. What I find unclear is how an entropic mechanism can explain the gravitational pull-to-the-center in cold empty AdS, or to a conventional zero temperature massive body in its (non-degenerate) ground state. How can an entropic theory be compatible with the periodic oscillations of the distance between the sun and a comet in an elongated orbit?

In contrast to coarse-grained thermal entropy, complexity and operator size can oscillate, especially for non-chaotic or weakly chaotic systems. By the complexity-volume correspondence, the oscillating complexity may manifest itself as periodic motion.

[13] I am grateful to Daniel Harlow for discussions about this point.

The motion of a particle in empty AdS_2, discussed in section 8 is an example.

Returning to the case of a black hole, entropy approaches its maximum value well before the scrambling time, but as shown in [1] and [2], under the influence of gravity, the infalling momentum increases exponentially until the scrambling time has been reached. Again it is not obvious how an entropic theory would deal with this.

It is quite possible that these remarks represent my own misunderstanding of Verlinde's theory.

Finally I would like to emphasize the importance of generalizing the concept of size to a wider class of gauge-gravity dualities. In a strongly coupled CFT it is not obvious what the fundamental constituents are, that are counted when we speak of size. I hope to come back to this issue in the future.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I thank Henry Lin and Ying Zhao for very helpful discussions about both the heuristic and formal arguments in this paper. The paper would not have been written without the many discussions I had with Alex Streicher, in which he explained his results on the growth of size in SYK, and related issues. This research was supported by NSF Award Number 1316699. The article has appeared as a preprint.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2020.00262/full#supplementary-material

REFERENCES

1. Susskind L. Why do things fall? *arxiv [Preprint] arXiv:1802.01198 [hep-th].* (2018)
2. Brown AR, Gharibyan H, Streicher A, Susskind L, Thorlacius L, Zhao Y. Falling toward charged black holes. *Phys Rev D.* (2018) 98:126016. doi: 10.1103/PhysRevD.98.126016
3. Lin H, Maldacena J, Zhao Y. *Symmetries Near the Horizon.* (2019). Available online at: http://pirsa.org/displayFlash.php?id=19040126
4. Roberts DA, Stanford D, Streicher A. Operator growth in the SYK model. *J High Energy Phys.* (2018) 1806:122. doi: 10.1007/JHEP06(2018)122
5. Qi XL, Streicher A. Quantum epidemiology: operator growth, thermal effects, and SYK. *arxiv [Preprint] arXiv:1810.11958 [hep-th].* (2019). doi: 10.1007/JHEP08(2019)012
6. Brown AR, Gharibyan H, Lin HW, Susskind L, Thorlacius L, Zhao Y. Complexity of Jackiw-Teitelboim gravity. *Phys Rev D.* (2019) 99:046016. doi: 10.1103/PhysRevD.99.046016
7. Maldacena J, Stanford D. Remarks on the Sachdev-Ye-Kitaev model. *Phys Rev D.* (2016) 94:106002. doi: 10.1103/PhysRevD.94.106002
8. Maldacena J, Stanford D, Yang Z. Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space. *PTEP.* (2016) 2016:12C104. doi: 10.1093/ptep/ptw124
9. Maldacena J, Stanford D, Yang Z. Diving into traversable wormholes. *Fortsch Phys.* (2017) 65:1700034. doi: 10.1002/prop.201700034
10. Kitaev A, Suh SJ. Statistical mechanics of a two-dimensional black hole. *arxiv [Preprint] arXiv:1808.07032 [hep-th].*
11. Yang Z. The quantum gravity dynamics of near extremal black holes. *arxiv [Preprint] arXiv:1809.08647 [hep-th].* (2018). doi: 10.1007/JHEP05(2019)205
12. Leichenauer S. Disrupting entanglement of black holes. *Phys Rev D.* (2014) 90:046009. doi: 10.1103/PhysRevD.90.046009
13. Aharonov Y, Susskind L. Charge superselection rule. *Phys Rev.* (1967) 155:1428. doi: 10.1103/PhysRev.155.1428
14. Brown AR, Roberts DA, Susskind L, Swingle B, Zhao Y. Holographic complexity equals bulk action? *Phys Rev Lett.* (2016) 116:191301. doi: 10.1103/PhysRevLett.116.191301
15. Carmi D, Chapman S, Marrochio H, Myers RC, Sugishita S. On the time dependence of holographic complexity. *J High Energy Phys.* (2017) 1711:188. doi: 10.1007/JHEP11(2017)188
16. Maldacena J, Qi XL. Eternal traversable wormhole. *arxiv [Preprint] arXiv:1804.00491 [hep-th].* (2018)
17. Anous T, Sonner J. Phases of scrambling in eigenstates. *arxiv [Preprint] arXiv:1903.03143 [hep-th].* (2019). doi: 10.21468/SciPostPhys.7.1.003
18. Papadodimas K, Raju S. Remarks on the necessity and implications of state-dependence in the black hole interior. *Phys Rev D.* (2016) 93:084049. doi: 10.1103/PhysRevD.93.084049
19. Harlow D. The Ryu–Takayanagi formula from quantum error correction. *Commun Math Phys.* (2017) 354:865. doi: 10.1007/s00220-017-2904-z
20. Verlinde EP. On the origin of gravity and the laws of Newton. *J High Energy Phys.* (2011) 1104:029. doi: 10.1007/JHEP04(2011)029

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 19 May 2021
doi: 10.3389/fphy.2021.655857

Replacing the Notion of Spacetime Distance by the Notion of Correlation

*Achim Kempf**

Department of Applied Mathematics, University of Waterloo, Waterloo, ON, Canada

Spacetime is conventionally viewed as a stage on which actors, in the form of massive and massless matter, move. In this study, we explore what may lie beyond this picture. The starting point is the observation that quantum field fluctuations are the more strongly correlated the shorter their spacetime distance. The notion of spacetime distance can, therefore, be replaced by the notion of correlation strength. This suggests a new picture in which the abstract 2-point and multi-point correlations are the primary structure, a picture which is essentially information-theoretic. In the low energy regime, the secondary notions of spacetime and of matter would then emerge as approximate representations of the abstract correlators, namely, in the form of Feynman rules on curved spacetime.

Keywords: quantum gravity, spectral geometry, shannon sampling theory, quantum information, general relativity

OPEN ACCESS

Edited by:
Caslav Brukner,
University of Vienna, Austria

Reviewed by:
Maximilian Lock,
Institute for Quantum Optics and Quantum Information (OAW), Austria
Luis C. Barbado,
University of Vienna, Austria

***Correspondence:**
Achim Kempf
akempf@uwaterloo.ca

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: *19 January 2021*
Accepted: *19 April 2021*
Published: *19 May 2021*

Citation:
Kempf A (2021) Replacing the Notion of Spacetime Distance by the Notion of Correlation. Front. Phys. 9:655857. doi: 10.3389/fphy.2021.655857

1. INTRODUCTION

The discoveries of general relativity and quantum theory, each, required the abandoning of major misconceptions. Today, the fact that it has turned out to be extremely hard to unify quantum theory and general relativity suggests that at least one more major misconception will need to be overcome. But what deeply held belief about the nature of spacetime and matter may need to be abandoned to clear the path for the development of the theory of quantum gravity?

One belief regarding spacetime and matter is that, while they do interact, they are fundamentally different, with spacetime representing a stage, which is itself dynamical, on which actors, in the form of massive or massless matter, move. The present study, which lays out ideas first presented orally in [1], asks if the stage-and-actors picture could be a misconception that needs to be abandoned, and it explores one possibility for what new picture lies beyond.

2. PROBING THE DESCRIPTION OF SPACETIME OF GENERAL RELATIVITY

For inspiration, we can take hints from some of the currently most successful descriptions of spacetime. One general relativistic description of a spacetime is as a pair, (M, g), where M is a differentiable manifold and g is a Lorentzian metric. Equivalently, a spacetime is often described as a manifold with Christoffel symbols, Γ, or a connection 1-form, ω. Also, equivalently, general relativity describes a spacetime as a pair (M, σ), where M is the differentiable manifold and $\sigma(x, x') = \frac{1}{2}\tilde{g}(x, x')$ is the Synge world function. Here, $\tilde{g}(x, x')$ is the geodesic distance between the events x and x', as far as that distance is unique. The Synge function, σ, contains all information about a spacetime since it allows one to recover its metric [2–4]:

$$g_{\mu\nu}(x) = \lim_{x \to x'} \sigma(x, x')_{;\mu\nu} = - \lim_{x \to x'} \frac{\partial}{\partial x^{\mu}} \frac{\partial}{\partial x'^{\nu}} \sigma(x, x') \tag{1}$$

For later reference, notice that Equation (1) proves that knowing the bi-scalar function $\tilde{g}$ in an infinitesimal neighborhood of its diagonal is sufficient. Finally, we should add that, to complete the general relativistic descriptions of spacetimes, Einstein also provided an exact mapping between the mathematical concepts and concrete physical measurements, based on rods and clocks [5].

To search for hints at what may lie beyond general relativity, let us now probe general relativity for curiosities or odd features in the way that it describes spacetime. One curious feature of the description of spacetime of general relativity is that it utilizes two different notions of distance. One notion of distance is that of coordinate distance. Finite coordinate distances are not covariant, but infinitesimal coordinate distances are used, covariantly, to define the topology (in the sense of open neighborhoods) of the manifold, which in turn is used to define the notions of continuity and differentiability of the manifold and also to define the limit taking for derivatives and integrals. The second notion of distance is the geodesic distance.

Further, this leads to the curious feature of general relativity that the topology of a spacetime manifold (in the sense of open sets) is ignorant of the drama of the light cone. Points that are arbitrarily close neighbors with respect to the topology of the manifold (in the sense of open sets) can be on either side of a light cone, which is the difference between being fully causally connected or not at all.

Underlying the presence of two notions of distance in general relativity is another curious feature of general relativity. On the one hand, general relativity describes space and time identically, except for a minus sign in the metric. On the other hand, space and time are to be measured using rods and clocks, but rod-like and clock-like physical instruments appear to differ by substantially more than a minus sign. It also appears odd that rods and clocks should play a fundamental role in general relativity, given that they each measure finite and, therefore, non-covariant coordinate distances. Further, nature does not provide rods or clocks in Einstein's sense at sub-atomic scales.

For now, let us take away the hint that it is worth considering to replace rods and clocks in general relativity in some way with tools that are more canonical.

One may ask, for example, whether it is useful to replace traditional clocks that define a notion of time by counting periodic processes with clocks that measure a notion of time that is based on the exponential decay of unstable particles. Intuitively, this would amount to switching from clocks describable by the oscillations of a complex exponential function to a decay-based clock that measures time through a decay-describing real exponential function, thereby possibly accounting for the extra sign in the signature of the metric. Still, even such decay-based, rather than oscillation-based, clocks would appear to differ from rods in their physical appearance by more than a minus sign. We will in the present study, therefore, not follow these lines.

Instead, let us explore how both clocks, as well as rods, could be replaced by tools that are more canonical in the sense that they allow us to determine spacetime distances directly instead of inferring them from measurements of spatial and temporal distances which are themselves not covariant. This could provide us not only with a more canonical method for measuring the spacetime distance between two events. We should thereby also obtain a new method to map a spacetime's curvature. This is because, as Equation (1) showed, knowledge of the (even just infinitesimal) spacetime distances is sufficient to calculate the metric.

3. MEASURING SPACETIME DISTANCES BY MEANS OF CORRELATORS

There may exist multiple ways to replace rod-like and clock-like tools by more canonical tools for measuring spacetime distances. In the present study, the idea is to replace rods and clocks with quantum field vacuum fluctuations. This is possible because the quantum fluctuations of a field are correlated and the strength of the correlation decays with the magnitude of the spacetime distance. The strength of the correlation can, therefore, serve as a measure of the spacetime distance.

For an example of a correlator of the fluctuations of a quantum field, let us recall the Feynman propagator of a free massless scalar. In flat spacetime, it reads:

$$G_F(x,x') = \langle 0|T\hat{\phi}(x)\hat{\phi}(x')|0\rangle \tag{2}$$

$$= -\int \frac{d^4p}{(2\pi)^4}\frac{e^{-ip_\mu(x^\mu-x'^\mu)}}{p_\mu p^\mu + i\epsilon} \tag{3}$$

$$= \frac{1}{4i\pi^2}\frac{1}{(x_\mu - x'_\mu)(x^\mu - x'^\mu) - i\epsilon}. \tag{4}$$

In this case, as well as in curved spacetime [3, 4, 6], the correlator $G_F(x,x')$ is finite both inside and outside the lightcone. On the lightcone, it diverges and changes sign. Important to note here is that as we move away from the lightcone, i.e., as we increase the magnitude of the spacetime distance, the smaller the absolute value of $G_F(x,x')$ becomes, i.e., the weaker do the correlations become. Let us review the reasons why the correlator decays away from the lightcone into both the timelike and the spacelike regions.

First, inside the lightcone, the correlations are caused by the propagation of perturbations. In the course of the propagation, the perturbation spreads out and, therefore, weakens[1] at a rate that depends on the spacetime dimension. Hence, inside the lightcone, the correlator decays for increasing timelike distances.

Second, the correlations outside the lightcone exist because the vacuum is a spatially entangled state. To see why the correlations decay away from the lightcone, i.e., here for increasing spacelike distances, let us recall the simple case of a quantized, massive Klein Gordon field. It can be viewed as consisting of one degree of freedom $\hat{\phi}(x,t)$ at each position, x. Holding a position x fixed, its degree of freedom, $\hat{\phi}(x,t)$, would obey an independent quantum harmonic oscillator equation

$$\ddot{\hat{\phi}}(x,t) = -m^2\hat{\phi}(x,t) \tag{5}$$

[1] It weakens, at least, initially. For finite propagation distances in curved spacetime, there can be, for example, lensing-type and echo effects (see, e.g., [7]).

if it were not for the existence of the Laplacian term in the Klein Gordon equation:

$$\ddot{\hat{\phi}}(x,t) - \Delta\hat{\phi}(x,t) = -m^2\hat{\phi}(x,t) \tag{6}$$

The Laplacian term couples spatially neighboring harmonic oscillators. Therefore, the ground state of these coupled harmonic field oscillators is an entangled state, hence the correlations. Since the Laplacian couples only neighboring field oscillators, these correlations decay with the spacelike distance at a rate that is dimension dependent. In 1+3 dimensions, the decay is polynomial for massless fields and exponential for massive fields[2] (see, e.g., [8–12] for early work, [13] for more recent work, and [14] for a recent review of the related topic of holography and quantum information).

So far, we have established that a correlator of quantum field fluctuations, such as the Feynman propagator, $G_F(x,x')$, can be used as a measure of spacetime distances, or at least of small spacetime distances. From Equation (1), we know that knowledge of small spacetime distances, in the form of knowledge of the Synge function $\sigma(x,x')$ near its diagonal, is sufficient to calculate the metric. It was shown in [15] that, knowledge of $G_F(x,x')$ near its diagonal suffices to calculate the metric tensor:

$$g_{\mu\nu}(x) = -\frac{1}{2}\left(\frac{\Gamma(D/2-1)}{4\pi^{D/2}}\right)^{\frac{2}{D-2}} \lim_{x\to y}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial y^\nu}\left(G_F(x,y)^{\frac{2}{2-D}}\right) \tag{7}$$

Here, D is the spacetime dimension[3]. The fact that the metric tensor can be calculated from the Feynman propagator can also be seen by this consideration: Knowledge of the Feynman propagator implies knowledge of the lightcones because the propagator diverges and changes sign on the lightcones. But knowledge of the lightcones of a spacetime manifold determines the metric tensor of the spacetime up to a conformal factor, as shown in [16]. The propagator also provides the remaining conformal factor through its finite decay near the lightcone.

For completeness, it is worth mentioning that the reconstruction of the metric from the Feynman propagator does not depend on the vacuum state. This is important because different observers may identify different states as their vacuum state. In the absence of a unique vacuum state, Feynman propagators can differ by homogeneous solutions to their equation of motion, such as the Klein Gordon equation $\Box G_F = \delta/\sqrt{(}g)$. The metric is calculated, as shown in Equation (7), by differentiating a negative power of the Feynman propagator, i.e., by differentiating a positive power of the wave operator, $\Box$. The matrix elements of $\Box$ are independent of which homogeneous solution one may choose to define a Feynman propagator, i.e., a right inverse, G_F of $\Box$. Concretely, in Equation (7), any choice of $i\epsilon$ prescription for the propagator drops out because $i\epsilon$ prescriptions are in the denominator, but since the propagator appears to a negative power, ϵ is in the numerator. Hence, the limit $\epsilon \to 0$ can be taken before using Equation (7) to calculate the metric.

[2]This is the case at least for small distances. At large distances, on curved spacetimes, there may again occur, for example, analogs of lensing-type effects.
[3]The case $D = 2$ is special and has a different expression, see [15].

Our conclusion so far is that a classical spacetime, i.e., a Lorentzian manifold, can be viewed as a pair (M, G_F), where G_F is a Feynman propagator of a scalar field. A key difference between describing a spacetime using a pair (M,σ) or a pair (M, G_F) is that, in the former case, the traditional measurement of a geodesic distance requires the use of rods and clocks along the geodesic. In contrast, in the latter case, we replace rods and clocks, which are non-canonical human artifacts, by the naturally occurring fluctuations of a quantum field. The correlations in the quantum fluctuations of a field are sufficiently modulated by the underlying curvature of spacetime to enable the reconstruction of the metric of the spacetime. Similarly, it should be possible to use spinorial and tensorial Feynman propagators, after scalar contractions, to determine the metric.

In practice, the measurement of a field correlator, such as a Feynman propagator, would require, in principle, the detection and counting of quantum field fluctuations, a difficult notion. With present technology, quantum fluctuations of the vacuum of the electromagnetic field can be measured with some accuracy in table-top quantum homodyne detectors [17, 18]. Quantum optical measurements of the correlations of spacelike or timelike separated electromagnetic quantum vacuum fluctuations may become feasible in table-top experiments. In principle, with sufficient accuracy, such types of experiments could pick up gravity-caused modulations of the functional form of a Feynman propagator. From the Feynman propagator, the metric could then be calculated. In principle, therefore, such experiments, if mobile and sufficiently accurate, could be used to map the curvature of spacetime. In a more indirect sense, particle accelerators, such as the LHC, can be interpreted as devices that test Feynman rules and to determine the functional form of Feynman rules, including the Feynman propagators. With sufficient accuracy, a mobile particle accelerator could, therefore, also be used to map the curvature of a spacetime by determining a gravity-modulated Feynman propagator.

To conclude, we arrived at the finding that, instead of using rods and clocks to map the curvature of a spacetime manifold, as Einstein envisaged, we could map the curvature of a spacetime manifold by measuring $G_F(x,x')$ close to its diagonal, i.e., by measuring the local correlations of quantum field fluctuations. This is because the Feynman propagator $G_F(x,x')$ can then be used to recover the traditional metric-based description of a spacetime through Equation (7). Intuitively, this is possible because the strength of the correlations of the quantum field fluctuations that is encoded in the propagator is a proxy for the covariant distance, the correlations being the stronger the smaller the covariant distance.

4. REPLACING THE NOTION OF DISTANCE BY THE NOTION OF CORRELATION

So far, we replaced one set of tools to map a spacetime's curvature with another set of tools to map a spacetime's curvature, namely by replacing rods and clocks with the correlator of quantum field fluctuations. We, thereby, assumed that there exists an underlying Lorentzian spacetime to be mapped. We now return

to the main objective, which is to challenge the validity of the picture of a spacetime-stage that hosts matter-actors.

To this end, we begin by asking what if the reasonable assumption is true that, in nature, there is no spacetime in the exact sense of a Lorentzian manifold? In this case, what we described above as the reconstruction of a spacetime from a Feynman propagator can only be approximate.

This suggests exploring the possibility that the primary structure of nature is not that of a Lorentzian spacetime-stage with matter-actors in the form of quantum fields but that the primary structure of nature consists of abstract correlators. While the abstract correlators would describe all regimes of nature, only in some regime, which may be called the "low energy" regime, the abstract correlators could be approximately represented in the sense that they could be viewed, at least approximately, as arising from quantum field fluctuations on a curved spacetime, as shown in [19].

If the abstract 2-point and multi-point correlation functions are the primary structure, nature would be information-theoretic in nature. The notions of a spacetime and matter would then be secondary in the sense that these notions only emerge in the low energy regime as convenient notions for describing the structure of the approximate mathematical representations of the abstract correlators in terms of Quantum Field Theories (QFTs) on a spacetime.

In the low energy regime, it would be the Feynman rules (to tree level) and, ultimately, the full sums of Feynman graphs of the standard model of particle physics on curved spacetime, which would serve as a good approximate mathematical representation of the abstract correlators. At high energies, the abstract correlators would not possess a representation that makes them appear to arise from the quantum fluctuations of fields on a background Lorentzian spacetime. Not only would there be no notion of spacetime but also no notion of matter belonging to definite species of fields that could live on a spacetime. Instead, the abstract correlators would need to be thought of as mere structures that may be best described information theoretically. In this study, we cannot answer the question of what determines the structure of the abstract correlators, as this question is as hard as asking what determines the dimension of spacetime, the structure of the standard model field content and interactions, and what lies beyond.

Let us now return to the aim of challenging the widely-held picture of a spacetime-stage and matter-actors, in order to perhaps get a glimpse of a possible new picture that could lie beyond. To this end, let us consider how, in this new picture, the derived notions of a spacetime stage and matter actors would be seen as breaking down toward the Planck scale, from the perspective of an experimenter who approaches the Planck scale from low energies: correlators, such as a propagators, should become less and less knowable at high energies and small distances. For example, to measure a correlator, such as a Feynman propagator, with some accuracy requires, in principle, a large number of measurements since statistics needs to be accumulated to obtain a reliable value for a correlator. Repeated measurements can be spaced out in small regions of spacetime, but, as these regions are chosen smaller (speaking in the conventional picture), interactions increase, significant renormalization is needed, and eventually a natural ultraviolet cutoff may arise, limiting the knowability of the statistics of the quantum fluctuations. If so, from the perspective of the traditional picture of a spacetime-stage with matter-actors, the Planck scale would not be a regime of exotic phenomena or of wild quantum fluctuations of spacetime and matter. Instead, the Planck scale might appear as a regime of poor statistics. The statistics of the correlators, or Feynman propagators, would be too poor to even approximately[4] assign a classical metric. From the information-theoretic perspective of the abstract correlators, this phenomenon of inaccessibility of information in the ultraviolet may appear, for example, as a bandlimitation for matter fields, which in turn induces a corresponding "bandlimitation" on the knowability, by means of matter-based measurements, of spacetime curvature (see, e.g., [20–28]).

5. IDENTIFYING THE GEOMETRIC DEGREES OF FREEDOM

We have arrived at a picture in which correlators, such as a Feynman propagator, are primary, with a metric spacetime and quantum fields emerging as derived, approximate concepts that provide a useful language for the "low energy" regime. Let us for now focus on this low energy regime, defined as the regime where both the metric-based and the correlator-based descriptions are valid.

In this low energy regime, we can now identify a problem that persists when transitioning from the metric-based picture to the new abstract correlator-based picture: the problem is that a correlator, such as a Feynman propagator, $G_F(x, y)$, is still a function of arbitrary parameters, much like the metric. The propagator is, therefore, encoding its geometric information highly redundantly. This is because, much like the metric, the affine connection and the Synge function, the functional form of a propagator changes under diffeomorphisms. This leaves us with the task to mod out the diffeomorphism group if we wish to isolate the geometric degrees of freedom, i.e., to identify the Lorentzian structure.

For a first attempt at extracting the diffeomorphism invariant information contained in a Feynman propagator, let us start by recalling that, functional analytically, the Feynman propagator is a right inverse

$$WG_F = \delta \tag{8}$$

of its wave operator, W, such as

$$W := \sqrt{g}(\Box + m^2). \tag{9}$$

The wave operator is a self-adjoint operator which, as we discussed above with Equation (7), inherits all geometric

[4]Crudely, the notion of a classical spacetime arising approximately, in some regime, from abstract correlators could be compared to the notion of a classical path arising approximately, in some regime, from a quantum particle's wave function.

information from the propagator. As a self-adjoint operator, the wave operator possesses a real spectrum, $spec(W)$, and this spectrum at first sight appears to be what we are looking for, namely a set of invariants under the diffeomorphism group. We arrive at Lorentzian spectral geometry (see, e.g., [29–32]), the discipline that asks: to what extent does the spectrum of a wave operator determine a Lorentzian manifold?

In fact, $spec(W)$ is not a large enough set of invariants to identify the Lorentzian manifold. There are two basic reasons. The first reason is that the spectra of the typically hyperbolic wave operators tend to be continuous and therefore particularly information poor. The spectra of hyperbolic wave operators can be made discrete with suitable infrared cutoffs. This necessitates choices of boundaries and boundary conditions. These, however, are strongly affecting the resulting spectra and, thereby, obscuring the extraction of geometric information from the spectra. The second reason for why $spec(W)$ is not a large enough set of invariants to identify a Lorentzian manifold, even if suitably discretized *via* an IR cutoff, is more fundamental. The reason is that $spec(W)$ is a set of invariants under the action of the entire unitary group in the function space, which is a larger group than the diffeomorphism group since it also contains, for example, Fourier transforms. This means that $spec(W)$ can be, and generally is, smaller than the set of invariants under only the diffeomorphism group.

As an aside, let us briefly discuss an approach [24, 33] to overcoming this problem by introducing the tool of infinitesimal spectral geometry. Conventional spectral geometry aims to solve the highly non-linear problem of determining to what extent the spectrum of an operator on a manifold determines the metric of the manifold. Infinitesimal spectral geometry aims to solve the simpler linear problem of determining to what extent an infinitesimal change of the spectrum of an operator on a manifold determines the corresponding infinitesimal change in the metric of the manifold. Infinitesimal changes are then iterated to obtain finite changes of the curvature of the manifold from finite changes of the spectrum, as far as well defined. This approach yields a new perspective for why the set of geometric invariants $spec(W)$ is generally incomplete: Only in two dimensions is the metric essentially scalar. In higher dimensions, perturbations of the metric are truly tensorial and, therefore, cannot be covariantly expanded in the eigenbasis of a scalar wave operator. This suggests, as a remedy, to work with the spectra, not of scalar wave operators but, of wave operators of covariant symmetric 2-tensors, since this will guarantee that any small change in the spacetime metric can be covariantly expanded in the eigenbasis of the wave operator. To this end, Feynman propagators of spin-2 particles that are composites could be used, as gravitons are the only expected non-composite spin-2 particles.

We now propose a new approach to extracting the geometric, diffeomeorphism invariant information from the abstract correlators. To this end, let us retrace the steps below Equation (9). We started with the knowledge that the Feynman propagator G_F and its wave operator W contain the complete information about the metric if given in the position representation. Due to diffeomorphism invariance, they do so in a highly redundant way. We, therefore, considered the spectrum of the wave operator, since it consists of diffeomorphism invariants that carry geometric information, though not the complete set of geometric information. In other words, we observed that, while the wave operator contains all geometric information when given in a position basis, it does not contain the complete geometric information when given in its eigenbasis, i.e., when we only know its spectrum.

This tells us that knowing the wave operator in its eigenbasis (i.e., knowing nothing but its spectrum) and, in addition, also knowing a unitary transformation from that eigenbasis to a position basis is sufficient to calculate the metric. This is because we can then transform the Feynman propagator or wave operator from the eigenbasis of the wave operator into a position basis and, from there, arrive at the metric using Equation (7).

The question is, therefore, if we can find such a unitary transformation on the basis of knowing only the abstract correlators. The answer is yes. To see this, recall that, so far, we have only utilized the information contained in the abstract 2-point correlators, i.e., in the Feynman propagators. The abstract n-point correlators for $n > 2$, which we have not yet used, happen to contain exactly the information that is needed to calculate unitaries that map from the eigenbasis of the wave operator of a propagator to a position bases. The reason is that these multi-point correlators describe the vertices of interactions and these vertices are local. This means that, from the abstract correlators in an arbitrary basis, we can always calculate unitary transformations to position bases, namely by diagonalizing these vertices (as operators from a n-fold tensor product of the space of fields into itself, with n depending on the valence of the vertex). Only in the position basis are the vertices of the Feynman rules of a local quantum field theory diagonal, i.e., only in a position representation are the vertices proportional to products of Dirac deltas. For example, the 3-vertex of $\lambda\phi^4$ theory is usually given in the momentum basis, but it can be expressed in any basis of the space of fields, for example, in a Bargmann Fock basis [19]. In a position basis, and only in a position basis, is the vertex diagonal in the sense that it takes the form

$$V(w,x,y,z) = -i\lambda\delta(w-x)\delta(y-z)\delta(w-z) \qquad (10)$$

which expresses the locality of the interaction.

In conclusion, we have, therefore, arrived at a new method to obtain the metric from correlation functions, namely, from knowledge of a propagator and a vertex of a QFT. Crucially, the new method can be used if the propagator and vertex are given in any arbitrary basis in the function space or also if they are given basis independently. Given the propagator and vertex, the method consists in determining a basis in which the vertex is diagonal (a position basis, therefore), then transforming the propagator into that basis and finally deriving the metric from the propagator. Unlike infinitesimal spectral geometry, the new method, therefore, works straightforwardly for spacetimes of any dimension and signature. As for the assumption of the diagonalizability of the vertex, the vertices of all physical theories are local and therefore diagonalizable (i.e., possess representations as products of Dirac deltas), at least in the low energy regime, which is where we are deriving a metric from the

correlators. The dissolution of the notion of a spacetime manifold as one approaches the Planck scale can then manifest itself mathematically as the non-diagonalizability of the vertex, i.e., in the appearance of non-diagonal terms in the vertex correlators in any basis.

We remark that it can now be seen from a new perspective how, for example at the Planck scale, abstract n-point correlators can fail to possess a representation in terms of a quantum field theory whose interactions are local and which lives on a classical curved spacetime. This happens in regimes where the n-point correlators no longer admit even an approximate diagonalization.

The new approach to extracting the geometric information from n-point correlators can be viewed as a generalization of spectral geometry: conventional spectral geometry studies to what extent the shape of a manifold can be inferred from the spectrum of a wave operator of a free field that lives on the manifold. The new approach is to consider not free but interacting fields on the manifold, or even just one field that is self-interacting, e.g., though a $\hat{\phi}^4$ interaction. This yields non-trivial n-point correlators for $n \geq 2$. Independently of the basis in which these Feynman rules are given, they contain the basis-independent information of (a) the spectrum of the propagator and (b) the changes of basis from the eigenbasis of the propagators to the position bases, defined as those bases in which the vertices are diagonal, i.e., local. Together, these two sets of basis independent and, therefore, diffeomorphism independent information form a complete set of invariants to describe a metric manifold.

It should be interesting to see if, in acoustic spectral geometry, this translates into the ability to hear the shape of a thin curved vibrating object if drumming it weakly as well as strongly enough to invoke non-linear oscillations.

6. OUTLOOK

There are, of course, open questions regarding the picture in which abstract correlators are primary, with the conventional picture of a spacetime stage that hosts matter actors only emerging in certain regimes as useful but approximate representations of the abstract correlators. For example, one may ask whether there are new prospects for deriving the dimensionality of spacetime and, regarding dimensionality, what the relationship to holography could be. In this context, it is worth considering the fact that any first quantized or suitably UV and IR regularized second quantized theory formulated in one number of spatial dimensions can be unitarily mapped into an equivalent first or second quantized theory in any other chosen number of spatial dimensions. The reason is that the Hilbert spaces of first quantized theories with a finite number of degrees of freedom are separable, i.e., they possess countable Hilbert bases. Quantum field theories, after suitable UV and IR cutoffs, also possess only a finite number of degrees of freedom and their Hilbert spaces are, therefore, also separable. All separable infinite-dimensional Hilbert spaces, however, are unitarily equivalent. (We are assuming here that the regularizations are not so drastic that they reduce the dimension of the Hilbert spaces to finite numbers since finite-dimensional Hilbert spaces are unitarily equivalent only if their finite dimensions match.)

For example, using Cantor's diagonal counting, the countable eigenbasis of a 1-dimensional harmonic oscillator can be unitarily mapped into the also countable Hilbert basis of a 2-dimensional harmonic oscillator, or, e.g., into the countable eigenbasis of a hydrogen atom in a three dimensional box. Of course, what is local in one theory will generally not be local in the unitarily equivalent theory. Similarly, the equivalence of a regularized second quantized theory in one number of spatial dimensions to a regularized second quantized (or first quantized!) theory in an arbitrary different number of spatial dimensions is guaranteed, i.e., it is not special *per se*. What can make such an equivalence special is if the two equivalent theories in question are each of interest in their own right.

From the perspective of the picture where abstract correlators are primary, the determination in which regime the correlators can be represented as arising, approximately, as the correlations of quantum fluctuations of fields on a spacetime tells us in effect what dimension of spacetime and what matter content the given abstract correlators describe in some regime. In order to investigate these questions, a technical challenge will be to develop functional analytic methods to describe Feynman rules, for example, those of a scalar $\hat{\phi}^4$ theory, basis independently, for example, in terms of the spectra of wave operators and the unitaries that map the eigenbasis of a wave operators into bases in which the vertices are (essentially) diagonal. This analysis should help identify those functional analytic properties of the vertices, i.e., of the n-point correlators, that determine the regime in which they can be viewed as at least approximately diagonalizable. Knowing those functional analytic properties could help explore possible structures that determine the abstract correlators.

Presumably, the natural language to study such questions about the abstract correlators is information theory. For example, as we briefly discussed, the presence of a natural ultraviolet cutoff could manifest itself in the abstract correlators as a form of bandlimitation, in which case generalized Shannon sampling theory [21–28, 34], which is related to minimum length uncertainty principles [19, 35, 36], could provide useful tools.

Among the many open questions is also how to interpret widely-separated entangled systems in the conventional picture. In the new picture, where spacetime distance is, by definition, inferred from correlations, such systems would appear to be "close" by definition, i.e., they may be considered "close" without needing an appeal, for example, to conventional wormholes [37].

From the new perspective where abstract correlators are primary, it should also be interesting to explore possible links to candidate quantum gravity theories, as shown in e.g., [38, 39] and also in [40] which is based on the Synge function and quantum indefinite causal structures, as shown in e.g., [41], as well as to studies that aim to link the structure of the standard model to algebraic structures and discrete spacetime models e.g., [42, 43]. Of relevance here could also be the in-depth investigations into the relationship between the possible dynamics of matter and the correspondingly possible dynamics of gravity, as shown in [44]. It will be interesting as well to explore possible connections to the physics and formalisms of quantum reference frames and

related studies of notions of distance *via* correlations between observables [45–47].

Finally, the perspective where abstract correlators are primary is philosophically close to various approaches, such as relational quantum mechanics, [48] and, in particular, the approach stated in [49]. The approach in [49] is not concerned with quantum field theory or curved spacetime. However, its central observation could be of interest to the approach proposed here: In [49], Mermin showed that the set of correlators among any chosen complete set of subsystems of a system provides a complete tomography of the state of the system, hence the motivation there to consider correlators as primary. It should be interesting to explore how or to what extent this result can be applied to quantum field theories, although limitations to localizability, as e.g., described by Malament's theorem, make the consideration of localized subsystems difficult in quantum field theory, even in the low energy regime. Worth mentioning are also attempts at describing nature information theoretically based on the idea of zeroth, first, and second quantizing the notion of a binary alternative [50, 51].

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article. Further inquiries can be directed to the author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

The author acknowledges support through a Discovery Grant of the National Science and Engineering Council of Canada (NSERC), a Discovery Project grant of the Australian Research Council (ARC), and a Google Faculty Research Award. The author is grateful for feedback from Barbara Šoda, Jason Pye, Maria Papageorgiou, Flaminia Giacomini, and two referees.

REFERENCES

1. Kempf A. *Oral Presentation at the Workshop Interplay of Quantum Information, Foundations and Gravity, May 28 - June 1, 2018 at the Austrian Academy of Sciences*. Vienna. A video recording of this talk held on June 1, 2018 is at: https://phaidra.univie.ac.at/detail/o:1095799?mode=full. Most of the results of the present paper were first presented in this talk.
2. Synge JL. *Relativity: The General Theory*. Amsterdam: North Holland Publishing (1960). p. 56.
3. Fulling SA., *Aspects of Quantum Field Theory in Curved Space-Time*. Cambridge: London Mathematical Society, LMSST17; Cambridge University Press (1989).
4. Birrell ND, Davies PCW. *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press (1982).
5. Einstein A. *Grundzüge der Relativitätstheorie*. 5th ed. Braunschweig: Vieweg (1981).
6. Fewster CJ, Verch R. The necessity of the Hadamard condition. *Class Quant Gravity*. (2013) 30:235027. doi: 10.1088/0264-9381/30/23/235027
7. Jonsson RH, Aruquipa DQ, Casals M, Kempf A, Martin-Martinez E. Communication through quantum fields near a black hole. *Phys Rev D*. (2020) 101:125005. doi: 10.1103/PhysRevD.101.125005
8. Valentini A. Non-local correlations in quantum electrodynamics. *Phys Lett A*. (1991) 153:321. doi: 10.1016/0375-9601(91)90952-5
9. Reznik B. Entanglement from the vacuum. *Found Phys*. (2003) 33:167. doi: 10.1023/A:1022875910744
10. Reznik B, Retzker A, Silman J. Violating bell's inequalities in vacuum. *Phys Rev A*. (2005) 71:042104. doi: 10.1103/PhysRevA.71.042104
11. Sorkin RD. On the entropy of the vacuum outside a horizon. *arXiv preprint* arXiv:1402.3589. doi: 10.1007/978-94-009-6469-3
12. Srednicki M. Entropy and area. *Phys Rev Lett*. (1993) 71:666. doi: 10.1103/PhysRevLett.71.666
13. Papageorgiopu M, Pye J. Impact of relativity on particle localizability and ground state entanglement. *J Phys A Math Theor*. (2019) 52:375304. doi: 10.1088/1751-8121/ab3593
14. Banks T. Holographic space-time and quantum information. *Front Phys*. (2020) 8:111. doi: 10.3389/fphy.2020.00111
15. Saravani M, Aslanbeigi S, Kempf A. Spacetime curvature in terms of scalar field propagators. *Phys Rev D*. (2016) 93:045026. doi: 10.1103/PhysRevD.93.045026
16. Hawking SW, Ellis GFR. *The Large Scale Structure of Space-Time*. Cambridge: Cambridge University Press (1973).
17. Bachor H-A, Ralph TC. *A Guide to Experiments in Quantum Optics*. 2nd ed. Weinheim: Wiley-VCH Verlag (2004).
18. Walls DF, Milburn GJ. *Quantum Optics*. 2nd ed. Heidelberg: Springer (2004).
19. Kempf A. On quantum field theory with nonzero minimal uncertainties in positions and momenta. *J Math Phys*. (1997) 38:1347. doi: 10.1063/1.531814
20. Shannon CE, Weaver W. *The Mathematical Theory of Communication*. Urbana; Chicago, IL: University of Illinois Press (1963).
21. Kempf A. Fields over unsharp coordinates. *Phys Rev Lett*. (2000) 85:2873. doi: 10.1103/PhysRevLett.85.2873
22. Kempf A. Covariant information-density cutoff in curved space-time. *Phys Rev Lett*. (2004) 92:221301. doi: 10.1103/PhysRevLett.92.221301
23. Kempf A. Information-theoretic natural ultraviolet cutoff for spacetime. *Phys Rev Lett*. (2009) 103:231301. doi: 10.1103/PhysRevLett.103.231301
24. Kempf A. Spacetime could be simultaneously continuous and discrete, in the same way that information can be. *New J Phys*. (2010) 12:115001. doi: 10.1088/1367-2630/12/11/115001
25. Kempf A. Quantum gravity on a quantum computer? *Found Phys*. (2014) 44:472–82. doi: 10.1007/s10701-013-9735-3
26. Pye J, Donnelly W, Kempf A. Locality and entanglement in bandlimited quantum field theory. *Phys Rev D*. (2015) 92:105022. doi: 10.1103/PhysRevD.92.105022
27. Chatwin-Davies A, Kempf A, Martin RTW. Natural covariant planck scale cutoffs and the cosmic microwave background spectrum. *Phys Rev Lett*. (2017) 119:031301. doi: 10.1103/PhysRevLett.119.031301
28. Kempf A. Quantum gravity, information theory and the CMB. *Found Phys*. (2018) 48:1191–203. doi: 10.1007/s10701-018-0163-2
29. Yazdi YK, Letizia M, Kempf A. Lorentzian spectral geometry with causal sets. *Class Quantum Grav*. (2021) 38:015011. doi: 10.1088/1361-6382/abc3f8
30. Kac M. Can one hear the shape of a drum? *Am Math Mon*. (1966) 73:1.
31. Weyl H. *Ueber die Asymptotische Verteilung der Eigenwerte*. Nachrichten der Koeniglichen Gesellschaft der Wissenschaft Goettingen (1911). p. 110-7.
32. Datchev K, Hezari H. *Inverse Problems and Applications: Inside Out II. Mathematical Sciences Research Institute Publications Series Vol. 60*. Cambridge: Cambridge University Press(2012).
33. Aasen D, Bhamre T, Kempf A. Shape from sound: toward new tools for quantum gravity. *Phys Rev Lett*. (2013) 110:121301. doi: 10.1103/PhysRevLett.110.121301
34. Šoda B, Kempf A. in preparation.

35. Kempf A, Mangano G, Mann RB. Hilbert space representation of the minimal length uncertainty relation. *Phys Rev D*. (1995) 52:1108. doi: 10.1103/PhysRevD.52.1108
36. Scardigli F. Generalized uncertainty principle in quantum gravity from micro-black hole gedanken experiment. *Phys Lett B*. (1999) 452:39–44. doi: 10.1016/S0370-2693(99)00167-7
37. Maldacena J, Susskind L. Cool horizons for entangled black holes. *Fortsch Phys*. (2013) 61:781–811. doi: 10.1002/prop.201300020
38. Rovelli C. Loop quantum gravity. *Liv Rev Relat*. (2008) 11:5. doi: 10.12942/lrr-2008-5
39. Loll R. Quantum gravity from causal dynamical triangulations: a review. *Class Quantum Grav*. (2020) 37:013002. doi: 10.1088/1361-6382/ab57c7
40. Jia D. World quantum gravity: an approach based on Synge's world function. *arXiv preprint* arXiv:1909.05322 (2019).
41. Giacomini F, Castro-Ruiz E, Brukner Č. Indefinite causal structures for continuous-variable systems. *New J Phys*. (2016) 18:113026. doi: 10.1088/1367-2630/18/11/113026
42. Furey C. *Standard model physics from an algebra?* (Ph.D. Thesis). University of Waterloo, Waterloo, ON, Canada.
43. Farnsworth S, Boyle L. Rethinking Connes' approach to the standard model of particle physics via non-commutative geometry. *New J Phys*. (2015) 17:023021. doi: 10.1088/1367-2630/17/2/023021
44. Schuller FP. Constructive gravity: foundations and applications. *arXiv preprint* arXiv:2003.09726 (2020).
45. Vanrietvelde A, Hoehn PA, Giacomini F, Castro-Ruiz E. A change of perspective: switching quantum reference frames via a perspective-neutral framework. *Quantum*. (2020) 4:225. doi: 10.22331/q-2020-01-27-225
46. Giacomini F, Castro-Ruiz E, Brukner C. Quantum mechanics and the covariance of physical laws in quantum reference frames. *Nat Commun*. (2019) 10:494. doi: 10.1038/s41467-018-08155-0
47. Castro-Ruiz E, Giacomini F, Belenchia A, Brukner C. Quantum clocks and the temporal localisability of events in the presence of gravitating quantum systems. *Nat Commun*. (2020) 11:1–12. doi: 10.1038/s41467-020-16013-1
48. Rovelli C. Relational qusantum mechanics. *Int J Theor Phys*. (1996) 35:1637–78. doi: 10.1007/BF02302261
49. Mermin ND. What is quantum mechanics trying to tell us? *Am J Phys*. (1998) 66:753–67. doi: 10.1119/1.18955
50. von Weizsäcker CF. *Aufbau der Physik*. München: Carl Hanser Verlag (1985).
51. Lyre H. Quantum theory of Ur-objects as a theory of information. *Int J Theor Phys*. (1995) 34:1541–52. doi: 10.1007/BF00676265

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Astronomy and Space Sciences

PERSPECTIVE
published: 07 June 2021
doi: 10.3389/fspas.2021.563450

Generalized Holographic Principle, Gauge Invariance and the Emergence of Gravity à la Wilczek

Andrea Addazi[1,2], Pisin Chen[3,4,5†], Filippo Fabrocini[6†], Chris Fields[7†], Enrico Greco[8†], Matteo Lulli[9†], Antonino Marcianò[10,11†] and Roman Pasechnik[12†]*

[1]Center for Theoretical Physics, College of Physics Science and Technology, Sichuan University, Chengdu, China, [2]INFN Sezione Roma Tor Vergata, Rome, Italy, [3]Leung Center for Cosmology and Particle Astrophysics, National Taiwan University, Taipei, Taiwan, [4]Department of Physics, National Taiwan University, Taipei, Taiwan, [5]Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory, Stanford University, Stanford, CA, United States, [6]College of Design and Innovation, Tongji University, Shanghai, China, [7]Caunes Minervois, Paris, France, [8]Institut de Chimie Radicalaire, Aix-Marseille University, Marseille, France, [9]Department of Mechanics and Aerospace Engineering, Southern University of Science and Technology, Shenzhen, China, [10]Department of Physics and Center for Field Theory and Particle Physics, Fudan University, Shanghai, China, [11]INFN Laboratori Nazionali di Frascati, Frascati, Italy, [12]Department of Astronomy and Theoretical Physics, Lund University, Sölvegatan, Sweden

We show that a generalized version of the holographic principle can be derived from the Hamiltonian description of information flow within a quantum system that maintains a separable state. We then show that this generalized holographic principle entails a general principle of gauge invariance. When this is realized in an ambient Lorentzian space-time, gauge invariance under the Poincaré group is immediately achieved. We apply this pathway to retrieve the action of gravity. The latter is cast à la Wilczek through a similar formulation derived by MacDowell and Mansouri, which involves the representation theory of the Lie groups SO(3, 2) and SO(4, 1).

Keywords: Wilczek gravity, black hole information loss problem, emergent gravity, gauge invariance, holographic principle

OPEN ACCESS

Edited by:
Mohamed Chabab,
Cadi Ayyad University, Morocco

Reviewed by:
Ignazio Licata,
Institute for Scientific Methodology (ISEM), Italy
Mohammed Daoud,
Ibn Tofail University, Morocco

****Correspondence:***
Antonino Marciano
marciano@fudan.edu.cn

†ORCID:
Pisin Chenc
orcid.org/0000-0001-5251-7210
Filippo Fabrocini
orcid.org/0000-0002-6972-1020
Chris Fields
orcid.org/0000-0002-4812-0744
Enrico Greco
orcid.org/0000-0003-1564-4661
Matteo Lulli
orcid.org/0000-0002-6172-0197
Antonino Marciano
orcid.org/0000-0003-4719-110X
Roman Pasechnik
orcid.org/0000-0003-4231-0149

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Astronomy and Space Sciences

Received: *18 May 2020*
Accepted: *04 January 2021*
Published: *07 June 2021*

Citation:
Addazi A, Chen P, Fabrocini F, Fields C, Greco E, Lulli M, Marcianò A and Pasechnik R (2021) Generalized Holographic Principle, Gauge Invariance and the Emergence of Gravity à la Wilczek. Front. Astron. Space Sci. 8:563450. doi: 10.3389/fspas.2021.563450

1 INTRODUCTION

Almost one hundred years of attempts to quantize gravity suggest that physical perspective may be responsible for this failure (Garay, 1995). While continuing to seek an UV-complete theory of either General Relativity (GR) or one of its possible extensions (Polchinski, 1998; Rovelli, 2004; Modesto, 2012; Modesto and Rachwal, 2014), an alternative option is to look at gravity as an emergent phenomenon (Jacobson, 1995; Barcelo et al., 2005; Van Raamsdonk, 2010; Verlinde, 2011; Swingle and Van Raamsdonk, 2014; Chiang et al., 2016; Oh et al., 2018). Among many possible instantiations of this simple idea stands a paradigm of emergence that aims at recovering gravity via its analogical similarity with Yang-Mills gauge theories. As remarked by Chen-Ning Yang, while electromagnetism is evidently a gauge theory, and the fact that gravity can be seen as such a theory is universally accepted, how this exactly happens to be the case must be still clarified. Notable explorations along these lines have been provided in the past by Weyl (1918), and more recently by MacDowell and Mansouri (1977), and Chamseddine et al. (1977), with subsequent improvements by Stelle and West (1979).

At the same time, we heuristically note that gravity may naturally encode principles of information theory. Such consideration naturally follows pondering that gravity is the field that is involved in the very definition of both masses and spacetime distances, and that specifies the

propagation velocities of point-like particles, and hence of information, through the geodesic equations. Thus it is reasonable to pursue a fundamental theory of gravity from this perspective. Indeed, the underlying graph-structure of information networks is a set of nodes and links—this is reminiscent of the basis of the states in Loop Quantum Gravity (Rovelli, 2004).

There have been huge achievements in the direction of a quantum-information based theory of gravity, with several different attempts developed so far—see e.g. Faulkner et al. (2014). More generally, deep links between quantum information theory and an "emergent" quantum theory of observable physical systems have been developed by many studies (Chiribella et al., 2016; Hamma and Markopoulou, 2011; Flammia et al., 2009). It is not within the present scope to summarize this vast literature. Instead, we focus on a specific alternative approach: we show that when the holographic principle is reformulated from a semi-classical to a fully general, quantum-theoretic principle, gravity emerges as a gauge theory along the lines of the gauge formulation of gravity, as proposed by Wilczek (1998).

We start by showing in **Section 2** that a generalized holographic principle (GHP) characterizes information transfer within any finite quantum system in a separable state. The HP is recovered from this more general, purely-quantum principle by requiring covariance. We then show in **Section 3** that compliance with the GHP entails gauge invariance under the Poincaré group in an ambient Lorentzian space-time. Hence the gauge principle has purely quantum-theoretic roots and characterizes all finite systems in separable states. We use this to retrieve the action of gravity in **Section 4**. In **Section 5**, we provide, as an example, an emergent theory of gravity, a theory of Yang-Mills gauge fields and Higgs pentaplets that is cast à la Wilczek. This is a formulation similar to a previous one envisaged by MacDowell and Mansouri, which involves the representation theory of the Lie group $SO(4,1)$, but without explicit symmetry breaking. We finally summarize some conclusions in **Section 6**, and suggest that the AdS/CFT and dS/CFT correspondences may naturally arise within this framework.

2 GENERALIZED HOLOGRAPHIC PRINCIPLE FOR FINITE QUANTUM SYSTEMS

2.1 Historical Remarks on the Genesis of the Holographic Principle

Probably the most direct way to summarize the Holographic Principle (HP) is via its original statement by 't Hooft (1993):

> "given any closed surface, we can represent all that happens inside it by degrees of freedom on this surface itself."

The path that led to the formulation of the HP can be traced from the Bekenstein's area law (Bekenstein, 2004) for black holes (BH),

$$S = \frac{A}{4}, \tag{2.1}$$

where S denotes the thermodynamic entropy of a BH and A its horizon area in Planck units. Bekenstein conjectured the existence of an upper bound, S itself, to the entropy of any physical system contained within a bounded volume:

> "the entropy contained in any spatial region will not exceed the area of the region's boundary."

Historically, this conjecture was first instantiated by Susskind (1995), who implemented a mapping from volume to surface degrees of freedom for a general closed system. This was based on the assumption that all light rays that are normal to any element within the volume are also normal to the surface. Bousso (2002) then showed that it is actually covariance that induces the holographic limit on information transfer by light; he further provided several counterexamples showing the failure of a straightforward interpretation of the HP as a spacelike entropy bound. Instead, Bousso formulated a covariant entropy bound:

$$S(L(\Sigma)) \leq \frac{A(\Sigma)}{4}, \tag{2.2}$$

with $A(\Sigma)$ denoting the area in Planck units of a [typically but not necessarily (Bousso, 2002)] closed surface Σ, and $L(\Sigma)$ any light-sheet of Σ, defined as any collection of converging light rays that propagate from Σ toward some focal point away from Σ. The bound **(2.2)** then refers to the entropy of the light-sheet $L(\Sigma)$. This covariant formulation of the HP holds for the light-sheets of any surface Σ. BH emerge as special cases, for which the equality in **(2.2)** holds.

We note that both **(2.1)** and **(2.2)** are semiclassical. The limits on the entropy S that they impose are "quantum" only in their reliance on Planck units and hence a finite value of $\hbar$. The entropy itself is classical and of statistical origin, but the finite value of $\hbar$ restricts this thermodynamic entropy within the volume enclosed by Σ. In the context of general relativity (GR), Σ is a continuous classical manifold enclosing a continuous classical volume characterized by a real-valued metric. As 't Hooft (1993) points out, the HP renders $S(L(\Sigma))$ independent of the metric inside Σ:

> "The inside metric could be so much curved that an entire universe could be squeezed inside our closed surface, regardless how small it is. Now we see that this possibility will not add to the number of allowed states at all."

It bears emphasis that "allowed states" in this context are thermodynamic states, i.e. states that can be counted by measuring energy transfer between the system and its external environment. As made fully explicit by Rovelli in the case of BH (Rovelli, 2017; Rovelli, 2019), states that are effectively isolated (e.g. isolated for some time interval much larger than relevant interaction times) from the external environment do not contribute to $S(L(\Sigma))$.

While the demonstration by Maldacena (1998) of a formal duality acting as an equivalence, at the level of the encoded information, between string quantum gravity on d-dimensional anti-de Sitter (AdS) spacetime and conformal quantum field theory (CFT) on its $d-1$-dimensional boundary has made the HP a centerpiece of quantum gravity research, its physical motivation remains that of 't Hooft's conjecture, namely the inaccessibility of the BH interior summarized by the Bekenstein area law **(2.1)**. The HP is conjectured to be fully general, although it is quite mysterious why this should be the case. As Bousso (2002) remarks, the HP retains a counterintuitive meaning:

> "an apparent law of physics that stands by itself, both uncontradicted and unexplained by existing theories that may still prove incorrect or merely accidental, signifying no deeper origin."

Our goal in the next section is to place the HP on a much deeper intuitive footing, by generalizing it from a semi-classical to a fully quantum principle, one that is entirely independent of geometric considerations.

2.2 Information Transfer in Finite, Separable Systems

Let $\mathbf{S} = AB$ be a closed quantum system characterized by a Hilbert space $\mathcal{H}_\mathbf{S}$ of finite dimension $dim(\mathcal{H}_\mathbf{S})$, and suppose that over some sufficiently long time interval $\overline{\tau}$, $\mathbf{S}$ maintains a separable state, i.e. $|\mathbf{S}\rangle = |AB\rangle = |A\rangle|B\rangle$ for all $t \leq \overline{\tau}$, where t is a time parameter characterizing $\mathbf{S}$. We can then write a Hamiltonian:

$$H_\mathbf{S} = H_A + H_B + H_{AB} \tag{2.3}$$

where H_{AB} is the A-B interaction. Separability allows $H_{AB} = 0$ but requires $H_A, H_B \neq 0$.

We now assume $H_{AB} \neq 0$ and choose bases for A and B such that, for all $t \leq \overline{\tau}$:

$$H_{AB}(t) = \beta^k k_B T^k \sum_i \alpha_i^k(t) M_i^k, \tag{2.4}$$

where $k = A$ or B, $i = 1 \ldots N$ for finite $N \leq dim(\mathcal{H}_S)$, the $\alpha_i^k(t)$ are real functions with codomains $[0,1]$ such that:

$$\sum_i \int_{\Delta t} dt\, \alpha_i^k(t) = \Delta t \tag{2.5}$$

for every finite Δt, k_B is Boltzmann's constant, T^k is k's temperature, $\beta^k \geq \ln 2$ is an inverse measure of k's average per-bit thermodynamic efficiency[1] that depends on the internal dynamics H_k, and the M_i^k are Hermitian operators with binary eigenvalues. Given separability, we can interpret these M_i^k as "measurement" operators that each transfer 1 bit between A and B. Here the condition $\beta^k \geq \ln 2$ assures compliance with Landauer's principle (Landauer, 1961): each bit transferred from A to B by the action of some operator M_i^A is paid for by the transfer of an energy $\beta^B k_B T^B$ from B to A and vice-versa. "Irreversible recording" of the transferred bits by A and B corresponds[2] to state changes:

$$|A\rangle|_t \to |A\rangle|_{t+\Delta t} \quad \text{and} \quad |B\rangle|_t \to |B\rangle|_{t+\Delta t} \tag{2.6}$$

that maintain the separability of $\mathbf{S}$. Given **(2.5)**, the action required for k to transfer N bits in time Δt is:

$$\int_{\Delta t} dt\,(\imath\hbar)\ln\mathcal{P}(t) = N\beta^k k_B T^k \Delta t \tag{2.7}$$

where $\mathcal{P}(t) = \exp(-(\imath/\hbar)H_{AB}t)$. Informational symmetry clearly requires $\beta^A T^A = \beta^B T^B$ during any finite Δt.

Let us now consider an interval $\tau \ll \overline{\tau}$ during which A and B exchange exactly N bits. In any such interval, the thermodynamic entropy $S(B)|_\tau$ measured by A is clearly N bits; the entropy $S(A)|_\tau$ measured by B is similarly N bits. Coarse-graining time to an interval $\Delta t = n\tau \ll \overline{\tau}$ to allow n N-bit measurements, both measured entropies remain N bits. Hence we have:

Theorem 1. Given any finite-dimensional quantum system $\mathbf{S} = AB$ that maintains a separable state $|AB\rangle = |A\rangle|B\rangle$ for $t \leq \overline{\tau}$, the information $S(B)$ obtainable by A during any finite interval $\Delta t \ll \overline{\tau}$ is independent of H_B.

Proof. The information $S(B)$ obtainable by A during any finite interval $\Delta t \ll \overline{\tau}$ is just the information transferred by H_{AB}, which is specified entirely independently of H_B. Indeed H_B and hence $H_\mathbf{S}$ can be varied arbitrarily, provided that B has sufficient degrees of freedom to maintain the separability of $\mathbf{S}$, without affecting H_{AB}.

Note that if the assumption of separability is dropped and two subsystems cannot be distinguished, **Eq. 2.6** fails, the von Neumann entropy of $|AB\rangle$ remains zero, and no information is transferred by H_{AB}.

2.3 The HP Is a Special Case of the Generalized Holographic Principle

Theorem 1 places a principled restriction on information transfer *within* any separable quantum system; as noted above, the notion of information transfer within a non-separable (i.e. entangled) quantum system is meaningless. The HP is a principled restriction on information transfer *within* a semiclassical system that is separable by definition. Hence the two should be related. This relation can be made explicit by stating:

> Generalized Holographic Principle (GHP): Given any finite-dimensional quantum system $\mathbf{S} = AB$ meeting the conditions of Theorem 1, the thermodynamic entropies of A and B over a coarse-grained time, over which A and B only interact through **Eq. 2.4**, are $S(B) = S(A) = N$ bits, where N is the number of operators in the representation **(2.4)** of H_{AB}.

We note that this GHP is formulated entirely independently of geometric assumptions; in particular, it is prior to any assumption of general covariance.

[1]Here the efficiency relates to the thermodynamic transformations triggering the exchange of information bits among the two subsystems A and B.

[2]Notice that irreversibility is connected to the efficiency bound $1/\beta^k < 1$.

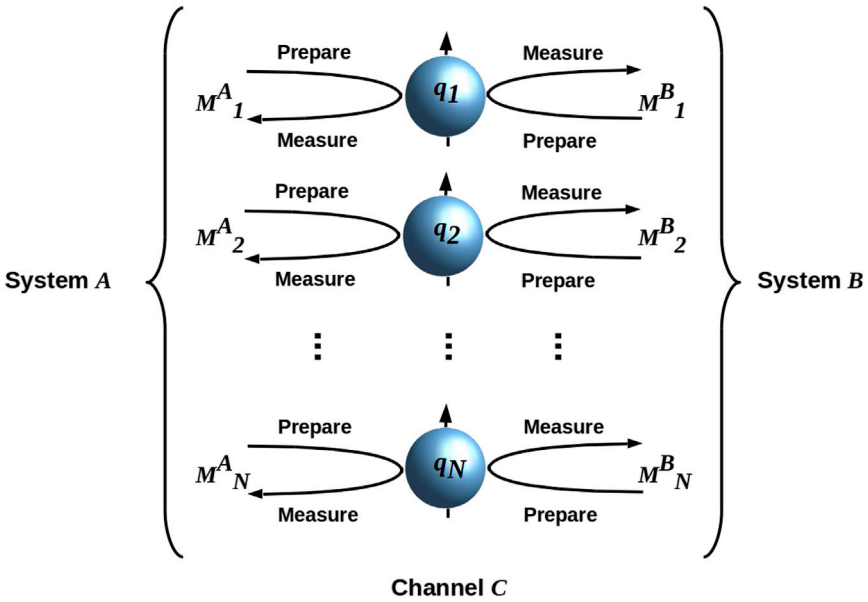

FIGURE 1 | Systems *A* and *B* exchange bits via an ancillary array of non-interacting qubits. Bit values are preserved if a quantum reference frame (here, a *z* axis) is shared *a priori*. Adapted from Fields and Marcianò (2020); CC-BY license.

To make the physical meaning of the GHP clear, let us consider a specific example. Suppose A and B interact by alternately preparing and measuring the states of N shared, non-interacting qubits as shown in **Figure 1**. We can consider that, in a time interval τ, A prepares the N qubits in her choice of basis, i.e. using her M_i^A and then B makes measurements in his choice of basis, i.e. using his M_j^B. In the next interval τ', B prepares and A measures, and so forth. The prepared and measured bit values will be preserved, i.e. A and B will employ the same "language," only if they share a basis, in this case a z axis, which functions as a shared quantum reference frame (QRF) (Bartlett et al., 2007; Fields and Marcianò, 2019). Clearly, however, $S(B) = S(A) = N$ bits in every interval of length at least 2τ, independently of whether A and B share a QRF.

The qubit-mediated interaction shown in **Figure 1** still makes no geometric assumptions. If we now imagine, however, that the array of qubits is embedded at maximal density in an ancillary real 2-dimensional surface Σ, and further require that the bit values generated by the actions of the M_i^A (respectively, M_j^B) must be transferred to a distant point within A (respectively, B) by photons (or any other quantum carrier consistent with the local symmetry/invariance that is present), **Eqs. (2.1)** and **(2.2)**, i.e. the usual covariant HP, results by the reasoning of Bekenstein (2004), Bousso (2002). The surface Σ can, in this case, naturally be interpreted as a "boundary" between A and B at which they interact. The self-interactions H_A and H_B are then naturally interpreted as characterizing the "bulk" of A and B, respectively.

We note that the GHP provides, when H_{AB} is assumed to act across an A-B boundary, an immediate and intuitive explanation of the decoherence of B relative to A and vice-versa (Fields and Marcianò, 2019; Fields, 2019). Hence the GHP provides a natural account of the "emergence of classicality" within separable quantum systems: if $|AB\rangle$ is separable as $|A\rangle|B\rangle$, "classicality" characterizes the bit values "encoded on" the A-B boundary, i.e. the boundary at which H_{AB} acts. There are, in other words, no classical systems, just classical information.

Both Theorem 1 and the GHP above are formulated for fixed N. Generalizing to the case of N varying slowly, i.e. remaining piecewise constant in time for intervals $\tau \ll \Delta t \ll \overline{\tau}$, is straightforward. Therefore, only the constant N case is needed in what follows.

3 POINCARÉ SYMMETRIES AND GAUGE INVARIANCE

3.1 The Generalized Holographic Principle Requires Gauge Invariance for Finite, Separable Systems

Theorem 1 and hence the GHP restricts access to information, and so states an invariance: the information $S(B)$ is invariant under changes in H_B (and vice-versa), provided H_{AB} remains fixed and separability is maintained. Gauge invariance for the "bulk" Hamiltonians H_A and H_B clearly follows.

Theorem 2. In any $\mathbf{S} = AB$ compliant with the GHP, the bulk interactions H_A and H_B are gauge invariant.

Proof. The situation is completely symmetrical, so considering either H_A or H_B alone is sufficient. Gauge symmetry for H_A can only fail if a local coordinate change, i.e. a local change in basis vectors for $\mathcal{H}_A$, is observable, i.e. has an effect on H_{AB}. Any such effect is ruled out by Theorem 1, which is satisfied by all systems compliant with the GHP.

Note that gauge invariance here depends explictly on separability: if $|AB\rangle \neq |A\rangle|B\rangle$, i.e. **S** is in an entangled state, the notion of a "bulk" Hamiltonian H_A (or H_B) is meaningless.

Theorem 2, like Theorem 1 and the GHP, involves no assumptions about geometry. We introduce these below, with QED as an initial example.

3.2 QED and the Consequences of the Generalized Holographic Principle

As a specific example, consider a finite system AB, with A comprising photons described by the usual electromagnetic vector field $A_\mu(x)$ and B comprising fermionic particles, e.g. electrons described by the Dirac field $\psi(x)$, with x a real space-time coordinate. Clearly A and B only interact via a Hamiltonian H_{AB}. The numbers of photons and electrons can be arbitrarily large, so the usual approximation of infinite degrees of freedom can be adopted for simplicity with no physical (i.e. observable) consequences.

We can make the presentation more precise at the mathematical level to illustrate the independence of observable results from coordinate (i.e. basis) transformations, even in the presence of the ancillary space-time geometry with points labelled by x. The local gauge freedom for the choice of the vector field A_μ, which generalizes for quantum fields, i.e. systems with infinite degrees of freedom embedded in an ancillary space x, the invariance with respect to the choice of basis we discussed above, can be defined as in:

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \partial_\mu \lambda(x), \tag{3.1}$$

where $\lambda(x)$ is a scalar function that is continuous with its first derivatives. Stepping out of the redundancy contained in **(3.1)**, and denoting the components of the vector fields as $A_\mu = \{\phi, \mathbf{A}\}$, the electric and the magnetic fields, invariant under **(3.1)**, can be defined respectively as $\mathbf{E} = \dot{\mathbf{A}} - \nabla\phi$ and $\mathbf{B} = \nabla \wedge \mathbf{A}$.

The quantization of the theory can be achieved following the standard path integral procedure. The partition function for the system A, namely the $U(1)$ gauge sector, casts:

$$\mathcal{Z}_A[A] = \int \mathcal{D}A_\mu \, e^{i\mathcal{S}(A)} \tag{3.2}$$

in which $\mathcal{D}A_\mu$ denotes the path integral measure over the copies of the gauge field, while $\mathcal{S}(A)$ denotes the classical action. The expectation value in the path integral of the theory of any functional observable $\mathcal{O}[A]$ is invariant under the gauge transformation **(3.1)**. It is straightforward to show this fundamental property by comparing the expectation value of $\mathcal{O}[A]$ for different choices of the gauge fixing. On the other hand, different choices of the gauge fixing correspond to different choices of (local) observers, namely bases. But within these circumstances, the GHP implies that the path integral of any $\mathcal{O}[A]$ must be gauge invariant: choosing an observer is choosing a measurement basis, i.e. choosing a set of operators M_i^k in **(2.4)**, a choice that is independent of A by Theorem 1. In simpler words, the *GHP implies gauge invariance*.

The redundancy due to the gauge transformations, i.e. choices of M_i^k, can be factored out by fixing the gauge functional G, and then imposing gauge invariance. This is achieved by inserting in the path integral **(3.2)** the resolution of the identity:

$$1 = \int \mathcal{D}\lambda \delta\big(G\big(A^\lambda\big)\big) \left|\det \frac{\delta G\big(A^\lambda\big)}{\delta\lambda(x)}\right| \tag{3.3}$$

where as customary $A^\lambda_\mu(x) = A_\mu(x) + \partial_\mu\lambda(x)$. The simplest choice of gauge functional is provided by the Lorentz functional $G(A) = \partial_\mu A^\mu$, which implements the gauge invariance of the path integral. Notice that under gauge transformations, the Lorentz functional transforms as:

$$G(A) = \partial_\mu A^\mu \to G\big(A^\lambda\big) = \partial_\mu A^\mu + \Box\lambda\,. \tag{3.4}$$

Having all set up, we can easily show the invariance of the path-integral:

$$\begin{aligned}
\mathcal{Z}_{\mathbf{A}}[A] &= \mathcal{N} \int \mathcal{D}A_\mu \, e^{i\mathcal{S}_{\mathrm{A}}(A)} \int \mathcal{D}\lambda\delta\big(G\big(A^\lambda\big)\big) \left|\det\frac{\delta G\big(A^\lambda\big)}{\delta\lambda(x)}\right| \\
&= \mathcal{N}\left|\det\Box\right| \int \mathcal{D}A_\mu \, e^{i\mathcal{S}_{\mathrm{A}}(A)} \int \mathcal{D}\lambda\delta\big(G\big(A^\lambda\big)\big) \\
&= \mathcal{N}' \int \mathcal{D}\lambda\,\mathcal{D}A_\mu \, e^{i\mathcal{S}_{\mathrm{A}}(A)}\,\delta(G(A)) \\
&= \mathcal{N}'' \int \mathcal{D}A_\mu \, e^{i\mathcal{S}_{\mathrm{A}}(A)}\,\delta(G(A)),
\end{aligned} \tag{3.5}$$

where the normalization functions $\mathcal{N}$, $\mathcal{N}'$, and $\mathcal{N}''$ are not relevant, and have been safely disregarded.

The perspective provided by the GHP allows us to place a novel physical interpretation on the action of $G(A)$ in removing gauge redundancy, one that points toward a deeper understanding of the role of the spatial coordinate x in QFT. As noted above, from a GHP perspective, gauge redundancy is the redundancy in the choice of measurement operators M_i^k. This can equally be interpreted as redundancy in the choice of observers k. But k, in this case, is just a quantum system X that can be coupled to the quantum field A while maintaining a separable joint state $|AX\rangle$. The action of $G(A)$ renders these different observers redundant, effectively removing the dependence of (observations of) A on X. Hence we can now see what X is doing in QFT: it is enforcing separability. This is indeed an insight of Einstein (1948):

> Further, it appears to be essential for this arrangement of the things introduced in physics that, at a specific time, these things claim an existence independent of one another, insofar as these things "lie in different parts of space."

"Claiming an existence independent of one another" obviously requires separability.

It is well known that the gauge condition can be cast in a more general form, employing an arbitrary function f. In this latter case, the gauge functional:

$$G_f(A) = \partial_\mu A^\mu - f \tag{3.6}$$

actually introduces a family of gauge-fixing terms. The independence of the physical observables on the gauge fixing is recovered through a process of average that is realized by integrating over f the gauge fixing terms weighted with the factor $\exp\left(-\frac{\imath}{2\xi}\int d^4x f^2(x)\right)$, where ξ is a positive parameter. The path integral in **(3.5)** then recasts as:

$$\begin{aligned}\mathcal{Z}_{\mathbf{A}}[A] &= \mathcal{N}\int \mathcal{D}f\,\mathcal{D}A_\mu\, e^{\imath\mathcal{S}_{\mathrm{A}}(A)-\frac{\imath}{2\xi}\int d^4x f^2(x)} \\ &= \mathcal{N}\int \mathcal{D}A_\mu\, e^{\imath\mathcal{S}_{\mathrm{A}}(A)-\frac{\imath}{2\xi}\int d^4x\left(\partial_\mu A^\mu(x)\right)^2}\end{aligned} \tag{3.7}$$

This invariance of the partition function under different choices of the gauge fixing condition, i.e. different choices of f, percolates into the gauge invariance of the expectation value of any observable $\mathcal{O}$. This manifests immediately, as one can recognize from the easy passages

$$\begin{aligned}\langle\mathcal{O}[A]\rangle_f &= \mathcal{N}\int \mathcal{D}f\,\mathcal{D}A_\mu\, e^{\imath\mathcal{S}_{\mathrm{A}}(A)-\frac{\imath}{2\xi}\int d^4x f^2(x)}\,\mathcal{O}[A] \\ &= \mathcal{N}\int \mathcal{D}A_\mu\, e^{\imath\mathcal{S}_{\mathrm{A}}(A)-\frac{\imath}{2\xi}\int d^4x\left(\partial_\mu A^\mu(x)\right)^2}\,\mathcal{O}[A] \\ &= \mathcal{N}\int \mathcal{D}g\,\mathcal{D}A_\mu\, e^{\imath\mathcal{S}_{\mathrm{A}}(A)-\frac{\imath}{2\xi}\int d^4x g^2(x)}\,\mathcal{O}[A] \\ &= \langle\mathcal{O}[A]\rangle_g\end{aligned} \tag{3.8}$$

where f and g are two different gauge-fixings.

We may take into account now the other interaction partner in QED, the system B composed by Dirac matter fields, for simplicity electrons. The path integral formulation of the system, composed by only one fermionic species ψ, then casts:

$$\mathcal{Z}_B[\psi,\overline{\psi}] = \int \mathcal{D}\psi\,\mathcal{D}\overline{\psi}\, e^{\imath\mathcal{S}_{\mathrm{D}}(\psi,\overline{\psi})} \tag{3.9}$$

where

$$\mathcal{S}_{\mathrm{D}}(\psi,\overline{\psi}) = \int d^4x\,\overline{\psi}\left(\imath\gamma^\mu\partial_\mu - m\right)\psi \tag{3.10}$$

The observable quantities $\mathcal{O}[\psi,\overline{\psi}]$ are bilinear operators in the fermionic fields ψ and $\overline{\psi}$, which can be generally cast as $\mathcal{O}[\psi,\overline{\psi}] = \overline{\psi}\,O(\Gamma_I)\psi$, where the matrix O, with suppressed spinorial indices, depends on the elements of the Clifford algebra Γ_I, with $I = 1\ldots 16$.

As previously done for the bosonic system A, also for the system B we can introduce a local gauge transformation having the meaning of a transformation among observers. Of course, this transformation cannot change the values of $\mathcal{O}$, which lead us to state the necessity of the symmetry

$$\psi(x)\to\psi'(x) = e^{\imath q\lambda(x)}\psi(x),\quad \overline{\psi}(x)\to\overline{\psi}'(x) = e^{-\imath q\lambda(x)}\overline{\psi}(x), \tag{3.11}$$

where q stands for the charge parameter. This is a $U(1)$ transformation, the generator of which commutes with the matrix O, ensuring the gauge invariance of any observable $\mathcal{O}$ under **(3.11)**. By the Noether theorem, selecting $\lambda(x) = \alpha\in\mathbb{R}$ to individuate an infinitesimal global transformation, the conserved charge is easily recovered

$$Q = \int_\Sigma d^3x\,\psi^\dagger\psi,$$

where Σ is a spatial hypersurface. This allows to cast $U(1)$ transformations acting on the Hilbert space of the theory as $U = e^{\imath\alpha Q}$.

3.3 Extension to Gravity and Local Lorentz Invariance

So far we have first considered generic quantum systems with finite number of degrees of freedom, and stated the GHP within these simplified but completely general contexts, which do not necessarily require geometric concepts. In this sense, these notions shall be considered as pre-geometric. We have then extended our focus to continuous systems with an infinite number of degrees of freedom, focusing specifically on the paradigmatic example of QED, which is embedded on a flat Minkowski space-time. This embedding requires the addition of ancillary coordinates x into the description of the system, which are necessary to specify its evolution and fully capture the dynamics as it is observed by spatially-separated observers.

Let us now include gravity in this construction, extending the arguments previously exposed. Our joint system $\mathbf{S} = AB$ shall be now composed by the gravitational degrees of freedom, described by the gravitational field $g_{\mu\nu}(x)$, the configurational space of which constitutes the system A, and by the matter degrees of freedom, the fields[3] $\phi^{\aleph}(x)$, the configurational space of which individuates the system B. The GHP then imposes the gauge invariance of the gravitational field, once the ancillary spatial coordinates have been introduced, in exactly the same way as discussed above. The role of the spatial coordinates is, as in the case of QED, labeling the manifold in which separable physical (in this case, matter) systems, e.g. observers, that interact with the field A are embedded. Symmetries fully depend in this picture on the signature of the embedding space, which we assume to be Lorentzian, so as to distinguish among time [required already by **(2.4)**] and space (ancillary) coordinates. Thus the emergent symmetry will impose the invariance under local $SO(3,1)$ transformations: indeed, the underlying space structure we are considering has Lorentzian signature. This is also consistent with the fact that the tangent space to each point of the manifold is flat and Minkowski, and thus the whole construction specifies how the invariance under supertranslations in time and space emerges in this framework.

Besides local $SO(3,1)$ Lorentz symmetries, the theory of gravity also encodes symmetries under diffeomorphism. It is customary to deal with these latter in the Hamiltonian formulation of the theory. This requires considering Lorentzian manifolds $\mathcal{M}_4$ that are diffeomorphic to $\mathbb{R}\times\Sigma$. This property enables a slicing of $\mathcal{M}_4$ into space-like hypersurfaces Σ at instants of time $t\in\mathbb{R}$. This slicing is arbitrary, since there exist several possible choices to pick a diffeomorphism $\phi:\mathcal{M}_4\to\mathbb{R}\times\Sigma$. Thus different time

[3]Here we denote with $\phi^{\aleph}(x)$ any possible scalar, vectorial or spinorial matter fields.

coordinates τ can be defined on $\mathcal{M}_4$, as the pullback of t that is realized by ϕ, or in the mathematical jargon $\tau = \phi * t$. This corresponds to different clocks for different observers (in the language of §2.2, different "tick" intervals τ), which nevertheless must not affect the definition of the physical observables, which are gauge invariant and diffeomorphic invariant. The theory is then recast on Cauchy surfaces, on which the gravitational field is captured by the restriction of the gravitational field to the three-metric on the slice Σ, namely ${}^3g$, and to its "time" derivative, namely the extrinsic curvature K. These variables form the Cauchy data of the problem, and open the pathway to access the meaning of the ten components of the Einstein equations. Indeed, four of the Einstein equations turn out to be constraint equations that the Cauchy data must satisfy, while the other six are evolutionary equations that dictate the dynamics in time of the three-metric.

A time-like unit vector n must be then defined that is orthogonal to any tangent vector v on Σ. Considering the metric two-form g on $\mathcal{M}_4$, these two requirements amount to write in formulas that $g(n,n) = -1$ and $g(n,v) = 0$. The direction of the unit vector n is then defined to point towards the future. A derivative of any generic vector v on Σ can be then defined along any generic direction individuated by a vector u on Σ. This is simply attained by projecting on Σ and removing the component along the normal direction, i.e. $\nabla_u v = -g(\nabla_u v, n)n + [\nabla_u v + g(\nabla_u v, n)]$. The first term identifies the extrinsic curvature, $K(u,v)n = -g(\nabla_u v, n)n$, which measures the bending of the surface Σ in the ambient manifold $\mathcal{M}_4$, by quantifying the failure of a generic vector of $\mathcal{M}_4$ to be still tangent to Σ after we parallel translate it using the Levi–Civita connection on the ambient space $\mathcal{M}_4$. While considering the component tangential on Σ, we can write ${}^3\nabla_u v = \nabla_u v + g(\nabla_u v, n)$, since it introduces the Levi–Civita connection on Σ that is associated to the three metric ${}^3g$. One can show that this is a connection, and satisfies the Leibnitz rule.

Given this framework, denoted as ADM (Arnowitt et al., 1959) in the literature, we can now introduce a time coordinate $\tau = \phi^* t$ on $\mathcal{M}_4$, which individuates a generic foliation $\{\tau = s\}$. The vector field ∂_τ on $\mathcal{M}_4$ then admits the generic decomposition along a tangential direction to Σ and its normal, respectively individuated by the lapse function N and the shift vector $\mathbf{N}$, i.e. $\partial_\tau N n + \mathbf{N}$. We can finally move to consider the Einstein theory of gravity and its Hamiltonian structure, which is a purely constrained system. We can first move from the Einstein-Hilbert action, which we review in **Section 4**, and cast it in terms of the ADM variables, using the three-metric ${}^3g$, the Levi–Civita connection on Σ, namely ${}^3\Gamma$, the associated Riemann tensor ${}^3R^a_{\mathrm{bcd}}$ and the Ricci scalar ${}^3R$ on Σ. Then the Lagrangian of the Einstein–Hilbert action reads, modulo a boundary term:

$$\mathcal{L} = \sqrt{{}^3g}\, N\left[{}^3R + \mathrm{tr}\left(K^2\right) - (\mathrm{tr}K)^2\right] \tag{3.12}$$

which allows to define the symplectic structure of the system, namely:

$$q_{ij} = {}^3g_{ij} \quad p^{ij} = \frac{\delta\mathcal{L}}{\delta\dot{q}_{ij}} = \sqrt{q}\left[K^{ij} - (\mathrm{tr}K)q^{ij}\right] \tag{3.13}$$

with vanishing momenta conjugated to $\mathbf{N}$ and N, respectively $\mathbf{P} = 0$ and $P = 0$. In **Eq. 3.13** the extrinsic curvature is expressed in terms of the covariant derivatives on Σ, namely ${}^3\nabla$, and the ADM variables, as:

$$K_{ij} = \frac{1}{2N}\left(\dot{q}_{ij} - {}^3\nabla_i N_j - {}^3\nabla_j N_i\right). \tag{3.14}$$

The Hamiltonian density of the gravitational system, which can be calculated by the usual Legendre transform $\mathcal{H}(q_{ij}, p^{ij}) = p^{ij}\dot{q}_{ij} - \mathcal{L}$, finally provides the Hamiltonian of the system $H = \int_\Sigma \mathcal{H} d^3x$. This latter can be recognized to be a totally constrained system:

$$\mathcal{H} = \sqrt{q}\left(N\mathcal{C} + N^i\mathcal{C}_i\right), \tag{3.15}$$

with

$$\mathcal{C} = -{}^3R + q^{-1}\left(\mathrm{tr}\left(p^2\right) - \left(\mathrm{tr}\,p\right)^2\right), \quad \mathcal{C}_i = -2\,{}^3\nabla^j\left(q^{-1/2}p_{ij}\right). \tag{3.16}$$

For simplicity, we assumed the hyperspace Σ to be compact, so as to neglect the contribution otherwise provided by total divergences.

The first term of the Hamiltonian represents the Hamiltonian constraint, which generalizes time reparametrization, while the second term is the space-diffeomorphism constraint, respectively:

$$C(N) = \int_\Sigma N\mathcal{C}\sqrt{q}d^3x, \quad C(\mathbf{N}) = \int_\Sigma N^i\mathcal{C}_i\sqrt{q}d^3x. \tag{3.17}$$

Involving the continuous version of the Poisson brackets for the phase-space variables of the system, namely:

$$\{f, g\} = \int_\Sigma \left\{\frac{\partial f}{\partial p_{ij}(x)}\frac{\partial g}{\partial q^{ij}(x)} - \frac{\partial f}{\partial q^{ij}(x)}\frac{\partial g}{\partial p_{ij}(x)}\right\}\sqrt{q}\, d^3x,$$

we may recover the algebra of constraints for the gravitational system, known as Dirac algebra, namely:

$$\{C(\mathbf{N}),\, C(\mathbf{N}')\} = C([\mathbf{N}, \mathbf{N}']), \tag{3.18}$$

$$\{C(\mathbf{N}),\, C(N)\} = C(\mathbf{N}N'), \tag{3.19}$$

$$\{C(N),\, C(N')\} = C\left[\left(N\partial^i N' - N'\partial^i N\right)\partial_i\right]. \tag{3.20}$$

The scalar and the vector constraints entering the total Hamiltonian can be cast in terms of the Einstein tensor components, contracted with the normal n^μ to the hypersurface Σ, i.e.:

$$\mathcal{C} = -2G_{\mu\nu}n^\mu n^\nu, \quad \mathcal{C}_i = -2G_{\mu i}n^\mu,$$

while the spatial components G_{ij} source the Hamilton equation of the phase-space variable, which are also captured by the Hamilton equation $\dot{q}^{ij} = \{H, q^{ij}\}$ and $\dot{p}^{ij} = \{H, p^{ij}\}$.

4 EMERGENT POINCARÉ SYMMETRIES FROM AN EMERGENT GAUGE THEORY

There is a deep similarity among gauge symmetries and diffeomorphisms, which becomes manifest as soon as both the gauge theories and gravity are formulated as principle bundle

theories. This turns space-time symmetry into an emergent concept, similarly to what has been discussed in the previous section, while considering the consequences of the GHP. A celebrated framework in which gravity, and thus the Poincaré symmetries, are shown to be emergent from a gauge structure was provided by MacDowell and Mansouri. Nonetheless, the gauge symmetry is explicitly broken in this model. We briefly review here this theoretical framework, as a propedeutic element to the next section, where we review a model, due by Wilczek, in which gravity is emergent from a fully gauge-invariant theory.

4.1 Einstein–Hilbert Action

Before introducing MacDowell–Mansouri gravity, it is useful to remind the Palatini formulation of gravity in the Einstein–Hilbert action. This casts in terms of the metric $g_{\mu\nu}$, its inverse $g^{\mu\nu}$, and its first and second derivatives, i.e.:

$$S_{\rm EH} = \frac{1}{16\pi G}\int d^4x\sqrt{-g}\,(R-2\Lambda) \tag{4.1}$$

with R being the Ricci scalar. The Ricci scalar, encoding non-linearly first order derivatives and linearly second order ones, is defined as the contraction of the Riemann tensor, namely:

$$R_{\alpha\beta\mu\nu}g^{\alpha\nu}g^{\beta\mu} = R\,, \tag{4.2}$$

with the Riemann tensor expressed as:

$$R^{\alpha}_{\beta\mu\nu} = \partial_\mu\Gamma^{\rho}_{\nu\sigma} - \partial_\nu\Gamma^{\rho}_{\mu\sigma} + \Gamma^{\alpha}_{\mu\lambda}\,\Gamma^{\lambda}_{\nu\beta} - \Gamma^{\alpha}_{\nu\lambda}\,\Gamma^{\lambda}_{\mu\beta}, \tag{4.3}$$

with $\Gamma^{\rho}_{\mu\nu}$ Christoffel symbols. These latter are torsionless in the Einstein–Hilbert theory, namely $T^{\rho}_{\mu\nu} = \Gamma^{\rho}_{\mu\nu} - \Gamma^{\rho}_{\nu\mu} = 0$. Varying with respect to the Christoffel symbol, one obtains the expression in terms of the metric and its derivatives:

$$\Gamma^{\rho}_{\mu\nu} = \frac{1}{2}g^{\rho\sigma}\left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right). \tag{4.4}$$

A first-order formulation of the Einstein–Hilbert action of gravity is admitted in terms of the $SO(3,1)$-connection ω^{ab}_{μ} and the tetrad one-form (frame-field) e^{a}_{μ}, which is valued in the $SO(3,1)$ algebra and carries an internal vector-index in the fundamental representation of $SO(3,1)$. In terms of these fields, the action now casts:

$$S_{\rm EH} = \frac{1}{64\pi G}\int d^4x\,\epsilon_{abcd}\left(R^{ab}_{\mu\nu}\,e^{c}_{\rho}\,e^{d}_{\sigma} - \frac{\Lambda}{3}\,e^{a}_{\mu}\,e^{b}_{\nu}\,e^{c}_{\rho}\,e^{d}_{\sigma}\right)\epsilon^{\mu\nu\rho\sigma} \tag{4.5}$$

with:

$$\begin{aligned} R_{\mu\nu}{}^{ab} &= \partial_\mu\omega^{ab}_{\nu} - \partial_\nu\omega^{ab}_{\mu} + \omega^{a}_{\mu\,c}\,\omega^{cb}_{\nu} - \omega^{a}_{\nu\,c}\,\omega^{cb}_{\mu}, \\ T^{a}_{\mu\nu} &= \mathcal{D}^{\omega}_{\mu}\,e^{a}_{\nu} - \mathcal{D}^{\omega}_{\nu}\,e^{a}_{\mu}, \end{aligned} \tag{4.6}$$

which can be recast as $R^{ab} = d\omega^{ab} + \omega^{a}_{c}\wedge\omega^{cb}$ and $T^{a} = \mathcal{D}^{\omega}e^{a} = de^{a} + \omega^{a}_{c}\wedge e^{c}$.

A new topological invariant can be added to the action of gravity, without affecting the classical equation of motions. The Holst term can be added to the Einstein–Hilbert action, then leading to the new action that involves a real (Barbero–Immirzi) parameter γ, i.e.:

$$S_{\rm EH} = \frac{1}{64\pi G}\int d^4x\left(\epsilon^{cd}_{ab} + \frac{1}{\gamma}\delta^{cd}_{ab}\right)R^{ab}_{\mu\nu}\,e_{\rho c}\,e_{\sigma\,d}\,\epsilon^{\mu\nu\rho\sigma} - \frac{\Lambda}{3}\,\epsilon_{abcd}\,e^{a}_{\mu}\,e^{b}_{\nu}\,e^{c}_{\rho}\,e^{d}_{\sigma}\epsilon^{\mu\nu\rho\sigma} \tag{4.7}$$

the phase-space of which retains the symplectic form:

$$\{\gamma\omega^{a}_{i}(x),\,\mathcal{E}^{i}_{a}(y)\} = \gamma\delta(x-y)\delta^{j}_{i}\delta^{a}_{b} \tag{4.8}$$

where now $a,b = 1,2,3$, the connection reads $\gamma\omega^{a}_{i} = \omega^{0a}_{i} + \frac{\gamma}{2}\epsilon^{0abc}\omega_{ibc}$, the (Plebanski) two-form reads $\mathcal{E}^{i}_{a} = \frac{4}{G}\epsilon_{abc}\epsilon^{ijk}\,e^{b}_{j}\,e^{c}_{k}$, and as usual $\{\,,\,\}$ denote the Poisson brackets.

4.2 BF Formulation of Gravity

The Einstein–Hilbert–Holst action admits a formulation within the BF framework, as a deviation from the topological theory. The BF theory is defined as a G-principle bundle on a D-dimensional base manifold $\mathcal{M}_D$. The action is the Killing form contraction of the $\mathfrak{g}$ Lie algebra-valued (D-2)-form B and the field strength of the G-connection A. The Lagrangian then reads:

$$\mathcal{L}_{\rm BF} = {\rm tr}\,(B\wedge F[A])\,, \tag{4.9}$$

which specialized to the case of $SO(3,1)$ casts:

$$\mathcal{L}^{{\rm SO}(3,1)}_{\rm BF} = B^{ab}\wedge F^{ab}[A]\,, \tag{4.10}$$

where here $a,b = 1,\ldots 4$ are indices in the fundamental representation of $SO(3,1)$. This automatically provides the Einstein-Hilbert action, when the two-form is constrained to appear as a bi-vector, i.e. $B^{ab} = \epsilon^{ab}_{cd}\,e^{c}\wedge e^{d}$. The Einstein-Hilbert action for gravity, complemented with the Holst term, is then instantiated by the imposition of the so-called simplicity constraints on the B Lie algebra-valued two form, namely:

$$B^{ab} = \pm\left(\epsilon^{ab}_{cd} + \frac{1}{\gamma}\,\delta^{ab}_{cd}\right)e^{c}\wedge e^{d}.$$

4.3 MacDowell–Mansouri Action

Switching now to the MacDowell–Mansouri action, we introduce an extended (anti-de Sitter) group $SO(3,2)$. In the MacDowell–Mansouri action, this is explicitly broken down to its stabilizer, the Lorenz group $SO(3,1)$. The anti-de Sitter connections, composed by ten internal components, are labelled by indices of the fundamental representation of the extended group $A,B = 1,2,\ldots 5$ as $A^{AB} = A^{AB}_{\mu}dx^{\mu}$. This decomposes into $A^{AB}_{\mu} = (A^{ab}_{\mu},\,A^{a5}_{\mu})$, with $A^{ab}_{\mu} = \omega^{ab}_{\mu}$ and $A^{a5}_{\mu} = \ell^{-1}\,e^{a}$), given the identification:

$$\frac{\Lambda}{3} = -\frac{1}{\ell^2}\,.$$

Involving the $SO(3,2)$ algebra-valued connections, the decomposition is recognized to encode both the generators of the Lorentz transformations M_{ab} and the space-time translation P_a, i.e.:

$$A_{\mu} = \frac{1}{2}\omega^{ab}_{\mu}M_{ab} + \frac{1}{\ell}e^{a}_{\mu}P^{a} = \frac{1}{2}A^{AB}_{\mu}M_{AB}\,,$$

having identified $M_{a5} = P_a$.

According to this decomposition, the connection casts as:

$$A^{AB} = \begin{pmatrix} \omega^{ab} & \frac{1}{\ell}e^a \\ -\frac{1}{\ell}e^b & 0 \end{pmatrix}, \tag{4.11}$$

and correspondently the curvature 2-form, with indices contracted with the structure constants of the $SO(3,2)$ group:

$$F^{AB} = dA^{AB} + A^{AC} \wedge A_C^{\ B},$$

decomposes into the $SO(3,1)$ valued components:

$$F^{AB} = \begin{pmatrix} R^{ab} + \frac{1}{\ell^2}e^a \wedge e^b & \frac{1}{\ell}T^a \\ -\frac{1}{\ell}T^b & 0 \end{pmatrix}. \tag{4.12}$$

The MacDowell–Mansouri action deploys this extended formalism, but with the crucial underlying assumption of (explicit) symmetry breaking:

$$F^{AB} \to \overline{F}^{AB} = F^{ab}. \tag{4.13}$$

The Einstein–Hilbert action of gravity can then be encoded in a general framework, moving from the action:

$$S_{MM}[A] = \frac{\ell^2}{64\pi G}\int \mathrm{tr}\left(\overline{F} \wedge \star F\right) \tag{4.14}$$

where $\star$ denotes the gravitational Hodge dual. Using the curvature decomposition in **Eq. 4.12**, the action recasts as:

$$S_{\mathrm{MM}}[A] = \frac{\ell^2}{64\pi G}\int\left(R^{ab} + \frac{1}{\ell^2}e^a \wedge e^b\right) \wedge \left(R^{cd} + \frac{1}{\ell^2}e^c \wedge e^d\right)\epsilon_{abcd} \tag{4.15}$$

The action then entails the Einstein–Hilbert action, with the cosmological term, and some 4D Euler characteristic:

$$32\pi G\, S_{\mathrm{MM}}[A] = S_{\mathrm{EH}} + \frac{1}{2\ell^2}\int \epsilon_{abcd}\, e^a_\mu\, e^b_\nu\, e^c_\rho\, e^d_\sigma\, \epsilon^{\mu\nu\rho\sigma} + \frac{\ell^2}{2}\int \epsilon_{abcd}\, R_{\mu\nu}{}^{ab} R_{\rho\sigma}{}^{cd}\, \epsilon^{\mu\nu\rho\sigma}$$

The equations of motion read:

$$\left(R^{ab} \wedge e^c + \frac{1}{2\ell^2}e^a \wedge e^b \wedge e^c\right)\epsilon_{abcd} = 0, \quad T^a = 0. \tag{4.16}$$

The MacDowell–Mansouri theory admits a straightforward *BF* formulation, involving $\mathrm{so}(2,3)$-valued B two-forms. The action then reads:

$$S = \int_M \mathrm{tr}\left(B \wedge F - \frac{G\Lambda}{6}\overline{B} \wedge \star\overline{B}\right), \tag{4.17}$$

the equations of motion of which imply that: i) varying in δA^{AB}, the local $SO(2,3)$ Gauss constraint holds; ii) varying in δB_{a5}, torsion vanishes, i.e. $F^{a5} = T^a\ell^{-1} = 0$; iii) varying in $\delta\overline{B}_{ab}$, the relations that provide the MacDowell–Mansouri in **Eq. 4.14** is recovered, namely $F^{ab} = G\Lambda/3\epsilon^{abcd}B_{cd}$. This relation allows us to write the MacDowell–Mansouri action as the deformation of a topological gauge theory. The symmetry breaking, which is here explicit, occurs as regulated by a coefficient that is dimensionless, in natural units, and for current estimates of the cosmological constant value reads $G\Lambda \sim 10^{-68}$. This makes General Relativity adapt to be described, with excellent approximation, as the perturbative limit of a topological field theory.

Without entering into further details, we notice that the MacDowell–Mansouri theory can be cast in a similar fashion in terms of an internal de Sitter group $SO(4,1)$, again explicitly broken down to $SO(3,1)$. Within this latter case, the $SO(4,1)$-connection is decomposed into the generator of translations and the generators of rotations, namely the tetrads and the spin-connection $A^{AB} = (\omega^{ab}, \ell^{-1}\, e^a)$.

5 WILCZEK GRAVITY

Frank Wilczek proposed in Wilczek (1998) a model that is reminiscent of the theory formulated by MacDowell and Mansouri, with internal $SO(4,1)$, or equivalently $SO(3,2)$, gauge symmetry. The configuration variables are the gauge symmetry connection A^{AB}_α, and the internal scalar field ϕ^A, in the fundamental representation of the group. The crucial difference between the MacDowell–Mansouri model and the Wilczek model lies in the spontaneous symmetry breaking of the internal gauge group that is present in the latter. This indeed directly enables us to recover the metric structure of General Relativity from the principal bundle of the model proposed by Wilczek.

The Lagrangian considered in Wilczek (1998) is:

$$\mathcal{L}_{\mathrm{W}} = \kappa_3\, \epsilon^{\alpha\beta\gamma\delta}\epsilon_{ABCDE}\, F^{AB}_{\alpha\beta}\, \nabla_\gamma\phi^C\, \nabla_\delta\phi^D\, \phi^E, \tag{5.1}$$

in which $\nabla_\gamma\phi^C = \partial_\gamma\phi^C + A^C_{\gamma F}\phi^F$ denotes the $SO(4,1)$ gauge covariant derivative. The field strength is defined as:

$$F^{AB}_{\alpha\beta} = \partial_\alpha A^{AB}_\beta - \partial_\beta A^{AB}_\alpha + f^{AB}_{\ \ CDEF}A^{CD}_\alpha A^{EF}_\beta, \tag{5.2}$$

in which f^{ABCDEF} is the structure constant of SO(4,1), namely:

$$f^{AB\,LM\,PQ} = \eta^{BL}\eta^{AP}\eta^{MQ} - \eta^{AL}\eta^{BP}\eta^{MQ} - \eta^{BM}\eta^{AP}\eta^{LQ} + \eta^{AM}\eta^{BP}\eta^{LQ}. \tag{5.3}$$

Two novel terms with respect to the MacDowell–Mansouri action, were introduced in the Wilczek model:

1. The interaction potential of ϕ^A, namely:

$$\mathcal{L}_1 = \kappa_1\left(\eta_{AB}\phi^A\phi^B - v^2\right)^2.$$

By varying with respect to ϕ^A, this term is recognized to be stationarized either for $\phi^A = 0$ or when $|\phi| = v$. In the latter case, the choice $\phi^A = \delta^A_5 v$ implements the spontaneous symmetry breaking.

2. A term that constrains the determinant of the metric in the spontaneous symmetry broken phase:

$$\mathcal{L}_2 = \kappa_2 (J - \omega)^2 ,$$

in which:

$$J = \epsilon^{\alpha\beta\gamma\delta}\epsilon_{ABCDE}\,\phi^E\nabla_\alpha\phi^A\nabla_\beta\phi^B\nabla_\gamma\phi^C\nabla_\delta\phi^D. \tag{5.4}$$

This term is stationarized when $J = \omega$, implying the unimodularity of gravity (Bufalo et al., 2015). The spontaneous symmetry breaking induces in **(5.4)** a reduced expression for J, namely:

$$J = v^5\epsilon^{\alpha\beta\gamma\delta}\epsilon_{abcd}A^{a5}_\alpha A^{b5}_\beta A^{c5}_\gamma A^{d5}_\delta , \tag{5.5}$$

J denoting the determinant of the metric.

The total Lagrangian proposed by Wilczek then reads:

$$\mathcal{L}_W = \kappa_2 (J - \omega)^2 + \kappa_1 \left(\eta_{AB}\Phi^A\Phi^B - v^2\right)^2 + \kappa_3\,\epsilon^{\alpha\beta\gamma\delta}\epsilon_{ABCDE}\,F^{AB}_{\alpha\beta}\,\nabla_\gamma\phi^C\,\nabla_\delta\phi^D\,\phi^E.$$

In order to unveil the emergence of the gravity, we may instantiate the spontaneous symmetry breaking **Eq. 4.13** directly on the Lagrangian in **Eq. 5.1**, using the decompositions recovered in **Eqs. (4.11)** and **(4.12)**, and then finding:

$$\begin{aligned}\mathcal{L}_{\rm W} &= \kappa_3\, v^3\,\epsilon^{\alpha\beta\gamma\delta}\epsilon_{abcd}\left[\left(\partial_\alpha\omega^{ab}_\beta - \partial_\beta\omega^{ab}_\beta + f^{ab}_{cdef}\omega^{cd}_\alpha\omega^{ef}_\beta\right) - \Lambda\, e^a_\alpha e^b_\beta\right]e^c_\gamma e^d_\delta \\ &= \kappa_3\, v^3\,\epsilon^{\alpha\beta\gamma\delta}\epsilon_{abcd}\left[F^{ab}_{\alpha\beta} - \Lambda\, e^a_\alpha e^b_\beta\right]e^c_\gamma e^d_\delta.\end{aligned} \tag{5.6}$$

This equation corresponds to the Einstein–Hilbert action, introduced in **Eq. 4.5**, plus the cosmological constant term. The unimodular term is not essential for our arguments, and we can neglect it here.

We can recast the main term of the Lagrangian density proposed by Wilczek, rearranging as:

$$\mathcal{L} = \kappa\,\epsilon^{\alpha\beta\gamma\delta}F^{AB}_{\alpha\beta}B^{AB}_{\gamma\delta}, \tag{5.7}$$

where

$$B^{AB}_{\gamma\delta} = \nabla_\gamma\Phi^C\nabla_\delta\Phi^D\,\Phi^E\,\epsilon^{AB}_{CDE}$$

works as a simplicity constraint, which here drags the Wilczek model away from its topological phase.

In the Higgs condensate phase, the $B \wedge F$ term is (nothing but) reduced to the Einstein-Hilbert term, and an emergent diffeomorphism invariance is recovered starting from a topological invariance, which is finally broken. In this way, moving from a background independent theory, because of the flatness of the connection, after the Higgs condensate phase an emergent metric tensor is obtained.

6 CONCLUSIONS AND OUTLOOKS

We have shown here that a generalized version of the holographic principle can be derived from fundamental considerations of quantum information theory, in particular, the imposition of separability on a joint state. This GHP entails gauge invariance. We emphasized that as soon as this is instantiated in an ambient Lorentzian space-time, gauge invariance under the Poincaré group automatically follows. Indeed, following this pathway we can recover the action of gravity. We summarize several gauge-invariant models for gravity, including gravity cast à la Wilczek. This is a formulation of the Einstein theory of gravity similar to the one proposed by MacDowell and Mansouri, which involves the representation theory of the Lie groups $SO(3,2)$ and $SO(4,1)$.

As the GHP provides a natural and completely general distinction between "bulk" and "boundary" degrees of freedom, one that is independent of geometry, it would be worth to investigate whether the AdS/CFT and dS/CFT correspondences could fit within this framework. This would require complementing the GHP with the concept of the renormalization group flow. Indeed, group renormalization flow techniques might be actually considered to connect the fully symmetric $SO(3,2)$ and $SO(4,1)$ theory with the SO(3,1) broken symmetric phase.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Ideation by CF and AM. Development by all the authors.

REFERENCES

Arnowitt, R., Deser, S., and Misner, C. W. (1959). Dynamical Structure and Definition of Energy in General Relativity. *Phys. Rev.* 116 (5), 1322–1330. doi:10.1103/PhysRev.116.1322

Barceló, C., Liberati, S., and Visser, M. (2005). Analogue Gravity. *Living Rev. Relativ.* 8, 12, 2005. [arXiv:gr-qc/0505065 [gr-qc]]. doi:10.12942/lrr-2005-12

Bartlett, S. D., Rudolph, T., and Spekkens, R. W. (2007). Reference Frames, Superselection Rules, and Quantum Information. *Rev. Mod. Phys.* 79, 555–609. doi:10.1103/revmodphys.79.555

Bekenstein, J. D. (2004). Black Holes and Information Theory. *Contemp. Phys.* 45 (1), 31–43. doi:10.1080/00107510310001632523

Bousso, R. (2002). The Holographic Principle. *Rev. Mod. Phys.* 74 (3), 825–874. doi:10.1103/revmodphys.74.825

Bufalo, R., Oksanen, M., and Tureanu, A. (2015). How Unimodular Gravity Theories Differ from General Relativity at Quantum Level. *Eur. Phys. J.* C75 (10), 477. doi:10.1140/epjc/s10052-015-3683-3

Chamseddine, A. H., and West, P. C. (1977). Supergravity as a Gauge Theory of Supersymmetry. *Nucl. Phys. B* 129, 39–44. doi:10.1016/0550-3213(77)90018-9

Chiang, H.-W., Hu, Y.-C., and Chen, P. (2016). Quantization of Spacetime Based on a Spacetime Interval Operator. *Phys. Rev. D* 93 (no.8), 084043, 2016. arXiv: 1512.03157 [hep-th]]. doi:10.1103/PhysRevD.93.084043

Chiribella, G., Perinotti, G. M., and Pernotti, P. (2016). "Quantum from Principles," in *Quantum Theory: Informational Foundations and Foils (Fundamental Theories of Physics 181)*. Editors G. Chiribella and R. W. Spekkens (Dordrecht: Springer), 171–221. doi:10.1007/978-94-017-7303-4_6

Einstein, A. (1948). *Quanten-mechanik Und Wirklichkeit*, 320–324. doi:10.1111/j.1746-8361.1948.tb00704.x

Faulkner, T., Guica, M., Hartman, T., Myers, R. C., and Van Raamsdonk, M. (2014). Search for Physics beyond the Standard Model in Multilepton Final States in Proton-Proton Collisions at $S\sqrt{s} = 13$ TeV. *JHEP* 03, 051, 2014. [arXiv:1312.7856 [hep-th]]. doi:10.1007/JHEP03

Fields, C. (2019). Decoherence as a Sequence of Entanglement Swaps. *Results Phys.* 12, 1888–1892. doi:10.1016/j.rinp.2019.02.007

Fields, C., and Marcianò, A. (2019). Sharing Nonfungible Information Requires Shared Nonfungible Information. *Quan. Rep.* 1, 252–259. doi:10.3390/quantum1020022

Fields, C., and Marcianò, A. (2020). Holographic Screens Are Classical Information Channels. *Quant. Rep.* 2, 326–336. doi:10.3390/quantum2020022

Flammia, S. T., Hamma, A., Hughes, T. L., and Wen, X. G. (2009). Topological Entanglement Rényi Entropy and Reduced Density Matrix Structure. *Phys. Rev. Lett.* 103, 261601, 2009. [arXiv:0909.3305 [cond-mat.str-el]]. doi:10.1103/PhysRevLett.103.261601

Garay, L. J. (1995). Quantum Gravity and Minimum Length. *Int. J. Mod. Phys. A.* 10, 145–165. [arXiv:gr-qc/9403008 [gr-qc]]. doi:10.1142/S0217751X95000085

Hamma, A., and Markopoulou, F. (2011). Background-independent Condensed Matter Models for Quantum Gravity. *New J. Phys.* 13, 095006, 2011. [arXiv:1011.5754 [gr-qc]]. doi:10.1088/1367-2630/13/9/095006

Jacobson, T. (1995). Thermodynamics of Spacetime: The Einstein Equation of State. *Phys. Rev. Lett.* 75, 1260–1263. [gr-qc/9504004]. doi:10.1103/PhysRevLett.75.1260

Landauer, R. (1961). Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.* 5, 183–191. doi:10.1147/rd.53.0183

MacDowell, S. W., and Mansouri, F. (1977). Unified Geometric Theory of Gravity and Supergravity. *Phys. Rev. Lett.* 38 (14), 739–742. doi:10.1103/physrevlett.38.739

Maldacena, J. (1998). The Large *N* Limit of Superconformal Field Theories and Supergravity. *Adv. Theor. Math. Phys.* 2, 231–252. doi:10.4310/atmp.1998.v2.n2.a1

Modesto, L., and Rachwał, L. (2014). Super-renormalizable and Finite Gravitational Theories. *Nucl. Phys. B* 889, 228–248. [arXiv:1407.8036 [hep-th]]. doi:10.1016/j.nuclphysb.2014.10.015

Modesto, L. (2012). Super-renormalizable Quantum Gravity. *Phys. Rev. D* 86, 044005, 2012. [arXiv:1107.2403 [hep-th]]. doi:10.1103/PhysRevD.86.044005

Oh, E., Park, I. Y., and Sin, S.-J. (2018). Complete Einstein Equations from the Generalized First Law of Entanglement. *Phys. Rev. D* 98 (no. 2), 026020, 2018. arXiv:1709.05752 [hep-th]]. doi:10.1103/PhysRevD.98.026020

Polchinski, J. (1998). *String Theory*. Cambridge University Press. doi:10.2172/5055719

Rovelli, C. (2017). *Black Holes Have More States than Those Giving the Bekenstein-Hawking Entropy: A Simple Argument*. Preprint. arXiv:1710:00218 [gr-qc].

Rovelli, C. (2004). *Quantum Gravity*. Cambridge University Press. doi:10.1017/CBO9780511755804

Rovelli, C. (2019). The Subtle Unphysical Hypothesis of the Firewall Theorem. *Entropy* 21, 839. doi:10.3390/e21090839

Stelle, K. S., and West, P. C. (1979). de Sitter Gauge Invariance and the Geometry of the Einstein-Cartan Theory. *J. Phys. A: Math. Gen.* 12, L205. doi:10.1088/0305-4470/12/8/003

Susskind, L. (1995). The World as a Hologram. *J. Math. Phys.* 36, 6377–6396. doi:10.1063/1.531249

Swingle, B., and Van Raamsdonk, M. (2014). *Universality of Gravity from Entanglement*. High Energy Physics - Theory. arXiv:1405.2933 [hep-th].

't Hooft, G. (1993). "Dimensional Reduction in Quantum Gravity," in *Salamfestschrift. Singapore*. Editors A. Ali, J. Ellis, and S. Randjbar-Daemi (Singapore: World Scientific), 284–296.

Van Raamsdonk, M. (2010). Building up Spacetime with Quantum Entanglement. *Gen. Relativ Gravit.* 42, 2323–2329. arXiv:1005.3035 [hep-th]]. doi:10.1142/S0218271810018529

Verlinde, E. P. (2011). On the Origin of Gravity and the Laws of Newton. *JHEP* 1104, 029, 2011. [arXiv:1001.0785 [hep-th]]. doi:10.1007/JHEP04(2011)029

Weyl, H. (1918). "Gravitation and Electricity. Sitzungsber Preuss Akad Wiss Berlin 1918, 465," reprinted in *The Principle of Relativity*. Editors A. Einstein and F. A. Davis (New York: Dover).

Wilczek, F. (1998). Riemann-Einstein Structure from Volume and Gauge Symmetry. *Phys. Rev. Lett.* 80 (22), 4851–4854. doi:10.1103/physrevlett.80.4851

Conflict of Interest: 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.

The handling editor declared a past co-authorship with one of the authors AM.

frontiers in Astronomy and Space Sciences

ORIGINAL RESEARCH
published: 10 June 2021
doi: 10.3389/fspas.2021.571282

Prelude to Simulations of Loop Quantum Gravity on Adiabatic Quantum Computers

*Jakub Mielczarek**

Institute of Theoretical Physics, Jagiellonian University, Kraków, Poland

The article addresses the possibility of implementing spin network states, used in the loop quantum gravity approach to Planck scale physics on an adiabatic quantum computer. The discussion focuses on applying currently available technologies and analyzes a concrete example of a D-Wave machine. It is introduced a class of simple spin network states which can be implemented on the Chimera graph architecture of the D-Wave quantum processor. However, extension beyond the currently available quantum processor topologies is required to simulate more sophisticated spin network states. This may inspire new generations of adiabatic quantum computers. A possibility of simulating loop quantum gravity is discussed, and a method of solving a graph non-changing scalar (Hamiltonian) constraint with the use of adiabatic quantum computations is proposed. The presented results establish a basis for the future simulations of Planck scale physics, specifically quantum cosmological configurations, on quantum annealers.

Keywords: quantum gravity, quantum computation, Planck scale, quantum annealing, adiabatic quantum algorithm

OPEN ACCESS

Edited by:
Jan De Boer,
University of Amsterdam, Netherlands

Reviewed by:
Francesca Vidotto,
Western University, Canada
Sayantan Choudhury,
National Institute of Science Education and Research (NISER), India

***Correspondence:**
Jakub Mielczarek
jakub.mielczarek@uj.edu.pl

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Astronomy and Space Sciences

Received: *10 June 2020*
Accepted: *25 May 2021*
Published: *10 June 2021*

Citation:
Mielczarek J (2021) Prelude to Simulations of Loop Quantum Gravity on Adiabatic Quantum Computers. Front. Astron. Space Sci. 8:571282. doi: 10.3389/fspas.2021.571282

1 INTRODUCTION

One can distinguish two main applications of quantum computers. The first is data processing associated with the implementation of quantum algorithms, e.g., quantum machine learning protocols (Biamonte et al., 2017). The second concerns simulations of quantum systems.

Simulating quantum systems using quantum computers is fundamentally different from what simulations performed at classical computers are. While classical simulations rely on either discretization of a given physical system or an adequate algebraic analysis, the simulations performed on quantum computers allow to *imitate* a given quantum system. This kind of *exact simulation* of a quantum system has been a subject of discussion in the seminal Feynman's article (Feynman, 1982).

In order to understand better what we mean by exact simulations, let us consider the case of Planck scale physics. Here, the relevant degrees of freedom are defined at length scales of the order of the Planck length $l_{\rm Pl} \sim 10^{-35}$ m. Despite significant advances made in theoretical understanding and experimental techniques, the Planck scale physics remains empirically directly inaccessible.

On the other hand, concrete examples of theories describing elementary quantum gravitational degrees of freedom exist. One such approach is loop quantum gravity (LQG) (Rovelli, 1998; Ashtekar and Lewandowski, 2004; Rovelli and Vidotto, 2014). In LQG, background-independent degrees of freedom can be defined, and predictions can be made (Agullo et al., 2012; Barrau et al., 2014). Even if the degrees of freedom under consideration are experimentally not directly accessible, one can consider their *projection* onto another physical realization, which will imitate its quantum behavior (see **Figure 1**). Under the assumption that quantum mechanics is valid at the Planck scale, the

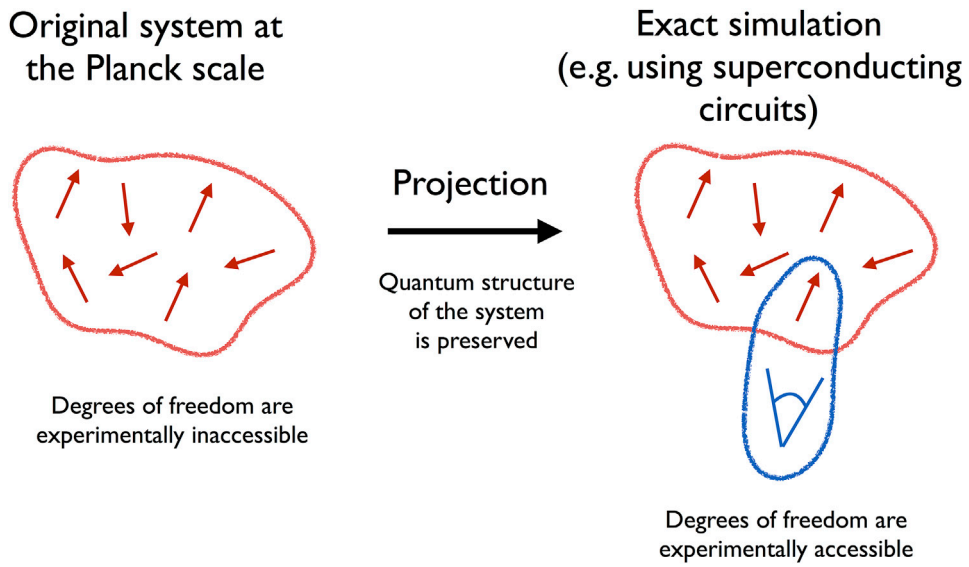

FIGURE 1 | Pictorial presentation of the relation between the original quantum system defined at the Planck scales and its exact simulation. The exact simulation is performed with the use of projection of the original quantum system onto the architecture of a quantum computer. In contrast to the Planck scale system, measurements of the quantum degrees of freedom can be performed at the level of the quantum simulation.

projected realization is quantum-mechanically equivalent to the original system. The only difference is appropriately rescaled couplings adjusted to the physical nature of the simulator (built, e.g., with the use of superconducting qubits). Such projection of one quantum system onto its equivalent imitation allows performing what we previously called exact simulations. As it has already been mentioned, the quantum simulations are very different from what we usually consider as physical simulations, where for instance, a given differential equation is discretized and then implemented on a computer with the use of appropriate algorithms. In contrast, in the case of exact simulations, one actually performs experiments on a quantum system which is defined as being equivalent (quantum mechanically) to a part (or a whole) of the original quantum system.

The aim of this article is to investigate the possibility of performing exact quantum simulations of the spin network states, which are used to construct the Hilbert space in the LQG approach to quantum gravity. Our interest is focused on applying existing adiabatic quantum computers, a commercial example of which is one provided by the D-Wave Systems company.

We consider conceptual issues related to the possibility of simulating Loop Quantum Gravity based on the considered fixed graph case. Specifically, the implementation of the scalar constraint is analyzed, and a toy model of such a procedure is presented. We conclude with an outlook of the following steps to be done in the research direction initiated in this article.

The idea of employing the spin network states in quantum computations already appeared in the literature [see Refs Marzuoli and Rasetti (2005), Kauffman and Lomonaco (2008), Jordan (2010)]. However, the potential of applying spin networks for universal quantum data processing was considered only. Up to the best of our knowledge, the issue of relating spin networks with adiabatic quantum computations was not considered before. While this article was in the final stage of preparing a study in which an LQG spin network is implemented on a molecular quantum simulator appeared. A simulation of quantum fluctuations of a 5-node spin network in the kinematical regime was performed (Li et al., 2019). Here, we will consider the same type of spin network in **Section 4** in the context of solving a prototype scalar constraint with the use of adiabatic quantum computations.

2 ADIABATIC QUANTUM COMPUTING

The last years have brought significant progress in the development of quantum computing technologies (Campbell et al., 2017). First, quantum computers have been commercialized and made available in a cloud or as independent hardware units. In both cases, the currently most advanced commercial technologies were possible to achieve thanks to the development of superconducting quantum circuits (You and Nori, 2005). In particular, the IBM Q universal quantum computers built using 5 and 20 superconducting qubits have been developed. However, from the point of view of exact simulations discussed in the Introduction, another type of quantum computer seems to be more suitable to use currently—namely the *adiabatic quantum computers* (Kadowaki and Nishimori, 1998).

Adiabatic quantum computers, in contrast to the universal ones, are designed to solve a specific problem of finding the

minimum of a Hamiltonian H_I of a coupled system of qubits (spins). In the process of finding the minimum of H_I one employs a time-dependent Hamiltonian in the form

$$H(\lambda) = (1-\lambda)H_B + \lambda H_I, \tag{1}$$

where H_B is the so-called base Hamiltonian, which is characterized by a simple and easy to prepare ground state. In practice, the base Hamiltonian is often equal to $H_B = \sum_i \sigma_i^x$, so that the ground state corresponds to the alignment of spins in the x direction. Then, the value of λ is changed adiabatically from $\lambda = 0$ to $\lambda = 1$, so that while the system that is initially in a non-degenerate ground state will remain in this state. Therefore, if the process is done correctly, the system ends up in the minimum of the Hamiltonian H_I. The process of transition from $\lambda = 0$ to $\lambda = 1$ involves quantum tunneling and is called *quantum annealing*. The characteristic time scales which allow keeping the system close to the adiabatic state are dependent on what kind of H_I Hamiltonian is considered. This issue is closely related to the efficiency of the quantum annealing-based algorithms to the classical ones [see Ref Biamonte et al. (2017)] for discussion of this subject).

In practical implementations, the most considered form of H_I corresponds to the Ising problem:

$$H_I = \sum_{\langle i,j\rangle} b_{ij}\sigma_i^z\sigma_j^z + \sum_i h_i\sigma_i^z, \tag{2}$$

where b_{ij} are coupling between spins and h_i quantify interactions of spins with external magnetic field. The summation $\langle i,j\rangle$ is defined such that it does not repeat over pairs. The values of couplings specify the problem to be solved while readouts of z components of the spins in the final state provide an outcome of the quantum computation (simulation).

From the mathematical viewpoint, the class of problems which can be solved in that way is the so-called Quadratic Unconstrained Binary Optimization (QUBO), which typically is of the NP-hard type. This is because, when we look a the problem from the classical perspective by measuring the orientations of spins along the z-axis, the two values of $\sigma_i^z \to s_i \in \{-1, 1\}$ are allowed. In consequence, for the system of N classical spins, there are 2^N configurations to be explored. Therefore, in general, finding a ground state requires exponential growth of time with the number of spins (N), which is the case for NP-hard problems.

The physical implementation of the QUBO problem using the quantum annealing procedure is provided by the D-Wave machine. In this realization, the spins (qubits) are created with the use of superconducting circuits in the form of CC JJ RF-SQUIDs (Harris et al., 2010) built with the use of Josephson junctions composed of Niobium in the superconducting state. The qubit basis states are defined employing two different orientations of quantum of magnetic flux across the superconducting circuit. SQUID (superconducting quantum interference device) introduces interactions between the qubits based circuits called *couplers*, which introduce the $b_{ij}\sigma_i^z\sigma_i^z$ factors in the Hamiltonian **(2)**. The values of parameters h_i can be controlled with the use of external magnetic fluxes. However, not all values of b_{ij} and h_i are allowed but only some fractional values from the range $[-1, 1]$. The readout of the final quantum states of qubits is performed with the use of sensitive magnetometers built with the use of SQUIDs.

In the D-Wave quantum annealer, the superconducting qubits q_i are arranged into 8-qubit blocks forming the so-called *Chimera* architecture. Each block consists of 16 couplings between 8 spins. As a consequence, not all qubits are coupled. The topology of couplings between qubits in a single block is presented in **Figure 2**. In the so far most advanced version of the D-Wave machine (the D-Wave 2000) the 8-qubit blocks form a 16 × 16 matrix (256 blocks in total), leading to 2048 qubits.

3 SPIN NETWORKS

Let us now proceed to discuss spin networks. We begin with a brief overview of how spin networks appear in the loop quantum gravity approach to quantum gravity. Then we introduce a class of spin networks that is possible to implement on the Chimera architecture of D-Wave quantum processors.

The fundamental elements of the LQG approach to quantum gravity are holonomies of Ashtekar connection $A \in \mathfrak{su}(2)$ along a curve $e(\lambda)$, with $\lambda \in [0, 1]$:

$$h[A, e] = \mathcal{P}e^{\int_e A}, \tag{3}$$

where $\mathcal{P}$ denotes path ordering. Performing gauge transformations, generated by the so-called Gauss constraint, the Ashtekar connection transforms as $A \to A_g = g^{-1}dg + g^{-1}Ag$. The corresponding transformation of holonomy is $h[A, e] \to h[A_g, e] = g(e(0))h[A, e]g(e(1))^{-1}$. The fact that the transformations of holonomies contribute only at the boundaries of e implies that gauge invariant objects are provided by the Wilson loops $W[A, e] := \mathrm{tr}(h[A, e])$.

The key idea behind LQG is to build a Hilbert space of the theory out of Wilson loops. However, such a basis is, in general, over-complete. A solution to the problem comes from the construction of *spin-networks* which are a certain linear combination of products of Wilson loops (Rovelli and Smolin, 1990). Such an approach guarantees that both the Gauss constraint (ensuring local gauge invariance) is satisfied by the base states, and the Hilbert space is complete. By introducing an equivalence relation between topologically equivalent spin networks, the so-called diffeomorphism constraint can be satisfied. There is eventually a scalar (Hamiltonian) constraint, which has to be satisfied by physical states. In this section, we will focus on spin network states satisfying both Gauss and diffeomorphism constraints. The issue of satisfying the scalar constraint (with the use of adiabatic quantum computing) will be discussed in the next section.

The spin network is formally a graph composed of edges E and nodes N with spin labels at the edges and the so-called *intertwiners* at the nodes. The spin labels correspond to irreducible representations of the $SU(2)$ group such that triangle inequalities (reflecting the Gauss constraint) are satisfied at the nodes. The intertwiners correspond to invariant subspaces at the nodes, which we will discuss in more detail below.

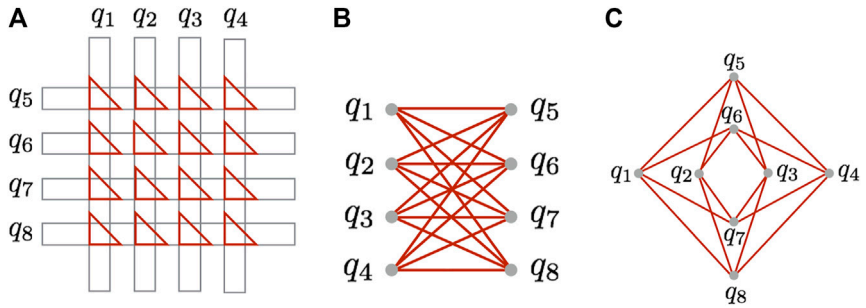

FIGURE 2 | Three different representations of the structure of couplings between eight q_i qubits forming an elementary block of the D-Wave processor: **(A)** The physical representation of quibits as closed superconducting loops. The couplers between the qubits are represented by triangles. **(B)** The Chimera graph structure of couplings between the eight qubits. **(C)** Representation of a single block which is useful when interconnections in the array of blocks are considered.

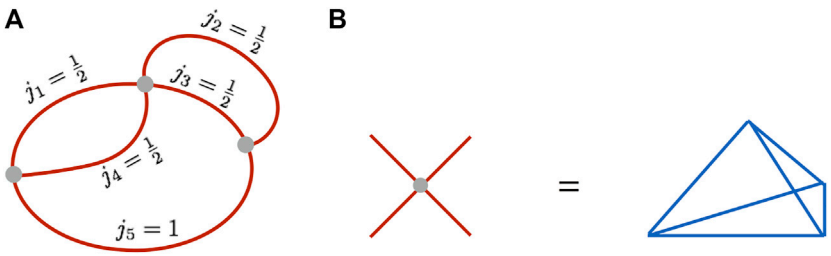

FIGURE 3 | A) An exemplary spin network. **(B)** In the geometric picture the 4-valent node is dual to a tetrahedron (3-simplex).

In **Figure 3** we present an exemplary spin network composed of 3-valent and 4-valent nodes. An essential feature of the nodes is that 3-valent nodes do not carry a volume element while 4-valent nodes and higher valent nodes are associated with 3-volume (in the sense that eigenvalues of the volume operator in such a state are non-vanishing).

The particular case is a 4-valent node that is dual to a tetrahedron (3-simplex). One can imagine that the vertex is located in the center of the tetrahedra, while each of the associated links intersects with one of the surfaces (see block b) in **Figure 3**).

In this article, we are considering the case of spin networks composed of 4-valent vertices only and spin labels corresponding to fundamental representations of the $SU(2)$ group i.e. $j = \frac{1}{2}$. The reason for that is that in such a case, the Hilbert space at each vertex is a tensor product of four 1/2 spins, which can be decomposed into irreducible representations in the following way:

$$H_{\frac{1}{2}} \otimes H_{\frac{1}{2}} \otimes H_{\frac{1}{2}} \otimes H_{\frac{1}{2}} = H_0 \oplus H_0 \oplus 3H_1 \oplus H_2. \tag{4}$$

There are, therefore, two possibilities in which the spins can add up to zero. In consequence, the invariant subspace for such a vertex is two dimensional:

$$\dim \mathrm{Inv}\left(H_{\frac{1}{2}} \otimes H_{\frac{1}{2}} \otimes H_{\frac{1}{2}} \otimes H_{\frac{1}{2}}\right) = 2. \tag{5}$$

We associate the two dimensional invariant space with the qubit space. The nature of the qubit associated with the 4-vertex under consideration is graphically presented in **Figure 4**.

Worth mentioning here is that there is a freedom of choice of channel in which recoupling of spin labels at the vertex is performed. In particular, in the s channel the base states can be expressed in terms of the four 1/2 spin base states ($|\uparrow\rangle$ and $|\downarrow\rangle$) as follows (Feller and Livine, 2016):

$$\begin{aligned} |0_s\rangle &= \frac{1}{2}\left(|\uparrow\downarrow\uparrow\downarrow\rangle + |\downarrow\uparrow\downarrow\uparrow\rangle - |\uparrow\downarrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\downarrow\rangle\right) \\ |1_s\rangle &= \frac{1}{\sqrt{3}}\left(|\uparrow\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\uparrow\rangle \right. \\ &\quad \left. - \frac{|\uparrow\downarrow\uparrow\downarrow\rangle + |\downarrow\uparrow\downarrow\uparrow\rangle + |\uparrow\downarrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\downarrow\rangle}{2}\right). \end{aligned} \tag{6}$$

The above definitions of the basis states for a qubit have recently been considered in the context of universal quantum computations in Refs. Li et al. (2019), Mielczarek (2019), Czelusta and Mielczarek (2021). Alternatively, it is convenient to construct our qubit states such that they are eigenvectors of the volume operator [see e.g. Ref Feller and Livine (2016) for details]:

$$\widehat{V} := \frac{\sqrt{2}}{3} l_{\mathrm{Pl}}^3 (8\pi\gamma)^{\frac{3}{2}} \sqrt{\left|\widehat{\vec{J}}_1 \cdot \left(\widehat{\vec{J}}_2 \times \widehat{\vec{J}}_3\right)\right|}, \tag{7}$$

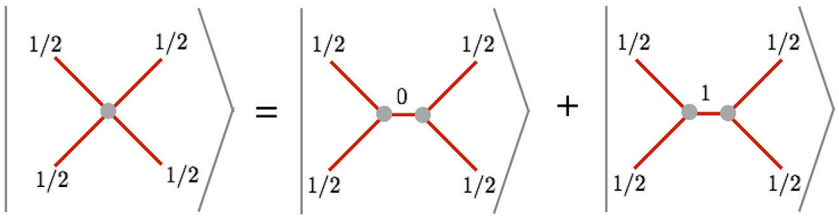

FIGURE 4 | The 4-valent node with spin labels equal to $j = 1/2$ corresponds to superposition of two graphs each with the different value of intertwiner. In this article, the two dimensional intertwiner space is associated with the qubit Hilbert space.

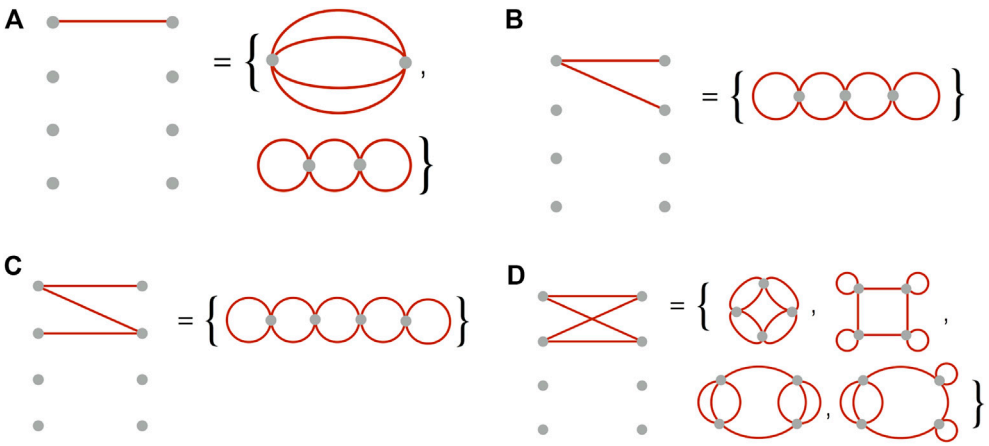

FIGURE 5 | Connected spin networks compatible with the Chimera architecture: **(A)** $N = 2$, **(B)** $N = 3$, **(C)** $N = 4$, **(D)** $N = 4$.

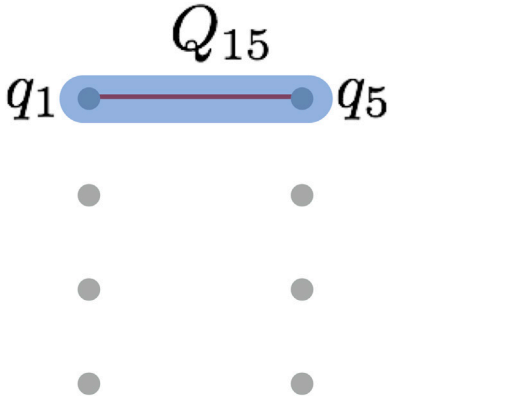

FIGURE 6 | A chain qubit. If the b_{15} coupling is sufficiently negative then the spins q_1 and q_5 will have tendency to align in the same direction.

where γ is the Barbero-Immirzi parameter (Barbero G, 1995; Immirzi, 1997) and the angular momentum operators $\overrightarrow{\mathcal{J}}_i = (\widehat{\mathcal{J}}_i^1, \widehat{\mathcal{J}}_i^2, \widehat{\mathcal{J}}_i^3)$ satisfy the $\mathfrak{su}(2)$ algebra: $[\widehat{\mathcal{J}}_k^a, \widehat{\mathcal{J}}_l^b] = i\delta_{kl}\epsilon_c^{ab}\widehat{\mathcal{J}}_k^c$, for $a, b, c \in \{1, 2, 3\}$. This can be achieved by considering the following superpositions of the states **Eq. 6**:

$$|1\rangle = \frac{1}{\sqrt{2}}(|0_s\rangle - i|1_s\rangle), \tag{8}$$

$$|0\rangle = \frac{1}{\sqrt{2}}(|0_s\rangle + i|1_s\rangle), \tag{9}$$

so that $\widehat{V}|1\rangle = +V_0|1\rangle$ and $\widehat{V}|0\rangle = -V_0|0\rangle$, where $V_0 := \frac{l_{\rm Pl}^3(8\pi\gamma)^{\frac{3}{2}}}{\sqrt{6\sqrt{3}}}$ is the minimal eigenvalue of the volume operator (Rovelli and Vidotto, 2014). The positive and negative signs of the eigenvalues distinguish between the two allowed orientations of a 3-simplex. Consequently, in LQG, the elementary volume can contribute with both signs: positive (for $|1\rangle$) or negative (for $|0\rangle$). However, it is expected that in the semiclassical limit, one of the contributions will dominate over the other. The eigenstates $|1\rangle$ and $|0\rangle$ are the qubit base states that we refer to in the rest of this article.

Having the definition of a qubit, one can consider different spin network topologies that are possible to implement directly with Chimera architecture. In **Figure 5** we present connected spin networks with the number of nodes equal to $N = 2, 3$ and 4, which can be directly embedded into the Chimera graph.

A single coupling between the qubits in the quantum processor architecture can be associated with one or more

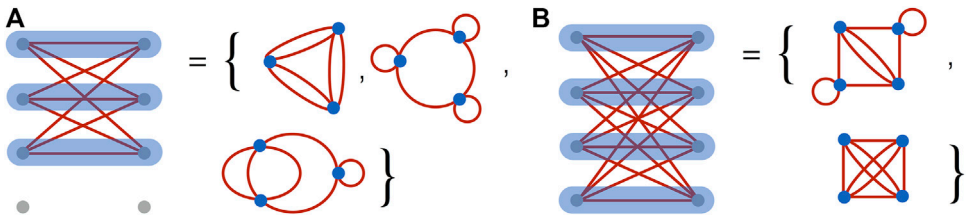

FIGURE 7 | Connected spin networks compatible with Chimera architecture: **(A)** $N = 3$, **(B)** $N = 4$.

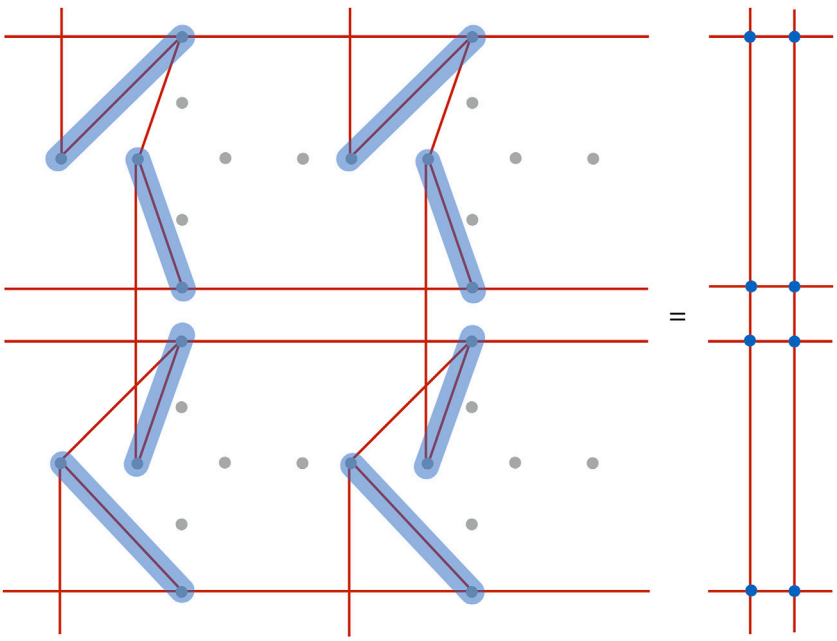

FIGURE 8 | Regular lattice spin network embedded with the use of chain qubits. Here, four 8 qubit blocks of the D-Wave processor are used.

links in the corresponding spin network. The difference between connections can be further encoded in the strength of the couplers. In particular, in the case a) in **Figure 5** there are two possible 4-valent spin networks can be associated with two coupled qubits. Relating a single coupler with a single link in the spin network is generically not possible. In the case of 4-valent nodes and a single block of D-Wave processor, the only possibility is given by the configuration represented in blocks b) and c) in **Figure 2**. The situation corresponds to a spin network with $N = 8$ qubits and $E = 16$ edges forming the Chimera graph.

As one can notice, the structure of Chimera architecture imposes significant restrictions on the possible associated spin network topologies. In particular, it is not possible to implement a "triangular" $N = 3$ spin network directly with the use of elementary qubits. In order to go beyond the limitations of the Chimera architecture, one can consider effective qubits (chain qubits) composed of two or more spins. If the coupling between the qubits is sufficiently negative, then the qubits will tend to align in the same direction, which is preferred energetically. In such a case, measurements can be performed on one of the elementary qubits contributing to the chain, while the remaining qubits can be considered ancilla qubits.

With the use of chain qubits (see **Figure 6**), the dictionary of spin networks can be extended further. In particular, previously inaccessible spin networks for $N = 3$ and $N = 4$ can now be constructed (see **Figure 7**).

Worth stressing is that different types of effective qubits can be considered and **Figure 7** represents only a one of many possible implementations of the spin networks under consideration.

A more extended example is a regular square lattice with the nearest neighbor connections. Implementation of the regular lattice spin network on the Chimera architecture of the quantum processor is shown in **Figure 8**.

The regular lattice configuration enables simulation of a 2D *Ising spin networks* discussed in Ref (Feller and Livine, 2016). which provide a toy model of extended quantum spacetime. The analysis of such configurations is especially interesting in the context of phase transitions and domain formation, which may reflect the emergence of semi-classical spacetime. We will come back to this issue in the next section. Furthermore, in

further studies, it would be interesting to investigate if the 3D hexagonal type *Ising spin network* can also be embedded into the architecture of the D-Wave processor.

4 SIMULATION OF LOOP QUANTUM GRAVITY

The spin network states discussed in the previous sections satisfy the Gauss constraint. Moreover, the diffeomorphism constraint can be imposed by considering equivalence classes under the action of spatial diffeomorphisms, which means that we equate all graphs with the same topology.

It remains the scalar (Hamiltonian) constraint, which is the most difficult one to satisfy. Finding a solution to the Hamiltonian constraint in the 3 + 1 D can be perceived as the most difficult open problem in LQG (Thiemann, 2006).

The scalar constraint reflects the fact that the total energy of the gravitational field is equal to zero. The scalar constraint is, in general, a graph-changing operator, which makes implementation of such a constraint a problematic task. However, the situation in which the action of the constraint preserves the graph structure may provide an intermediate step toward the solution of the full problem. The question is whether adiabatic quantum computation may be helpful here.

To address this question, let us observe that finding solutions to the classical constraint:

$$C \approx 0, \tag{10}$$

can be mapped into a problem of minimizing some Hamiltonian. Because the quantum annealing algorithm is just searching for the minimum of the spin Hamiltonian **(2)**, making use of adiabatic computations requires association of the minimum of the Hamiltonian with a solution of the constraint **(10)**. The simplest way to do it is to consider the Hamiltonian H in the following form:

$$H \propto C^2. \tag{11}$$

In this case, the Hamiltonian is bounded from below, and in the classical ground state, the constrain **(10)** is naturally satisfied.

Solution of the constraint **(10)** allows to extract physical states and construct a physical phase space Γ_{phys} (or physical Hilbert space) being a subset of kinematical phase space Γ_{kin}[1]. It is important to stress that the minimum energy states of the Hamiltonian **(11)** are just the physical states of the theory and they form Γ_{phys}. If there is no additional matter content, the states represent also a vacuum gravitational field configuration, described by the prototype scalar (Hamiltonian) constraint under considerations. The auxiliary Hamiltonian **(11)** is a simplified version of the *Master Constraint* introduced in LQG, being a square function of constraints [see Ref Thiemann (2006)].

There are, however, technical limitations in the implementation of the procedure proposed above. This is because, in the D-Wave annealer, only quadratic Hamiltonian functions are allowed. This implies that the scalar constraint cannot be higher than linear order in the spin variables. On the other hand, scalar constraints being of the higher than linear order in the spin variables is expected in the full LQG.

The most general type of the constraint that one consider in the context of D-Wave quantum computer is

$$C = \sum_{i=1}^{N} a_i s_i - c \approx 0, \tag{12}$$

with some parameters $a_i, c \in \mathbb{R}^2$ and where $s_i \in \{-1, 1\}$ are classical spin variables.

Here, for the sake of simplicity we will consider the case with $a_i = 1\ \forall i$, such that the prototype scalar constraint **(12)** takes the following form:

$$C = \sum_{i=1}^{N} s_i - c \approx 0, \tag{13}$$

with some parameter $c \in \{-N, -N+2, \ldots, N-2, N\}$. By squaring **(13)** we obtain:

$$C^2 = 2\left(\sum_{\langle i,j\rangle}^{N} s_i s_j + \sum_{i=1}^{N} (-c) s_i\right) + N + c^2. \tag{14}$$

From this one can propose that the Hamiltonian to be considered is:

$$H = \frac{C^2 - N - c^2}{2} = \sum_{\langle i,j\rangle}^{N} s_i s_j + h \sum_{i=1}^{N} s_i, \tag{15}$$

where $h = -c$. The obtained Hamiltonian corresponds to the QUBO problem with a complete graph and equal couplers between the qubits ($b_{ij} = 1$). In this model, the ground state (which corresponds to $C = 0$) energy is:

$$H_0 = -\frac{N + c^2}{2}. \tag{16}$$

There is one more important aspect illustrated by the model - a degeneracy of the ground state. Namely, there are, in general, multiple spin configurations which are minimizing the Hamiltonian (15). In the model under consideration, the vacuum degeneracy depends on both c and N. Given the c and N, finding the order of degeneracy is a combinatorial problem which can be reduced to determining the number of $(N + c)/2$

[1]Here, for simplicity, we define the kinematical phase space such that is obtained by solving Gauss and diffeomorphism constraints. This corresponds to all possible spin configurations at the nodes of a given 4-valent spin network. In the quantum theory, the kinematical Hilbert space H_{kin} with respect to the scalar constraint $\hat{C}|\phi\rangle \approx 0$ is a tensor product of qubit Hilbert spaces defined at N nodes of the spin network.

[2]Alternatively, one can consider a complex constraint $C = \sum_{i=1}^{N} z_i s_i - c \approx 0$ with $z_i, c \in \mathbb{C}$. Then, in order to obtain a real Hamiltonian on has to consider $H \propto |C|^2$. This can be extended further to the case of multi-constraint model, which is not discuss here.

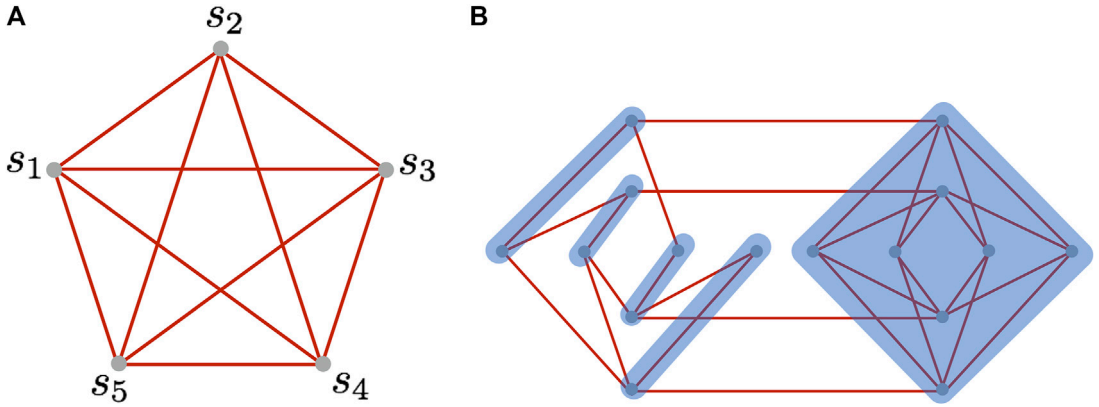

FIGURE 9 | (**A**) Pentagram spin network with $N = 5$. Each link of the graph is labeled with $j = 1/2$. The qubits (here modeled by the classical bits s_i) are defined at the nodes. **(B)** An exemplary embedding of the pentagram spin network on the two neighbor blocks of the D-Wave processor. The shadowed regions represent collective (chain) qubits which correspond to the nodes of the spin network.

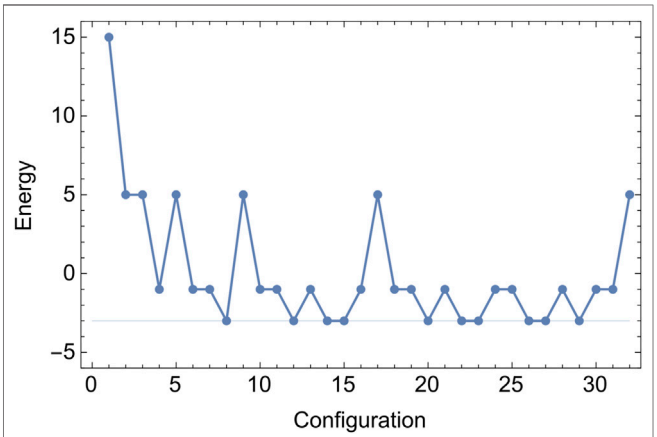

FIGURE 10 | Energy landscape of the pentagram spin network with $N = 5$ and $c = -1$. The ground state configurations satisfy the prototype scalar constraint **(13)** and span physical phase space Γ_{phys} of the model.

subset of a set composed of N elements, which is given by the binomial coefficient:

$$\binom{N}{\frac{N+c}{2}}. \tag{17}$$

Then, for a fixed N the maximal degeneracy is obtained for the choices

$$c = 2\lfloor N/2 \rfloor - N \vee c = 2\lceil N/2 \rceil - N, \tag{18}$$

where $\lfloor x \rfloor$ and $\lceil x \rceil$ are floor and ceiling functions respectively. Based on this, the corresponding maximal degeneracy is equal to

$$\binom{N}{\lfloor N/2 \rfloor} = \binom{N}{\lceil N/2 \rceil}. \tag{19}$$

The degeneracy is an essential quantity because it corresponds to the number of physical states satisfying the constraint **(13)**.

One can now make relation with the spin networks. For this purpose, let us recall that the spin Hamiltonian **(15)** corresponding to the constraint **(13)** describes a complete graph. Associating a spin coupler with a single link of the spin network, one can conclude that for the 4-valent nodes under considerations, the only complete spin network must have *pentagram* structure with $N = 5$ nodes (see block "a" in **Figure 9**)[3].

Such spin network corresponds to the geometry of a three-sphere. Furthermore, it turns out that introducing composite (chain) spins the pentagram spin network can be implemented using two neighbor blocks of the D-Wave processor. There are various ways to do so. One of them is presented in block "b" in **Figure 9**, where the shadowed regions correspond to the effective qubits composed of the elementary ones.

Finally, let us take a look at the energy landscape of the model. In **Figure 10** we plot energies corresponding to all of the spin configurations for $c = -1$.

The total number of spin configurations corresponds to dimensionality of the kinematical space: $\dim \Gamma_{kin} = 2^5 = 32$. On the other hand, the degeneracy of the vacuum of **(15)** gives us dimensionality of the physical space $\dim \Gamma_{phys} = \binom{5}{\frac{5-1}{2}} = 10$ (here we used **Eq. 17**). The physical space is a subset of kinematical space $\Gamma_{phys} \subset \Gamma_{kin}$ as expected.

In order to extract all the physical states with the use of adiabatic quantum simulations, the quantum annealing procedure has to be performed repeatedly. The outcome is a superposition of the ground states, and the procedure of measurement should select the particular ground states in the independent runs. However, as

[3]The restriction is because, in the considered case, all couplers have equal value. Therefore, the couplers have to be associated with the same number of links in the spin network. The simplest example is when a single coupler is associated with a single link. However, extensions to the other cases are possible if the general form of the constraint **(12)** is considered.

discussed in Refs (Matsuda et al., 2009; Mandrà et al., 2017) the type of quantum annealing procedure used in the D-Wave quantum computers may not be suited to identify all degenerate ground states. The studies suggest that an extension beyond the currently employed base Hamiltonians is needed to ensure that the ground state manifold is adequately sampled. Otherwise, the probability of finding some of the possible ground states may be suppressed.

Assuming that the physical states are selected, analysis of fluctuations of various observables is possible to perform. In the case under consideration, one of the interesting possibilities would be to investigate volume fluctuations. As we already mentioned, the base states corresponding to the 4-valent note qubits are eigenstates of the volume operator describing the same absolute volume but with different signs. It is, however, expected that in the classical limit, only one type of contribution would dominate such that averaged nonvanishing space volume will emerge. On the other hand, in a strongly quantum state, the positive and negative contributions can subtract one another, leading to the lack of the notion of classical geometry. Analysis of correlations of the spins in the physical states could, therefore, tell us whether domains of the same sign of volume are formed. If yes, that would be a sign of the emergence of semi-classical spacetime. Furthermore, the presence of long-range correlation would unavoidably allow associating a notion of length scale to the spin network configurations. Such observation would be a significant step toward the reconstruction of classical spacetime directly from the spin network states.

5 SUMMARY

The purpose of this article was to investigate the possibility of the implementation of spin networks on the architecture of commercially accessible adiabatic quantum computers (quantum annealers). In the studies, we focused our attention on spin networks with fixed spin labels ($j = \frac{1}{2}$) corresponding to the fundamental representation of the $SU(2)$ group. In such a case, the 4-valent nodes give rise to two-dimensional intertwiner space associated with the qubit Hilbert space. In the geometric picture, the 4-valent nodes of a spin network are dual to the 3D simplices, and the qubit bases states represent different orientations of a 3-simplex.

We have shown that it is possible to define spin networks on the architecture of the D-Wave quantum processor in the considered case. However, due to topological restrictions of the Chimera graph, not all spin networks are possible to implement with the use of elementary qubits. However, some obstacles can be overcome by introducing effective (chain) qubits composed of two or more elementary qubits. Such effective qubits allow implementing, e.g., regular 2D square lattice topology of the nearest neighbor type of interaction Ising model.

Furthermore, we proposed a method of solving scalar (Hamiltonian) constraints using quantum annealing. In the case of D-Wave quantum annealers, we have shown that a prototype constraint being a linear function of qubit variables, is possible to solve. The solutions of the constraint (i.e., physical states) are obtained as ground states of an appropriate Ising-type Hamiltonian. The procedure has been theoretically demonstrated for the pentagram spin network, which (as we have shown) can be embedded onto the architecture of the D-Wave processor. This opens a path to simulate simplified LQG models on available quantum annealers. However, one has to keep in mind that computational complexity of the approach not been addressed yet. Consequently, it is not known whether quantum annealers may provide the *quantum speed-up* for solving the LQG-related scalar constraints.

Various generalizations of the investigated class of spin networks are to be considered. In particular, the situation which is motivated by the semi-classical limit is when all spin labels are equal to some $j \gg \frac{1}{2}$, instead of $j = \frac{1}{2}$. In the case of arbitrary j, the dimension of the intertwiner space of a single 4-vertex is

$$\dim \mathrm{Inv}\left(H_j \otimes H_j \otimes H_j \otimes H_j\right) = 2j + 1. \tag{20}$$

Generalization to the case of higher than 4-valence of nodes can also be considered. In both cases, ancillary qubits have to be introduced appropriately, which is a more difficult task or, in some cases, perhaps even not possible to do. These and other issues related to quantum simulations of spin networks, especially in the context of LQG will be the subject of our further studies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

Author was supported from the Sonata Grant DEC-2014/13/D/ST2/01895 and from the Sonata Bis Grant DEC-2017/26/E/ST2/00763 of the National Science Centre Poland, and the Mobilność Plus Grant 1641/MON/V/2017/0 of the Polish Ministry of Science and Higher Education. Furthermore, this publication was made possible through the support of the ID# 61466 grant from the John Templeton Foundation, as part of the "The Quantum Information Structure of Spacetime (QISS)" Project (qiss.fr).

ACKNOWLEDGMENTS

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. Author would like to thank to Mario Flory for his careful reading of the manuscript and helpful comments.

REFERENCES

Agullo, I., Ashtekar, A., and Nelson, W. (2012). Quantum Gravity Extension of the Inflationary Scenario. *Phys. Rev. Lett.* 109, 251301. [arXiv:1209.1609 [gr-qc]]. doi:10.1103/physrevlett.109.251301

Ashtekar, A., and Lewandowski, J. (2004). Background Independent Quantum Gravity: A Status Report. *Class. Quan. Grav.* 21, R53–R152. [gr-qc/0404018]. doi:10.1088/0264-9381/21/15/r01

Barbero G., J. F. (1995). Real Ashtekar Variables for Lorentzian Signature Space-Times. *Phys. Rev. D* 51, 5507–5510. [arXiv:gr-qc/9410014 [gr-qc]]. doi:10.1103/physrevd.51.5507

Barrau, A., Cailleteau, T., Grain, J., and Mielczarek, J. (2014). Observational Issues in Loop Quantum Cosmology. *Class. Quan. Grav.* 31, 053001. [arXiv:1309.6896 [gr-qc]]. doi:10.1088/0264-9381/31/5/053001

Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., and Lloyd, S. (2017). Quantum Machine Learning. *Nature* 549, 195–202. doi:10.1038/nature23474

Campbell, E. T., Terhal, B. M., and Vuillot, C. (2017). Roads towards Fault-Tolerant Universal Quantum Computation. *Nature* 549, 172–179. doi:10.1038/nature23460

Czelusta, G., and Mielczarek, J. (2021). Quantum Simulations of a Qubit of Space. *Phys. Rev. D* 103 (4), 046001. [arXiv:2003.13124 [gr-qc]]. doi:10.1103/physrevd.103.046001

Feller, A., and Livine, E. R. (2016). Ising Spin Network States for Loop Quantum Gravity: a Toy Model for Phase Transitions. *Class. Quan. Grav.* 33 (6), 065005. [arXiv:1509.05297 [gr-qc]]. doi:10.1088/0264-9381/33/6/065005

Feynman, R. P. (1982). Simulating Physics with Computers. *Int. J. Theor. Phys.* 21, 467–488. doi:10.1007/bf02650179

Harris, R., Johansson, J., Berkley, A. J., Johnson, M. W., Lanting, T., Han, S., et al. (2010). Experimental Demonstration of a Robust and Scalable Flux Qubit. *Phys. Rev. B* 81, 134510. doi:10.1103/physrevb.81.134510

Immirzi, G. (1997). Quantum Gravity and Regge Calculus. *Nucl. Phys. B - Proc. Supplements* 57, 65–72. [arXiv:gr-qc/9701052 [gr-qc]]. doi:10.1016/s0920-5632(97)00354-x

Jordan, S. P. (2010). Permutational Quantum Computing. *Quan. Inf. Comput.* 10, 470–497. doi:10.26421/qic10.5-6-7

Kadowaki, T., and Nishimori, H. (1998). Quantum Annealing in the Transverse Ising Model. *Phys. Rev. E* 58, 5355–5363. doi:10.1103/physreve.58.5355

Kauffman, L. H., and Lomonaco, S. J., Jr. (2008). Spin Networks and Quantum Computation. *Bulg. J. Phys.* 35, 241–256.

Li, K., Li, Y., Han, M., Lu, S., Zhou, J., Ruan, D., et al. (2019). Quantum Spacetime on a Quantum Simulator. *Commun. Phys.* 2, 122. [arXiv:1712.08711 [quant-ph]]. doi:10.1038/s42005-019-0218-5

Mandrà, S., Zhu, Z., and Katzgraber, H. G. (2017). Exponentially Biased Ground-State Sampling of Quantum Annealing Machines with Transverse-Field Driving Hamiltonians. *Phys. Rev. Lett.* 118, 070502. doi:10.1103/PhysRevLett.118.070502

Marzuoli, A., and Rasetti, M. (2005). Computing Spin Networks. *Ann. Phys.* 318, 345–407. doi:10.1016/j.aop.2005.01.005

Matsuda, Y., Nishimori, H., and Katzgraber, H. G. (2009). Ground-state Statistics from Annealing Algorithms: Quantum versus Classical Approaches. *New J. Phys.* 11, 073021. doi:10.1088/1367-2630/11/7/073021

Mielczarek, J. (2019). Spin Foam Vertex Amplitudes on Quantum Computer - Preliminary Results. *Universe* 5 (8), 179. [arXiv:1810.07100 [gr-qc]]. doi:10.3390/universe5080179

Rovelli, C. (1998). Loop Quantum Gravity. *Living Rev. Relativ.* 1, 1. [arXiv:gr [gr-qc]]. doi:10.12942/lrr-1998-1

Rovelli, C., and Smolin, L. (1990). Loop Space Representation of Quantum General Relativity. *Nucl. Phys. B* 331, 80–152. doi:10.1016/0550-3213(90)90019-a

Rovelli, C., and Vidotto, F. (2014). *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory*. Cambridge, United Kingdom: Cambridge Monographs on Mathematical Physics.

Thiemann, T. (2006). The Phoenix Project: Master Constraint Programme for Loop Quantum Gravity. *Class. Quan. Grav.* 23, 2211–2247. [gr-qc/0305080]. doi:10.1088/0264-9381/23/7/002

You, J. Q., and Nori, F. (2005). Superconducting Circuits and Quantum Information. *Phys. Today* 58 (11), 42–47. doi:10.1063/1.2155757

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 18 June 2021
doi: 10.3389/fphy.2021.587083

Equivalence of Approaches to Relational Quantum Dynamics in Relativistic Settings

Philipp A. Höhn[1,2]*, *Alexander R. H. Smith*[3,4]* and *Maximilian P. E. Lock*[5]*

[1] Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan, [2] Department of Physics and Astronomy, University College London, London, United Kingdom, [3] Department of Physics, Saint Anselm College, Manchester, NH, United States, [4] Department of Physics and Astronomy, Dartmouth College, Hanover, NH, United States, [5] Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Vienna, Austria

OPEN ACCESS

Edited by:
Daniele Oriti,
Ludwig-Maximilians-University Munich, Germany

Reviewed by:
Mehdi Assanioussi,
University of Warsaw, Poland
Zhenbin Wu,
University of Illinois at Chicago, United States
Steffen Gielen,
The University of Sheffield, United Kingdom

***Correspondence:**
Philipp A. Höhn
philipp.hoehn@oist.jp
Alexander R. H. Smith
arhsmith@anselm.edu
Maximilian P. E. Lock
maximilian.lock@oeaw.ac.at

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: 24 July 2020
Accepted: 22 March 2021
Published: 18 June 2021

Citation:
Höhn PA, Smith ARH and Lock MPE (2021) Equivalence of Approaches to Relational Quantum Dynamics in Relativistic Settings. Front. Phys. 9:587083. doi: 10.3389/fphy.2021.587083

We have previously shown that three approaches to relational quantum dynamics—relational Dirac observables, the Page-Wootters formalism and quantum deparametrizations—are equivalent. Here we show that this "trinity" of relational quantum dynamics holds in relativistic settings per frequency superselection sector. Time according to a clock subsystem is defined via a positive operator-valued measure (POVM) that is covariant with respect to the group generated by its (quadratic) Hamiltonian. This differs from the usual choice of a self-adjoint clock observable conjugate to the clock momentum. It also resolves Kuchař's criticism that the Page-Wootters formalism yields incorrect localization probabilities for the relativistic particle when conditioning on a Minkowski time operator. We show that conditioning instead on the covariant clock POVM results in a Newton-Wigner type localization probability commonly used in relativistic quantum mechanics. By establishing the equivalence mentioned above, we also assign a consistent conditional-probability interpretation to relational observables and deparametrizations. Finally, we expand a recent method of changing temporal reference frames, and show how to transform states and observables frequency-sector-wise. We use this method to discuss an indirect clock self-reference effect and explore the state and temporal frame-dependence of the task of comparing and synchronizing different quantum clocks.

Keywords: relational quantum dynamics, problem of time, relational Dirac observables, Page-Wootters formalism, quantum deparametrizations, quantum clocks, quantum reference frames, relativistic quantum clocks

1. INTRODUCTION

In general relativity, time plays a different role than in classical and quantum mechanics, or quantum field theory on a Minkowski background. General covariance dispenses with a preferred choice of time and introduces instead a dynamical notion of time which depends on solutions to the Einstein field equations. In the canonical approach to quantum gravity this leads to the infamous *problem of time* [1–3]. Its most well-known facet is that, due to the constraints of the theory, quantum states of spacetime (and any matter contained in it) do not at first sight appear to undergo any time evolution, in seeming contradiction with everyday experience.

The resolution comes from one of the key insights of general relativity: any physical notion of time is relational, the degrees of freedom of the Universe evolve relative to one another [4–6]. This

insight has led to three main relational approaches to the problem of time, each of which seeks to extract a notion of time from within the quantum degrees of freedom, relative to which the others evolve:

(i) a Dirac quantization scheme, wherein relational observables are constructed that encode correlations between evolving and clock degrees of freedom [1, 2, 4, 7–38],
(ii) the Page-Wootters formalism, which defines a relational dynamics in terms of conditional probabilities for clock and evolving degrees of freedom [7, 25, 39–57], and
(iii) classical or quantum deparametrizations, which result in a reduced quantum theory that only treats the evolving degrees of freedom as quantum [1, 2, 7, 10, 30, 31, 58–65].

These three approaches have been pursued largely independently with the relation between them previously unknown. They have also not been without criticism, especially the Page-Wootters formalism. For example, Kuchař [1] raised three fundamental criticisms against this approach, namely that it:

(a) leads to wrong localization probabilities in relativistic settings,
(b) is in conflict with the constraints of the theory, and
(c) yields wrong propagators.

Concern has also been voiced that there is an inherent ambiguity in terms of which clock degrees of freedom one should choose, also known as the *multiple choice problem* [1–3, 66, 67]. Indeed, in generic general relativistic systems there is no preferred choice of relational time variable and different choices may lead to *a priori* different quantum theories.

In our recent work [7] we addressed the relation between these three approaches (i)–(iii) to relational quantum dynamics, demonstrating that they are, in fact, equivalent when the clock Hamiltonian features a continuous and non-degenerate spectrum and is decoupled from the system whose dynamics it is used to describe. Specifically, we constructed the explicit transformations mapping each formulation of relational quantum dynamics into the others. These maps revealed the Page-Wootters formalism (ii) and quantum deparametrizations (iii) as *quantum symmetry reductions* of the manifestly gauge-invariant formulation (i). In other words, the Page-Wootters formalism (ii) and quantum deparametrizations (iii) can be regarded as quantum analogs of gauge-fixed formulations of gauge-invariant quantities (i). Conversely, the formulation in terms of relational Dirac observables (i) constitutes the quantum analog of a gauge-invariant extension of the gauge-fixed formulations (ii) and (iii). More physically, these transformations establish (i) as a clock-choice-neutral (in a sense explained below), (ii) as a relational Schrödinger, and (iii) as a relational Heisenberg picture of the dynamics. Constituting three faces of the same quantum dynamics, we called the equivalence of (i)–(iii) the *trinity of relational quantum dynamics*.

This equivalence not only provides relational Dirac observables with a consistent conditional probability interpretation, but also resolves Kuchař's criticism (b) that the Page-Wootters formalism would be in conflict with the quantum constraints. Furthermore, the trinity resolves Kuchař's criticism (c) that the Page-Wootters formalism would yield wrong propagators, by showing that the correct propagators always follow from manifestly gauge-invariant conditional probabilities on the physical Hilbert space [7]. This resolution of criticism (c) differs from previous resolution proposals which relied on ideal clocks [25, 43, 68] and auxiliary ancilla systems [43] and can be viewed as an extension of [46].

The transformations between (i)–(iii) of the trinity also allowed us to address the multiple choice problem in Höhn et al. [7] by extending a previous method for changing temporal reference frames, i.e., clocks, in the quantum theory [30, 31, 47] (see also [32–34, 69]). The resolution to the problem lies in part in realizing that a solution to the Wheeler-DeWitt equation encodes the relations between all subsystems, including the relations between subsystems employed as clocks to track the dynamics of other subsystems; there are multiple choices of clocks, each of which can be used to define dynamics. Our proposal is thus to turn the multiple choice problem into a *feature* by having a multitude of quantum time choices at our disposal, which we are able to connect through quantum temporal frame transformations. This is in line with developing a genuine quantum implementation of general covariance [7, 30, 31, 38, 70–74]. This proposal is part of current efforts to develop a general framework of quantum reference frame transformations (and study their physical consequences [75–85]), and should be contrasted with other attempts at resolving the multiple choice problem by identifying a preferred choice of clock [53] (see [7] for further discussion of this proposal).

We did not address Kuchař's criticism (a) that the Page-Wootters formalism yields the wrong localization probabilities for relativistic models in Höhn et al. [7] as they feature clock Hamiltonians which are quadratic in momenta and thus generally have a degenerate spectrum, splitting into positive and negative frequency sectors. This degeneracy is not covered by our previous construction. While quadratic clock Hamiltonians are standard in the literature on relational observables [approach (i)] and deparametrizations [approach (iii)], see e.g., [4, 10, 11, 19, 29, 31], relativistic particle models have only recently been studied in the Page-Wootters formalism [approach (ii)] [45, 49–51]. However, Kuchař's criticism (a) that the Page-Wootters approach yields incorrect localization probabilities in relativistic settings has yet to be addressed. Since the Page-Wootters formalism encounters challenges in relativistic settings, given the equivalence of relational approaches implied by the trinity, one might worry about relational observables and deparametrizations too.

In this article, we show that these challenges can be overcome, and a consistent interpretation of the relational dynamics can be provided. To this end, we extend the trinity to quadratic clock Hamiltonians, thus encompassing many relativistic settings; we show that all the results of Höhn et al. [7] hold per frequency sector associated to the clock due to a superselection rule induced by the Hamiltonian constraint. Frequency-sector-wise, the relational dynamics encoded in (i) relational observables, (ii) the

Page-Wootters formalism, and (iii) quantum deparametrizations are thus also fully equivalent.

The key to our construction, as in Höhn et al. [7], is the use of a Positive-Operator Valued Measure (POVM) which here transforms covariantly with respect to the quadratic clock Hamiltonian [86–89] as a time observable. This contrasts with the usual approach of employing an operator conjugate to the clock momentum (i.e., the Minkowski time operator in the case of a relativistic particle). This covariant clock POVM is instrumental in our resolution of Kuchař's criticism (a) that the Page-Wootters formalism yields wrong localization probabilities for relativistic systems. We show that when conditioning on this covariant clock POVM rather than Minkowski time, one obtains a Newton-Wigner type localization probability [90, 91]. While a Newton-Wigner type localization is approximate and not fully Lorentz covariant, due to the relativistic localization no-go theorems of Perez-Wilde [92] and Malament [93] (see also [94, 95]), it is generally accepted as the best possible localization in relativistic quantum mechanics (In quantum field theory localization is a different matter [90, 94]). This demonstrates the advantage of using covariant clock POVMs in relational quantum dynamics [7, 44, 45, 96, 97]. The trinity also extends the probabilistic interpretation of relational observables: a Dirac observable describing the relation between a position operator and the covariant clock POVM corresponds to a Newton-Wigner type localization in relativistic settings.

Finally, we again use the equivalence between (i) and (iii) to construct temporal frame changes in the quantum theory. On account of superselection rules across frequency sectors, temporal frame changes can only map information contained in the overlap of two frequency sectors, one associated to each clock, from one clock "perspective" to another. We apply these temporal frame change maps to explore an indirect clock self-reference and the temporal frame and state dependence of comparing and synchronizing readings of different quantum clocks.

While completing this manuscript, we became aware of Chataignier [38], which independently extends some results of Höhn et al. [7] on the conditional probability interpretation of relational observables and their equivalence with the Page-Wootters formalism into a more general setting. However, a different formalism [37] is used in Chataignier [38], which does not employ covariant clock POVMs and therefore the two works complement one another.

Throughout this article we work in units where $\hbar = 1$.

2. CLOCK-NEUTRAL FORMULATION OF CLASSICAL AND QUANTUM MECHANICS

Colloquially, general covariance posits that the laws of physics are the same in every reference frame. This is usually interpreted as implying that physical laws should take the form of tensor equations. Tensors can be viewed as reference-frame-neutral objects: they define a description of physics prior to choosing a reference frame. They thereby encode the physics as "seen" by all reference frames at once. If one wants to know the numbers which a measurement of the tensor in a particular reference frame would yield, one must contract the tensor with the vectors corresponding to that choice of frame. In this way, the description of the same tensor looks different relative to different frames, but the tensor *per se*, as a multilinear map, is reference-frame-neutral. It is this reference-frame-neutrality of tensors which results in the frame-independence of physical laws.

The notion of reference frame as a vector frame is usually taken to define the orientation of a local laboratory of some observer. In practice, one often implicitly identifies the local lab (i.e., the reference system relative to which the remaining physics is described) with the reference frame. This is an idealization which ignores the lab's back-reaction on spacetime, interaction with other physical systems and possible internal dynamics, while at the same time assuming it to be sufficiently classical so that superpositions of orientations can be ignored. Such an idealization is appropriate in general relativity where the aim is to describe the large-scale structure of spacetime. However, in quantum gravity, where the goal is to describe the micro-structure of spacetime, this may no longer be appropriate [98]. More generally, we may ask about the fate of general covariance when we take seriously the fact that physically meaningful reference frames are in practice always associated with physical systems, and as such are comprised of dynamical degrees of freedom that may couple with other systems, undergo their own dynamics and will ultimately be subject to the laws of quantum theory. What are then the reference-frame-neutral structures?

In regard to this question, we note that the classical notion of general covariance for reference frames associated to idealized local labs is deeply intertwined with invariance under general coordinate transformations, i.e., passive diffeomorphisms. In moving toward non-idealized reference frames (or rather systems), we shift focus from coordinate descriptions to dynamical reference degrees of freedom, relative to which the remaining physics will be described. In line with this, we shift the focus from passive to active diffeomorphisms, which directly act on the dynamical degrees of freedom. This is advantageous for quantum gravity, where classical spacetime coordinates are *a priori* absent. A quantum version of general covariance should be formulated in terms of dynamical reference degrees of freedom [7, 30, 31, 38, 70–74].

The active symmetries imply a redundancy in the description of the physics. *A priori* all degrees of freedom stand on an equal footing, giving rise to a freedom in choosing which of them shall be treated as the redundant ones. The key idea is to identify this choice with the choice of reference degrees of freedom, i.e., those relative to which the remaining degrees of freedom will be described[1]. Accordingly, choosing a dynamical reference system amounts to removing redundancy from the description. As such, we may interpret the redundancy-containing description (in both the classical and quantum theory) as a *perspective-neutral* description of physics, i.e., as a global description of

[1]Indeed, we do not want to describe the reference degrees of freedom directly relative to themselves in order to avoid the self-reference problem [99, 100]. Nevertheless, through the perspective-neutral structure it is possible to construct indirect self-reference effects of quantum clocks through temporal frame changes, see Höhn et al. [7] and section 7.3.

physics prior to having chosen a reference system, from whose perspective the remaining degrees of freedom are to be described [7, 30, 31, 71, 72]. This perspective-neutral structure is thus proposed as the reference-frame-neutral structure for dynamical (i.e., non-idealized) reference systems.

In this article we focus purely on temporal diffeomorphisms and thus on temporal reference frames/systems, or simply clocks. In this case, we refer to the perspective-neutral structure as a *clock-(choice-)neutral* structure [7, 30, 31], which we briefly review here in both the classical and quantum theory. It is a description of the physics, prior to having chosen a temporal reference system relative to which the dynamics of the remaining degrees of freedom are to be described.

2.1. Clock-Neutral Classical Theory

Consider a classical theory described by an action $\mathcal{S} = \int_{\mathbb{R}} du\, L(q^a, dq^a/du)$, where q^a denotes a collection of configuration variables indexed by a. Such a theory exhibits *temporal diffeomorphism invariance* if the action $\mathcal{S}$ is reparametrization invariant; that is, $L(q^a, dq^a/du) \mapsto L(q^a, dq^a/du')du'/du$ transforms as a scalar density under $u \mapsto u'(u)$. The Hamiltonian of such a theory is of the form $H = N(u)\, C_H$, where $N(u)$ is an arbitrary lapse function and

$$C_H = \sum_a \frac{dq^a}{du}\, p_a - L \approx 0, \tag{1}$$

the so-called Hamiltonian constraint, is a consequence of the temporal diffeomorphism symmetry. This equation defines the constraint surface $\mathcal{C}$ inside the kinematical phase space $\mathcal{P}_{\text{kin}}$, which is parametrized by the canonical coordinates q^a, p_b. The $\approx$ denotes a weak equality, i.e., one which only holds on $\mathcal{C}$ [101, 102].

The Hamiltonian generates a dynamical flow on $\mathcal{C}$, which transforms an arbitrary phase space function f according to

$$\frac{df}{du} := \{f, C_H\} \tag{2}$$

and integrates to a finite transformation $\alpha^u_{C_H} \cdot f$, where for simplicity the lapse function has been chosen to be unity, $N(u) = 1$. Owing to the reparametrization invariance, this flow should be interpreted as a gauge transformation rather than true evolution [4, 11], and thus for an observable F to be physical, it must be invariant under such a transformation, i.e.,

$$\{F, C_H\} \approx 0. \tag{3}$$

Observables satisfying Equation (3) are known as (weak) *Dirac observables*.

In order to obtain a gauge-invariant dynamics, we have to choose a dynamical temporal reference system, i.e., a clock function $T(q^a, p_a)$, to parametrize the dynamical flow, Equation (2), generated by the constraint. We can then describe the evolution of the remaining degrees of freedom relative to $T(q^a, p_a)$. This gives rise to so-called relational Dirac observables (a.k.a. evolving constants of motion) which encode the answer to the question "what is the value of the function f along the flow generated by C_H on $\mathcal{C}$ when the clock T reads τ?" [4, 7, 11–19, 19–24, 26, 30, 31]. We will denote such an observable by $F_{f,T}(\tau)$. As shown in Dittrich [21, 22] and Dittrich and Tambornino [23, 24], these observables can be constructed by solving $\alpha^u_{C_H} \cdot T = \tau$ for u yielding the solution $u = u_T(\tau)$ and defining

$$\begin{aligned} F_{f,T}(\tau) &:= \alpha^u_{C_H} \cdot f \Big|_{u=u_T(\tau)} \\ &\approx \sum_{n=0}^{\infty} \frac{(\tau - T)^n}{n!} \left\{ f, \frac{C_H}{\{T, C_H\}} \right\}_n , \end{aligned} \tag{4}$$

where $\{f,g\}_n := \{\{f,g\}_{n-1}, g\}$ is the n^{th}-nested Poisson bracket subject to $\{f,g\}_0 := f$. The $F_{f,T}(\tau)$ satisfy Equation (3) and thus constitute a family of Dirac observables parametrized by τ. Such relational observables are so-called gauge-invariant extensions of gauge-fixed quantities [7, 21–24, 37, 102].

In generic models there is no preferred choice for the clock function T among the degrees of freedom on $\mathcal{P}_{\text{kin}}$, which is sometimes referred to as the *multiple choice problem* [1, 2]. Different choices of T will lead to different relational Dirac observables, as can be seen in Equation (4). All these different choices are encoded in the constraint surface $\mathcal{C}$ and stand *a priori* on an equal footing.

This gives rise to the interpretation of $\mathcal{C}$ as a clock-neutral structure. The temporal diffeomorphism symmetry leads to a redundancy in the description of $\mathcal{C}$: thanks to the Hamiltonian constraint the kinematical canonical degrees of freedom are not independent and due to its gauge flow there will only be $\dim \mathcal{P}_{\text{kin}} - 2$ independent physical phase space degrees of freedom. In particular, relative to any choice of clock function T one can construct $\dim \mathcal{P}_{\text{kin}} - 2$ independent relational Dirac observables using Equation (4) [101, 102]. Hence, the relational Dirac observables relative to any other clock choice T' can be constructed from them. Consequently, there is redundancy among the relational Dirac observables relative to different clock choices. Thus $\mathcal{C}$ yields a description of the physics prior to choosing and fixing a clock relative to which the gauge-invariant dynamics of the remaining degrees of freedom can be described. Specifically, no choice has been made as to which of the kinematical and physical degrees of freedom are to be considered as redundant. In analogy to the tensor case, $\mathcal{C}$ still contains the information about all clock choices and their associated relational dynamics at once; it yields a clock-neutral description.

Being of odd dimension $\dim \mathcal{P}_{\text{kin}} - 1$, $\mathcal{C}$ is also not a phase space. A proper phase space description can be obtained, e.g., through phase space reduction by gauge-fixing [7, 30, 31, 38, 58]. Given a choice of clock function T, we may consider the gauge-fixing condition $T = const$, which may be valid only locally on $\mathcal{C}$. Since $F_{f,T}(\tau)$ is constant along each orbit generated by C_H for each value of τ, we do not lose any information about the relational dynamics by restricting to $T = const$ and leaving τ free. By restricting to the relational observables $F_{f,T}(\tau)$ relative to clock T and by solving the two conditions $T = const$, $C_H = 0$, we remove the redundancy from among both the kinematical and physical degrees of freedom. The surviving reduced phase space description, which no longer contains the clock degrees of freedom as dynamical variables, can be interpreted as the

description of the dynamics relative to the temporal reference system defined by the clock function T. But now we keep track of time evolution not in terms of the dynamical T, but in terms of the parameter τ representing its "clock readings." In particular, the temporal reference system is not described relative to itself, e.g., one finds the tautology $F_{T,T}(\tau) \approx \tau$. Accordingly, choosing the "perspective" of a clock means choosing the corresponding clock degrees of freedom as the redundant ones and removing them. The theory is then deparametrized: it no longer contains a gauge-parameter u, nor a constraint, nor dynamical clock variables—only true evolving degrees of freedom.

2.2. Clock-Neutral Quantum Theory

Following the Dirac prescription for quantizing constrained systems [10, 11, 101, 102], one first promotes the canonical coordinates of $\mathcal{P}_{\text{kin}}$ to canonical position and momentum operators $\hat{q}^a$ and $\hat{p}_a$ acting on a *kinematical Hilbert space* $\mathcal{H}_{\text{kin}}$. The Hamiltonian constraint in Equation (1) is then imposed by demanding that physical states of the quantum theory are annihilated by the quantization of the constraint function

$$\hat{C}_H \, |\psi_{\text{phys}}\rangle = 0. \tag{5}$$

Solutions to this Wheeler-DeWitt-like equation may be constructed from kinematical states $|\psi_{\text{kin}}\rangle \in \mathcal{H}_{\text{kin}}$ via a group averaging operation [11, 103–107][2]

$$|\psi_{\text{phys}}\rangle = \delta(\hat{C}_H) \, |\psi_{\text{kin}}\rangle = \frac{1}{2\pi} \int_G du \, e^{-i\hat{C}_H u} \, |\psi_{\text{kin}}\rangle \,, \tag{6}$$

where G is the group generated by $\hat{C}_H$. Physical states are not normalizable in $\mathcal{H}_{\text{kin}}$ if they are improper eigenstates of $\hat{C}_H$ (i.e., if zero lies in the continuous part of its spectrum). However, they are normalized with respect to the so-called *physical inner product*

$$\langle \psi_{\text{phys}} | \phi_{\text{phys}} \rangle_{\text{phys}} := \langle \psi_{\text{kin}} | \delta(\hat{C}_H) | \phi_{\text{kin}} \rangle_{\text{kin}} \tag{7}$$

where $\langle \cdot | \cdot \rangle_{\text{kin}}$ is the kinematical inner product and $|\psi_{\text{kin}}\rangle, |\phi_{\text{kin}}\rangle \in \mathcal{H}_{\text{kin}}$ reside in the equivalence class of states mapped to the same $|\psi_{\text{phys}}\rangle, |\phi_{\text{phys}}\rangle$ under the projection in Equation (6). Equipped with this inner product, the space of solutions to the Wheeler-DeWitt equation in Equation (5) can usually be Cauchy completed to form the so-called *physical Hilbert space* $\mathcal{H}_{\text{phys}}$ [11, 103–107].

A gauge-invariant (i.e., physical) observable $\hat{F}$ acting on $\mathcal{H}_{\text{phys}}$ must satisfy the quantization of Equation (2)

$$\left[\hat{C}_H, \hat{F}\right] |\psi_{\text{phys}}\rangle = 0. \tag{8}$$

Such an observable $\hat{F}$ is a *quantum Dirac observable*.

Clearly, $\exp(-i\,u\,\hat{C}_H) \, |\psi_{\text{phys}}\rangle = |\psi_{\text{phys}}\rangle$, i.e., physical states do not evolve under the dynamical flow generated by the Hamiltonian constraint. This is the basis of the so-called *problem of time* in quantum gravity [1–3], and of statements that a quantum theory defined by a Hamiltonian constraint is timeless. However, such a theory is only "background-timeless," i.e., physical states do not evolve with respect to the "external" gauge parameter u parametrizing the group generated by the Hamiltonian constraint. Instead, it is more appropriate to regard the quantum theory on $\mathcal{H}_{\text{phys}}$ as a *clock-neutral* quantum theory: it is a global description of the physics prior to choosing an internal clock relative to which to describe the dynamics of the remaining degrees of freedom (as argued in [7, 30, 31]). Just as in the classical case, there will in general be many possible clock choices and the "quantum constraint surface" $\mathcal{H}_{\text{phys}}$ contains the information about all these choices at once; it is thus by no means "internally timeless."

The goal is to suitably quantize the relational Dirac observables in Equation (4), promoting them to families of operators $\hat{F}_{f,T}(\tau)$ on $\mathcal{H}_{\text{phys}}$. This involves a quantization of the temporal reference system T and it is clear that in the quantum theory different choices of T will also lead to different quantum relational Dirac observables. This will give rise to a multitude of gauge-invariant, relational quantum dynamics, each expressed with respect to the evolution parameter τ, which corresponds to the readings of the chosen quantum clock (and is thus *not* a gauge parameter). The quantization of relational observables is non-trivial, especially because Equation (4) may not be globally defined on $\mathcal{C}$, and depends very much on the properties of the chosen clock. Steps toward systematically quantizing relational Dirac observables have been undertaken (e.g., in [7, 17, 20, 37, 38]) and part of this article is devoted to further developing them for a class of relativistic models.

In analogy to the classical case, the clock-neutral description on the "quantum constraint surface" $\mathcal{H}_{\text{phys}}$ is redundant: since the constraint is satisfied, not all the degrees of freedom are independent. In particular, the sets of quantum relational Dirac observables relative to different clock choices—and thus different relational quantum dynamics—will be interdependent. The proposal is once more to associate the choice of clock with the choice of redundant degrees of freedom; moving to the "perspective" of a given clock means considering the quantum relational observables relative to it as the independent ones, and removing the (now redundant) dynamical clock degrees of freedom altogether. This works through a quantum symmetry reduction procedure, i.e., the quantum analog of phase space reduction, which is tantamount to a *quantum deparametrization* and has been developed in Höhn and Vanrietvelde [30], Höhn [31], and Höhn et al. [7] and will be further developed in section 5. In particular, this procedure is at the heart of changing from a description relative to one quantum clock to one relative to another clock, which we elaborate on in section 7. As such, quantum symmetry reduction is the key element of a proposal for exploring a quantum version of general covariance [7, 30, 31, 47, 70–72] and thereby also addressing the multiple choice problem in quantum gravity and cosmology [1, 2] (see also [28, 32–34, 64, 65, 69, 77, 108]).

[2] In contrast to [11, 103–107] and for notational simplicity, we refrain from using the more rigorous formulation in terms of Gel'fand triples and algebraic duals of (dense subsets of) Hilbert spaces. However, the remainder of this article could be put into such a more precise formulation.

3. QUADRATIC CLOCK HAMILTONIANS

Building upon the clock-neutral discussion, we now assume that the kinematical degrees of freedom described by $\mathcal{P}_{\rm kin}$ and $\mathcal{H}_{\rm kin}$ split into a clock C and an "evolving" system S, which do not interact. This will permit us to choose a temporal reference system in the next section, and thence define a relational dynamics in both the classical and quantum theories.

3.1. Classical Theory

Suppose the classical theory describes a clock C associated with the phase space $\mathcal{P}_C \simeq T^*\mathbb{R} \simeq \mathbb{R}^2$, and some system of interest S associated with a phase space $\mathcal{P}_S$, so that the kinematical phase space decomposes as $\mathcal{P}_{\rm kin} := \mathcal{P}_C \times \mathcal{P}_S$. We assume $\mathcal{P}_C$ to be parametrized by the canonical pair (t, p_t), but will not need to be specific about the structure of $\mathcal{P}_S$ (other than assuming it to be a finite dimensional symplectic manifold). Further suppose that the clock and system are not coupled, leading to a Hamiltonian constraint function that is a sum of their respective Hamiltonians[3]

$$C_H = H_C + H_S \approx 0, \tag{9}$$

where H_C is a function on $\mathcal{P}_C$ and H_S is a function on $\mathcal{P}_S$.

This article concerns clock Hamiltonians that are quadratic in the clock momentum, $H_C = s\,p_t^2/2$, where $s \in \{-1, +1\}$, so that the Hamiltonian constraint becomes

$$C_H = s\,\frac{p_t^2}{2} + H_S \approx 0. \tag{10}$$

This class of clock Hamiltonians appears in a wide number of (special and general) relativistic and non-relativistic models—see **Table 1** for examples. They are doubly degenerate; every value of H_C has two solutions in terms of p_t, except on the line defined by $p_t = 0$. Note that p_t is a Dirac observable.

The constraint in Equation (10) can be factored into two constraints, each linear in p_t [30, 31, 33]:

$$C_H = s\,C_+ \cdot C_-\,, \quad \text{for} \quad C_\sigma := \frac{p_t}{\sqrt{2}} + \sigma\,\sqrt{-s\,H_S}\,, \tag{11}$$

where we have introduced the degeneracy label $\sigma = \pm 1$. Note that Equation (10) forces $s\,H_S$ to take non-positive values on $\mathcal{C}$. In the $s = -1$ case, $\sigma = +1$ and $\sigma = -1$ define the positive and negative frequency modes in the quantum theory, respectively. For simplicity, we shall henceforth use this terminology for *both* $s = \pm 1$[4]. It follows that we can decompose the constraint surface into a positive and a negative frequency sector [30, 31]

$$\mathcal{C} = \mathcal{C}_+ \cup \mathcal{C}_-\,, \tag{12}$$

where $\mathcal{C}_\sigma$ is the set of solutions to $C_\sigma = 0$ in $\mathcal{P}_{\rm kin}$. The intersection $\mathcal{C}_+ \cap \mathcal{C}_-$ is defined by $p_t = H_S = 0$ (see **Figure 1** for an illustration).

[3]The assumption that clock and system do not interact does not hold in generic general relativistic systems. However, it is satisfied in some commonly used examples (see [7] for a discussion and **Table 1** for some examples).

[4]Hence, positive and negative frequency modes are defined by $p_t < 0$ and $p_t > 0$, respectively, which follows from setting $C_\sigma = 0$.

TABLE 1 | Examples of constraints of the form in Equation (10).

Non-relativistic particle and arbitrary system
$C_H = \frac{\boldsymbol{p}^2}{2m} + H_S$
Relativistic particle in inertial coordinates
$C_H = -p_t^2 + \boldsymbol{p}^2 + m^2$
Homogeneous isotropic cosmology with massless scalar field
$C_H = p_\phi^2 - p_\alpha^2 - 4k\,\exp(4\alpha)$
Homogeneous cosmology (vacuum Bianchi models)
$C_H = -\frac{1}{2}\bar{p}_0^2 + k_0\,\exp(2\sqrt{2}\bar{\beta}^0) + \frac{1}{2}p_+^2 + k_+\,\exp(-4\sqrt{3}\bar{\beta}^+)\; + \frac{k_-}{2}p_-^2$

Some examples of constraints of the form of Equation (10), i.e., with clock Hamiltonians quadratic in an appropriate canonical momentum. The last three (relativistic) examples each contain both cases $s = \pm 1$, depending on which degree of freedom is used to define the clock C. In the example of the Friedman-Lemaître-Robertson-Walker model with homogeneous massless scalar field we have used $\alpha := \ln a$, where a is the scale factor, and k is the spatial curvature constant [109–115] (here a choice of lapse function $N = e^{3\alpha}$ has been made and included in the definition of C_H). The shape of the Hamiltonian constraint for vacuum Bianchi models can be found (e.g., in [116]), and holds for types I, II, III, VIII, IX, and the Kantowski-Sachs models. Here $\bar{\beta}^0, \bar{\beta}^+, \bar{\beta}^-$ are linear combinations of the Misner anisotropy parameters and k_0, k_+, k_- are constants, each of which may be zero, depending on the model.

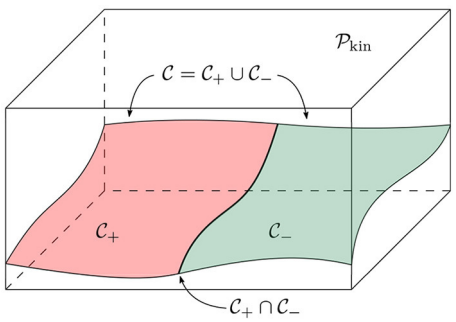

FIGURE 1 | Depicted are the surfaces $\mathcal{C}_+$ (red) and $\mathcal{C}_-$ (green) defined by $C_+ = 0$ and $C_- = 0$, respectively. The union of these surfaces is the constraint surface $\mathcal{C} = \mathcal{C}_+ \cup \mathcal{C}_- \subset \mathcal{P}_{\rm kin}$, while their intersection $\mathcal{C}_+ \cap \mathcal{C}_-$ is characterized by $p_t = H_S = 0$, and is depicted by the thick black line. We have assumed that H_S is not degenerate (see Figure 1 of [31] for a similar depiction when H_S is doubly degenerate).

3.2. Quantum Theory

The Dirac quantization of the kinematical phase space $\mathcal{P}_{\rm kin} = \mathcal{P}_C \times \mathcal{P}_S$ leads to the kinematical Hilbert space $\mathcal{H}_{\rm kin} \simeq \mathcal{H}_C \otimes \mathcal{H}_S$ describing the clock and system, where $\mathcal{H}_C \simeq L^2(\mathbb{R})$ and $\mathcal{H}_S$ is the Hilbert space associated with S. We assume the system Hamiltonian to be promoted to a self-adjoint operator $\hat{H}_S$ on $\mathcal{H}_S$. An element of $\mathcal{H}_{\rm kin}$ may be expanded in the eigenstates of the clock and system Hamiltonians as

$$|\psi_{\rm kin}\rangle = \sum\!\!\!\!\!\!\int_E \int_{\mathbb{R}} dp_t\; \psi_{\rm kin}(p_t, E)\; |p_t\rangle_C\, |E\rangle_S\,,$$

where the integral-sum highlights that $\hat{H}_S$ may either have a continuous or discrete spectrum[5].

Physical states of the theory satisfy Equation (5), which for the Hamiltonian constraint in Equation (10) becomes

$$\hat{C}_H\ |\psi_{\rm phys}\rangle = \left(s\,\frac{\hat{p}_t^2}{2}\otimes I_S + I_C\otimes \hat{H}_S\right)\,|\psi_{\rm phys}\rangle = 0. \qquad (13)$$

We assume here that this constraint has zero-eigenvalues, i.e., that solutions to Equation (13) exist. Note that this requires the spectrum of $s\,\hat{H}_S$ to contain non-positive eigenvalues, in analogy with the classical case.

Quantizing C_σ in Equation (11) yields $[\hat{C}_+, \hat{C}_-] = 0$, so that the group averaging projector in Equation (6) can be expressed as

$$\delta(\hat{C}_H) = \delta(s\,\hat{C}_+\,\hat{C}_-) = \frac{1}{2(-s\,\hat{H}_S)^{\frac{1}{2}}}\sum_\sigma \delta(\hat{C}_\sigma). \qquad (14)$$

The form of $\delta(\hat{C}_H)$ implies the decomposition of the physical Hilbert space into a direct sum of positive and negative frequency sectors $\mathcal{H}_{\rm phys} \simeq \mathcal{H}_+ \oplus \mathcal{H}_-$ (see also [31, 104]). Acting with the projector $\delta(\hat{C}_H)$ on an arbitrary kinematical state yields a physical state

$$\begin{aligned}|\psi_{\rm phys}\rangle &= \delta(\hat{C}_H)\,|\psi_{\rm kin}\rangle \\ &= \sum_\sigma \int\!\!\!\!\!\!\diagup_{E\in\sigma_{SC}} \frac{\psi_\sigma(E)}{(2|E|)^{1/4}}\ |p_{t,\sigma}(E)\rangle_C\,|E\rangle_S\,, \end{aligned} \qquad (15)$$

where $\psi_\sigma(E)$ are Newton-Wigner-type wave functions associated to the positive and negative frequency modes [90][6]:

$$\psi_\sigma(E) := \frac{\psi_{\rm kin}\left(p_{t,\sigma}(E), E\right)}{(2|E|)^{1/4}}, \qquad (16)$$

and we have defined the function $p_{t,\sigma}(E) := -\sigma\sqrt{2|E|}$ and spectrum

$$\begin{aligned}\sigma_{SC} &:= {\rm Spec}(\hat{H}_S)\cap{\rm Spec}(-\hat{H}_C) \\ &= \left\{E\in{\rm Spec}(\hat{H}_S)\,\middle|\,s\,E\le 0\right\}. \end{aligned} \qquad (17)$$

Physical states are normalized with respect to the physical inner product introduced in Equation (7)

$$\begin{aligned}\langle\psi_{\rm phys}|\phi_{\rm phys}\rangle_{\rm phys} &:= \langle\psi_{\rm kin}|\delta(\hat{C}_H)|\phi_{\rm kin}\rangle_{\rm kin} \\ &= \sum_\sigma \int\!\!\!\!\!\!\diagup_{E\in\sigma_{SC}} \psi^*_\sigma(E)\,\phi_\sigma(E), \end{aligned} \qquad (18)$$

which takes the usual form of non-relativistic quantum mechanics (σ-sector-wise), in line with the properties of Newton-Wigner-type wave functions. This observation will be crucial when discussing relativistic localization in section 6.

4. COVARIANT CLOCKS

4.1. Relational Dynamics With a Classical Covariant Clock

Exploiting the splitting of the degrees of freedom into clock C (our temporal reference system) and evolving system S, we now choose a *clock function* T on $\mathcal{P}_C$ relative to which we describe the evolution of S in terms of relational observables, as discussed in section 2.1.

We could simply choose the phase space coordinate $T = t$ as the clock function. It follows from Equation (2) that in the $s = -1$ case t runs "forward" on the positive frequency sector $\mathcal{C}_+$ and "backward" on the negative frequency sector $\mathcal{C}_-$ along the flow generated by C_H; for $s = +1$ the converse holds. Note that every point in $\mathcal{C}_+\cap\mathcal{C}_-$ corresponds to a static orbit of t (since $p_t = 0$ there), and t is therefore a maximally bad clock function on $\mathcal{C}_+\cap\mathcal{C}_-$. This leads to challenges in describing relational dynamics relative to t: inverse powers of p_t appear in the construction of relational observables encoding the evolution of system degrees of freedom relative to t when canonical pairs on $\mathcal{P}_S$ are used [18, 23, 30, 31][7]. One can solve this problem and obtain a well-defined relational dynamics by using affine (rather than canonical) pairs of evolving phase space coordinates on $\mathcal{P}_S$ in the construction of relational observables [31], or in the quantum theory by carefully regularizing inverse powers of p_t [30].

However, in this article we shall sidestep these challenges and provide an arguably more elegant solution. We choose a different clock function according to the classical covariance condition: that it be canonically conjugate to H_C. This has the consequence of incorporating the pathology at $p_t = 0$ into the clock function (which will nevertheless be meaningfully quantized in section 4.2), and leads to relational observables which work independently of the choice of phase space coordinates on $\mathcal{P}_S$. Solving $\{T, H_C\} = 1$, we find that a covariant clock function T must be of the form $T = s\,t/p_t + g(p_t)$, where $g(p_t)$ is an arbitrary function. Henceforth, we choose $g(p_t) = 0$ for simplicity, so that we have

$$T = s\,\frac{t}{p_t}\,. \qquad (19)$$

[5]The way we have written physical states implicitly assumes the non-positive part of the spectrum of $s\,\hat{H}_S$ to be non-degenerate. Were this not the case, additional degeneracy labels would be necessary. However, this would not otherwise affect the subsequent analysis. See Höhn [31] for an explicit construction of the flat FLRW model with a massless scalar field, whose Hamiltonian constraint can also be interpreted as a free relativistic particle, and thus features a 2-fold system energy degeneracy.

[6]The fourth root comes about because the Newton-Wigner wave function is usually defined for Klein-Gordon systems where what we call E is in fact the square of the energy $\omega_p := \sqrt{\vec{p}^2 + m^2}$. Note also that for the Klein-Gordon case one has a doubly degenerate system energy, which we are not considering here. In that case, it is more convenient to use a momentum, rather than an energy representation of physical states, and a distinct measure. Equation (16) can then indeed be interpreted as the usual Newton-Wigner wave function, written in terms of kinematical states. This will be discussed in more detail in section 6 (see also [31]).

[7]These challenges are related to those facing the definition of time-of-arrival operators in quantum mechanics [18, 23, 30, 117–119].

This clock function is well-defined everywhere, except on the line $p_t = 0$, where H_C is non-degenerate. It is clear that T runs "forward" everywhere on $\mathcal{C}$ for both $s = \pm 1$, except on $\mathcal{C}_+ \cap \mathcal{C}_-$.

The covariance condition, combined with our assumption that the clock does not interact with the system, implies that $\{T, C_H\} = 1$, which simplifies the form of the relational Dirac observables in Equation (4). For example, the relational observable corresponding to the question "what is the value of the system observable f_S when the clock T reads τ?" now takes the simple form [7, 21]

$$F_{f_S,T}(\tau) \approx \sum_{n=0}^{\infty} \frac{(\tau - T)^n}{n!} \{f_S, H_S\}_n. \tag{20}$$

4.2. Covariant Quantum Time Observable for Quadratic Hamiltonians

One might try to construct a time operator in the quantum theory by directly quantizing the covariant clock function in Equation (19) on the clock Hilbert space $\mathcal{H}_C \simeq L^2(\mathbb{R})$ [86, 120, 121]. Choosing a symmetric ordering, this yields

$$\hat{T} = s\,\frac{1}{2}\left(\hat{t}\,\hat{p}_t^{-1} + \hat{p}_t^{-1}\,\hat{t}\right). \tag{21}$$

Here, $\hat{p}_t^{-1}$ is defined in terms of a spectral decomposition such that $\hat{T}\,|p_t = 0\rangle$ is undefined, analogous to the classical case. While the operator $\hat{T}$ is canonically conjugate to the clock Hamiltonian, $[\hat{T}, \hat{H}_C] = i$, it is a symmetric operator that does not admit a self-adjoint extension [86, 88]. Since $\hat{T}$ is not self-adjoint, its status as an observable seems *a priori* unclear[8]. This is a manifestation of Pauli's objection against the construction of time observables in quantum mechanics: For $\hat{H}_C$ bounded below, there does not exist a self-adjoint operator satisfying $[\hat{T}, \hat{H}_C] = i$. Pauli's conclusion was that we are forced to treat time as a classical parameter, different to the way other observables (e.g., position and momentum) are treated [122].

However, it was later realized that by appealing to the more general notion of an observable offered by a POVM, a *covariant* time observable[9] E_T can be constructed whose first moment corresponds to the operator $\hat{T}$ [86, 87, 89]. Such a time observable is defined by a set of effect operator densities $E_T(dt) \geq 0$ normalized as $\int_{\mathbb{R}} E_T(dt) = I_C$, and self-adjoint effect operators $E_T(X) := \int_X E_T(dt)$ associated with the probability $\langle\psi_C|E_T(X)|\psi_C\rangle$ that a measurement of E_T yields an outcome $t \in X \subset \mathbb{R}$ given that the clock was in the state $|\psi_C\rangle \in \mathcal{H}_C$. In order to be a good time observable, it should satisfy the so-called *covariance condition*

$$E_T(X + t) = U_C(t) E_T(X) U_C^\dagger(t), \tag{23}$$

relative to the unitary group action $U_C(t) := e^{-i\hat{H}_C t}$ generated by the clock Hamiltonian. As we shall see shortly, this will give rise to a generalization of canonical conjugacy of the time observable and the clock Hamiltonian, and permit us to extend the approach to relational quantum dynamics based on covariant clock POVMs [7] to relativistic models. In particular, we obtain a valid quantum time observable despite the classical clock pathologies.

In the present case, such an observable can be constructed purely from the self-adjoint quantization of the clock Hamiltonian $\hat{H}_C$ and its eigenstates. The effect densities can be defined as a sum of "projections"

$$E_T(dt) = \frac{1}{2\pi} \sum_\sigma dt\ |t, \sigma\rangle\langle t, \sigma| \tag{24}$$

onto the *clock states* corresponding to the clock reading $t \in \mathbb{R}$ in the negative and positive frequency (i.e., positive and negative clock momentum) sector[10]

$$|t, \sigma\rangle := \int_{\mathbb{R}} dp_t\, \sqrt{|p_t|}\, \theta(-\sigma\, p_t)\, e^{-i\, t\, s\, p_t^2/2}\ |p_t\rangle\ . \tag{25}$$

The covariance condition in Equation (23) is ensured by the fact that the clock states transform as

$$|t + t', \sigma\rangle = U_C(t)\ |t', \sigma\rangle\,. \tag{26}$$

Note that the clock states are orthogonal to the pathological state $|p_t = 0\rangle$, so that the covariant time observable does not have support on it. The clock states are furthermore not mutually orthogonal:

$$\langle t', \sigma'|t, \sigma\rangle = \delta_{\sigma\sigma'}\left[\pi\,\delta\left(t - t'\right) - i\mathrm{P}\frac{1}{t - t'}\right], \tag{27}$$

where P denotes the Cauchy principal value. Hence $E_T(dt)$ is not a true projector. Nevertheless, the following lemma demonstrates that the clock states $|t, \sigma\rangle$ form an over-complete basis for the σ-frequency sector of $\mathcal{H}_C$, and in turn a properly normalized covariant time observable E_T on $\mathcal{H}_{\text{kin}}$.

[8]Using the commutation relation $[\hat{t}, \hat{p}_t^{-1}] = -i\hat{p}_t^{-2}$, which follows from multiplying $[\hat{t}, \hat{p}_t] = i$ from both sides with $\hat{p}_t^{-1}$, we can also write this operator as

$$\hat{T} = s\,\hat{p}_t^{-1}\left(\hat{t} - \frac{i}{2}\,\hat{p}_t^{-1}\right) \tag{22}$$

We note in passing that the operator $\hat{t} - \frac{i}{2}\,\hat{p}_t^{-1}$ is precisely the "complex time operator" derived in Bojowald et al. [32] (see also [33, 34, 77]) when constructing a relational Schrödinger picture for Wheeler-DeWitt type equations for constraints of the form Equation (13).

[9]We emphasize that the covariant time observable is a *kinematical* observable on $\mathcal{H}_{\text{kin}}$ and not a gauge-invariant Dirac observable on $\mathcal{H}_{\text{phys}}$, as by construction its moments will not commute with the constraint. In particular, it is a *partial* observable in the sense of [123].

[10]Compared to Braunstein et al. [89], we use a different definition of the degeneracy label σ (here adapted to positive and negative frequency modes), change the normalization slightly, fix the relative phase, keep the momentum eigenstates as energy eigenstates and introduce s. For notational simplicity, we also set an arbitrary function in Braunstein et al. [89] (accounting for a freedom in choosing the clock states) to zero. This is the quantum analog of the classical choice we made above, where we also set $g(p_t)$ in $T = t/p_t + g(p_t)$ to zero (see also **Supplementary Material** of [7]). It would, however, be straightforward to reinsert this $g(p_t)$ in each of the following expressions.

Lemma 1. *The clock states $|t,\sigma\rangle$ defined in Equation (25) integrate to projectors $\theta(-\sigma\,\hat{p}_t)$ onto the positive/negative frequency sector on $\mathcal{H}_C$*

$$\frac{1}{2\pi}\int_{\mathbb{R}} dt\ |t,\sigma\rangle\langle t,\sigma| = \theta(-\sigma\,\hat{p}_t) \tag{28}$$

and hence form a resolution of the identity as follows:

$$\int_{\mathbb{R}} E(dt) = \frac{1}{2\pi}\sum_{\sigma}\int_{\mathbb{R}} dt\ |t,\sigma\rangle\langle t,\sigma| = I_C\,. \tag{29}$$

Proof: The proof is given in **Supplementary Material**. □

The n^{th}-moment operator of the time observable E_T is defined as

$$\hat{T}^{(n)} := \int_{\mathbb{R}} E_T(dt)\,t^n = \frac{1}{2\pi}\sum_{\sigma}\int_{\mathbb{R}} dt\,t^n\,|t,\sigma\rangle\langle t,\sigma|\,. \tag{30}$$

While the effect operators of the clock POVM are self-adjoint, the moment operators for $n \geq 1$ are not due to the non-orthogonality in Equation (27). Nevertheless, the latter are viable quantum observables with a valid probability interpretation in terms of the POVM; the only price to pay is that the different measurement outcomes t are not perfectly distinguishable.

With this definition, we find that the first-moment operator $\hat{T}^{(1)}$ of E_T is in fact equal to the operator $\hat{T}$ in Equation (21). This was previously noticed in Holevo [86] and Busch et al. [88] (for the $s = +1$ case). This provides a concrete interpretation of the time observable E_T in terms of the classical theory—the time operator $\hat{T}^{(1)}$, namely the first moment of the time observable E_T, is the quantization of the classical clock function T in Equation (19).

Lemma 2. *The operator $\hat{T}$ and the first moment operator $\hat{T}^{(1)}$ of the covariant time observable E_T are equal, $\hat{T} \equiv \hat{T}^{(1)}$.*

Proof: The proof is given in **Supplementary Material**. □

Equation (30) demonstrates that the time operator $\hat{T}$ automatically splits into a positive and negative frequency part, in contrast to $\hat{t}$, the quantization of the phase space coordinate t.

Next, we find that while the clock states are not orthogonal, they are "almost" eigenstates of the covariant time operator $\hat{T}$ on each σ-sector:

Lemma 3. *The clock states $|t,\sigma\rangle$ defined in Equation (25) are* not *eigenstates of $\hat{T} = \hat{T}^{(1)}$. However, for all $|\psi\rangle \in \mathcal{D}(\hat{T})$, where $\mathcal{D}(\hat{T})$ is the domain of $\hat{T}$, they satisfy:*

$$\langle\psi|\ \hat{T}\ |t,\sigma\rangle = t\ \langle\psi|t,\sigma\rangle\ , \quad \forall\, t\in\mathbb{R}\,,\ \sigma = \pm 1\,.$$

Proof: The proof is given in **Supplementary Material**. □

This leads to a another result, which underscores why the covariance condition Equation (23) can be regarded as yielding a generalization of canonical conjugacy:

Lemma 4. *The nth-moment operator defined in Equation (30) satisfies $[\hat{T}^{(n)},\hat{H}_C] = i\,n\,\hat{T}^{(n-1)}$. Furthermore, $\forall\,|\psi\rangle \in \mathcal{D}(\hat{T}^n)$ we have $\hat{T}^{(n)}\,|\psi\rangle = \hat{T}^n\,|\psi\rangle$.*

Proof: The proof is given in **Supplementary Material**. □

We emphasize that the second statement of Lemma 4 does not hold on all of $\mathcal{H}_C$.

The effect density does not commute with the clock Hamiltonian, $[E_T(dt),\hat{H}_C] \neq 0$, which implies the time indicated by the clock (i.e., a measurement outcome of E_T) and the clock energy cannot be determined simultaneously. However, importantly, the following lemma shows that the clock reading and the frequency sector, i.e., the value of σ *can* be simultaneously determined.

Lemma 5. *The effect density $E_T(dt)$ of the covariant clock POVM and the projectors onto the σ-sectors commute: $[E_T(dt),\theta(-\sigma\,\hat{p}_t)] = 0$.*

Proof: The proof is given in **Supplementary Material**. □

Corollary 1. *Since the effect density integrates to the effect and moment operators, this entails that $[E_T(X),\theta(-\sigma\,\hat{p}_t)] = [\hat{T}^{(n)},\theta(-\sigma\,\hat{p}_t)] = 0$, for all $X \subset \mathbb{R}$ and $n \in \mathbb{N}$.*

The significance of this lemma and corollary is that they permit us to condition on the time indicated by the clock and the frequency sector simultaneously. This will become crucial when defining the quantum reduction maps below that take us from the physical Hilbert space to the relational Schrödinger and Heisenberg pictures which exist for each σ-sector. This lemma is thus an important for extending the quantum reduction procedures of Höhn et al. [7] to the class of models considered here.

5. THE TRINITY OF RELATIONAL QUANTUM DYNAMICS: QUADRATIC CLOCK HAMILTONIANS

Having introduced the clock-neutral structure of the classical and quantum theories in section 2, a natural partitioning of the kinematical degrees of freedom into a clock C and system S in section 3, and a covariant time observable E_T in section 4, we are now able to construct a *relational quantum dynamics*, describing how S evolves relative to C.

As noted in the introduction, we showed in Höhn et al. [7] that three formulations of relational quantum dynamics, namely (i) quantum relational Dirac observables, (ii) the relational Schrödinger picture of the Page-Wootters formalism, and (iii) the relational Heisenberg picture obtained through quantum deparametrization, are equivalent for models described by the Hamiltonian constraint in Equation (9) when the clock Hamiltonian has a continuous, non-degenerate spectrum; the three formulations form a *trinity* of relational quantum dynamics. Here we demonstrate that this equivalence extends to

TABLE 2 | Summary of different Hilbert spaces.

Clock C and system S Hilbert spaces
$\mathcal{H}_C$ and $\mathcal{H}_S$
Kinematical Hilbert space
$\mathcal{H}_{\text{kin}} \simeq \mathcal{H}_C \otimes \mathcal{H}_S$
Physical Hilbert space
$\mathcal{H}_{\text{phys}} \simeq \delta(\hat{C}_H)(\mathcal{H}_{\text{kin}}) = \mathcal{H}_+ \oplus \mathcal{H}_-$
Physical system Hilbert space
$\mathcal{H}_S^{\text{phys}} = \Pi_{\sigma_{SC}}(\mathcal{H}_S) \subseteq \mathcal{H}_S$

The various Hilbert spaces appearing in the construction of the trinity. The physical system Hilbert space is the subspace of $\mathcal{H}_S$ spanned by the energy eigenstates permitted upon solving the constraint. The σ-sector $\mathcal{H}_\sigma$ of $\mathcal{H}_{phys}$ is also defined through solutions to the constraint $\hat{C}_\sigma$. Finally, $\Pi_{\sigma_{SC}} = \theta(-s\,\hat{H}_S)$ is a projector onto the system subspace permitted upon solving the constraint in Equation (13).

constraints of the form in Equation (13), involving the doubly degenerate clock Hamiltonian[11].

Thanks to the direct sum structure of the physical Hilbert space $\mathcal{H}_{\text{phys}} = \mathcal{H}_+ \oplus \mathcal{H}_-$ and the separation of the clock moment operators (Equation 30), into non-degenerate positive and negative frequency sectors, all the technical results needed for establishing the equivalence in Höhn et al. [7] will hold per σ-sector for the present class of models. We will thus state some of the following results without proofs, referring the reader as approriate to the proofs of the corresponding results in Höhn et al. [7], which apply here per σ-sector. In particular, Corollary 1 implies that we are permitted to simultaneously condition on the clock reading and the frequency sector.

Lastly, we also provide a discussion of the relational quantum dynamics obtained through reduced phase space quantization. In this case, one deparametrizes the model classically relative to the clock function T, which amounts to a classical symmetry reduction. While the relational quantum dynamics thus obtained yields a relational Heisenberg picture resembling dynamics (iii) of the trinity, it is not always equivalent and thus not necessarily part of the trinity. For this reason, we have moved the exposition of reduced phase space quantization to **Supplementary Material**. It is however useful for understanding why the quantum symmetry reduction explained below is the quantum analog of classical phase space reduction through deparametrization. We emphasize that symmetry reduction and quantization do not commute in general [7, 124–129]. A related discussion for relativistic constraints can also be found in Kaminski et al. [29].

To aid the reader, we summarize the various Hilbert spaces appearing in the construction of the trinity in **Table 2**.

[11]We also refer the reader to the recent work [38], which we became aware of while completing this manuscript. It extends some of the results of Höhn et al. [7] as well, though using the different formalism developed in Chataignier [37]. It also does not employ covariant clocks in the case of quadratic clock Hamiltonians.

5.1. The Three Faces of the Trinity

5.1.1. Dynamics (i): Quantum Relational Dirac Observables

We now quantize the relational Dirac observables in Equation (20), substantiating the discussion of relational quantum dynamics in the clock-neutral picture in section 2.2 for Hamiltonian constraints of the form Equation (13). Quantization of relational Dirac observables has been studied when the quantization of the classical time function T results in a self-adjoint time operator $\hat{T}$ (see [4, 7, 11–20, 27–31, 35–38, 116] and references therein); however, when $\hat{T}$ fails to be self-adjoint, such as in Equation (21), a more general quantization procedure is needed.

Such a procedure was introduced in Höhn et al. [7] based upon the quantization of Equation (20) using covariant time observables. Applying this procedure to the present class of models described by quadratic clock Hamiltonians, we quantize the relational Dirac observables in Equation (20) using the nth-moment operators defined in Equation (30):

$$\begin{aligned}\hat{F}_{f_S,T}(\tau) &:= \int_{\mathbb{R}} E_T(dt) \otimes \sum_{n=0}^{\infty} \frac{i^n}{n!}(t-\tau)^n \left[\hat{f}_S, \hat{H}_S\right]_n \\ &= \sum_{\sigma} \int_{\mathbb{R}} \frac{dt}{2\pi}\, U_{CS}(t)\Big(|\tau,\sigma\rangle\langle\tau,\sigma| \otimes \hat{f}_S\Big) U_{CS}^{\dagger}(t) \\ &=: \sum_{\sigma} \mathcal{G}\left(|\tau,\sigma\rangle\langle\tau,\sigma| \otimes \hat{f}_S\right), \end{aligned} \tag{31}$$

where $[\hat{f}_S, \hat{H}_S]_n := [[\hat{f}_S, \hat{H}_S]_{n-1}, \hat{H}_S]$ is the nth-order nested commutator with the convention $[\hat{f}_S, \hat{H}_S]_0 := \hat{f}_S$, $U_{CS}(t) := \exp(-i\,t\,\hat{C}_H)$, and the second line follows upon a change of integration variable and invoking the covariance condition in Equation (26). The relational Dirac observable $\hat{F}_{f_S,T}(\tau)$ is thus revealed to be an incoherent average over the one-parameter non-compact gauge group G generated by the constraint operator $\hat{C}_H$ of the kinematical operator $|\tau,\sigma\rangle\langle\tau,\sigma| \otimes \hat{f}_S$, which is the system observable of interest $\hat{f}_S$ paired with the projector onto the clock reading τ and the σ-frequency sector. Such a group averaging is known as the G-twirl operation and we denote it $\mathcal{G}$ as in the last line of Equation (31). G-twirl operations have previously been mostly studied in the context of spatial quantum reference frames, e.g., see [130–132], but have also appeared in some constructions of quantum Dirac observables (e.g., see [7, 11, 37, 38, 106])[12]. As discussed in Höhn et al. [7], this G-twirl constitutes the quantum analog of a gauge-invariant extension of a gauge-fixed quantity.

[12]The recent [37, 38] also develop a systematic quantization procedure for relational Dirac observables, based on integral techniques rather than the sum techniques used here and in Höhn et al. [7], and which too yields an expression similar to the one in the second line of Equation (31). While the construction procedure in Chataignier [37, 38] encompasses a more general class of models (but implicitly assumes globally monotonic clocks too), it uses a more restrictive choice of clock observables which, in contrast to the covariant clock POVMs here and in Höhn et al. [7], are required to be self-adjoint. However, the two quantization procedures of relational observables are compatible and it will be fruitful to combine them.

The relational Dirac observables $\hat{F}_{f_S,T}(\tau)$ in Equation (31) constitute a one-parameter family of strong Dirac observables on $\mathcal{H}_{\text{phys}}$ (Theorem 1 of [7] whose proof applies here in each σ-sector):

$$[\hat{F}_{f_S,T}(\tau), \hat{C}_H] = 0, \quad \forall\, \tau \in \mathbb{R}\,. \tag{32}$$

We thus obtain a gauge-invariant relational quantum dynamics by letting the evolution parameter τ in the physical expectation values $\langle\psi_{\text{phys}}|\hat{F}_{f_S,T}(\tau)\,|\psi_{\text{phys}}\rangle_{\text{phys}}$ run.

The decomposition of $\hat{F}_{f_S,T}(\tau)$ in Equation (31) into positive and negative frequency sectors gives rise to a reducible representation of the Dirac observable algebra on the physical Hilbert space. More precisely, relational Dirac observables are superselected across the σ-frequency sectors, and the σ-sum in Equation (31) should thus be understood as a direct sum. To see this, consider the operator $\hat{Q} := \theta(-\hat{p}_t) - \theta(\hat{p}_t)$, where we recall that $\theta(-\sigma\,\hat{p}_t)$ is a projector onto the corresponding σ-sector. By construction $[\hat{Q}, \hat{C}_H] = 0$, which means that $\hat{Q}$ is a strong Dirac observable. Its eigenspaces, with eigenvalues $+1$ and -1, correspond to the positive and negative frequency sector subspaces $\mathcal{H}_+$ and $\mathcal{H}_-$. Furthermore, $\hat{Q}$ commutes with any relational Dirac observable $F_{f_S,T}(\tau)$ in Equation (31) on account of Lemma 5, which implies that Q and any self-adjoint $F_{f_S,T}(\tau)$ can be diagonalized in the same eigenbasis. This in turn implies the following superselection rule

$$\hat{F}_{f_S,T}(\tau) = \hat{F}^{+}_{f_S,T}(\tau) \oplus \hat{F}^{-}_{f_S,T}(\tau), \tag{33}$$

where $\hat{F}^{\sigma}_{f_S,T}(\tau) := \mathcal{G}\big(\,|\tau,\sigma\rangle\langle\tau,\sigma|\otimes\hat{f}_S\big) \in \mathcal{L}(\mathcal{H}_\sigma)$[13].

While there do exist states in the physical Hilbert space that exhibit coherence across the σ-frequency sectors, for example $|\psi_{\text{phys}}\rangle \sim |\psi_+\rangle + |\psi_-\rangle$, where $|\psi_\sigma\rangle \in \mathcal{H}_\sigma$, such coherence is not physically accessible because it does not affect the expectation value of any relational Dirac observable on account of the decomposition in Equation (33). In other words, superpositions and classical mixtures across the σ-frequency sectors are indistinguishable. Hence, superpositions of physical states across σ-sectors are mixed states and the pure physical states are those of either $\mathcal{H}_+$ or $\mathcal{H}_-$ (see also [103, 105] for a discussion on superselection in group averaging).

For example, this superselection rule manifests as a superselection across positive and negative frequency modes in the case of the relativistic particle and across expanding and contracting solutions in the case of the FLRW model with a massless scalar field in **Table 1** [31]. On account of the reducibility of the representation, one usually restricts to one frequency sector (e.g., see [113, 114, 116, 133]). One might conjecture that the analogous superselection rule in a quantum field theory would manifest as a superselection rule between matter and anti-matter sectors.

Superselection rules induced by the G-twirl are often interpreted as arising from the lack of knowledge about a reference frame, and that if an appropriate reference frame is used the superselection rule can be lifted [130]. This interpretation seems unsuitable here. Firstly, lifting the superselection rule would entail undoing the group averaging, in violation of gauge invariance. Secondly, such an interpretation is usually tied to an average over a given group action which parametrizes one's ignorance about relative reference frame orientations. By contrast, the origin of the superselection of Dirac observables here is not the group generated by the constraint, but is a consequence of a property of the constraint, i.e., the group generator. Indeed, the superselection rule above originates in the factorizability of the constraint and the ensuing decomposition of the projector onto the constraint (Equation 14). Both these properties rely on the absence of a $\hat{t}$-dependent term in the constraint Equation (13); if such a term is introduced, one generally finds $[\hat{C}_+, \hat{C}_-] \neq 0$, where the $\hat{C}_\sigma$ are the quantization of the classical factors, but we emphasize that $\hat{C}_H \neq s\,\hat{C}_+\hat{C}_-$ in that case. While such a modified constraint may generate the same group[14], no superselection rule across the σ-sectors would arise. The above superselection rule can thus not be associated with the lack of a shared physical reference frame. This resonates with the interpretation of the physical Hilbert space as a clock-neutral, i.e., temporal-reference-frame-neutral structure (see section 2.2).

Consider now the projector $\Pi_{\sigma_{SC}} = \theta(-s\,\hat{H}_S)$ from $\mathcal{H}_S$ to its subspace spanned by all system energy eigenstates $|E\rangle_S$ with $E \in \sigma_{SC}$; that is, those permitted upon solving the constraint Equation (13). We shall henceforth denote this system Hilbert subspace $\mathcal{H}^{\text{phys}}_S := \Pi_{\sigma_{SC}}(\mathcal{H}_S)$ and call it the *physical system Hilbert space*. We will obtain two copies of the physical system Hilbert space, one for each frequency sector. In analogy to the classical case, we introduce the *quantum weak equality* between operators, signifying that they are equal on the "quantum constraint surface" $\mathcal{H}_{\text{phys}}$:

$$\begin{aligned} &\hat{O}_1 \approx \hat{O}_2 \\ &\quad\Leftrightarrow\ \hat{O}_1\,|\psi_{\text{phys}}\rangle = \hat{O}_2\,|\psi_{\text{phys}}\rangle\,, \quad \forall\,|\psi_{\text{phys}}\rangle \in \mathcal{H}_{\text{phys}}\,. \end{aligned} \tag{34}$$

It follows from Lemma 1 of [7], whose proof applies here per σ-sector, that

$$\hat{F}_{f_S,T}(\tau) \approx \hat{F}_{\Pi_{\sigma_{SC}} f_S\, \Pi_{\sigma_{SC}},T}(\tau)\,, \tag{35}$$

[13] In particular, when the spectrum of $\hat{H}_S$ does not contain zero, the G-twirl $\mathcal{G}$ can on each σ-sector be weakly rewritten as a reduced G-twirl $\mathcal{G}_\sigma$, i.e., one generated by $\hat{C}_\sigma$, rather than $\hat{C}_H$. Indeed, it is easy to see that the observables in Equation (31) satisfy

$$\hat{F}^{\sigma}_{f_S,T}(\tau) = \mathcal{G}\big(\,|\tau,\sigma\rangle\langle\tau,\sigma|\otimes\hat{f}_S\big) \approx \delta(\hat{C}_H)(|\tau,\sigma\rangle\,\langle\tau,\sigma|\otimes\hat{f}_S),$$

where $\approx$ is the quantum weak equality introduced in Equation (34). Now use Equation (14) and notice that $\delta(\hat{C}_{-\sigma})\,|\tau,\sigma\rangle\otimes|E\rangle_S = 0$ when zero does not lie in the spectrum of $\hat{H}_S$. This observation yields

$$\begin{aligned} \hat{F}^{\sigma}_{f_S,T}(\tau) \;&\approx\; \frac{1}{2(-s\hat{H}_S)^{\frac{1}{2}}}\delta(\hat{C}_\sigma)(|\tau,\sigma\rangle\,\langle\tau,\sigma|\otimes\hat{f}_S), \\ &\approx\; \frac{1}{2(-s\hat{H}_S)^{\frac{1}{2}}}\,\mathcal{G}_\sigma(|\tau,\sigma\rangle\,\langle\tau,\sigma|\otimes\hat{f}_S)\,, \end{aligned}$$

where the last weak equality is restricted to $\mathcal{H}_\sigma$. When zero does lie in the spectrum of $\hat{H}_S$, the decomposition Equation (14) is not well-defined and one needs to regularize.

[14] For example, when $C_H = p_t^2 - H^2(q_i, p_i, t)$ and $H^2 > 0\ \forall\, t \in \mathbb{R}$.

are weakly equal relational Dirac observables. Hence, the relational Dirac observables in Equation (31) form weak equivalence classes on $\mathcal{H}_{\text{phys}}$, where $\hat{F}_{f_S,T}(\tau) \sim \hat{F}_{g_S,T}(\tau)$ if $\Pi_{\sigma_{SC}}\hat{f}_S\,\Pi_{\sigma_{SC}} = \Pi_{\sigma_{SC}}\hat{g}_S\,\Pi_{\sigma_{SC}}$. These weak equivalence classes are labeled by what we shall denote

$$\hat{f}_S^{\text{phys}} := \Pi_{\sigma_{SC}}\hat{f}_S\,\Pi_{\sigma_{SC}} \in \mathcal{L}\left(\mathcal{H}_S^{\text{phys}}\right), \tag{36}$$

for arbitrary $\hat{f}_S \in \mathcal{L}(\mathcal{H}_S)$, where $\mathcal{L}$ denotes the set of linear operators. For later use, we note that the algebras of the physical system observables $\hat{f}_S^{\text{phys}}$ on $\mathcal{H}_S^{\text{phys}}$ and the $\hat{F}_{f_S^{\text{phys}},T}(\tau)$ on $\mathcal{H}_{\text{phys}}$ are weakly homomorphic with respect to addition, multiplication and commutator relations. More precisely,

$$\hat{F}_{f_S^{\text{phys}}+g_S^{\text{phys}}\cdot h_S^{\text{phys}},T}(\tau) \approx \hat{F}_{f_S^{\text{phys}},T}(\tau) + \hat{F}_{g_S^{\text{phys}},T}(\tau)\cdot\hat{F}_{h_S^{\text{phys}},T}(\tau)$$

$\forall f_S, g_S, h_S \in \mathcal{L}(\mathcal{H}_S)$. This is a consequence of Theorem 2 of Höhn et al. [7] (whose proof again applies here per σ-sector). Together with Höhn et al. [7], this translates the weak classical algebra homomorphism defined through relational observables in Dittrich [21] into the quantum theory.

5.1.2. Dynamics (ii): The Page-Wootters Formalism

Suppose we are given a quantum Hamiltonian constraint Equation (5) which splits into a clock and system contribution as in Equation (9), but for the moment not necessarily assuming it to be of the quadratic form in Equation (13). Suppose further we are given some (kinematical) time observable on the clock Hilbert space, which need not necessarily be a clock POVM which is covariant with respect to the group generated by the clock Hamiltonian, but is taken to define the clock reading. Page and Wootters [39, 40, 134, 135] proposed to extract a relational quantum dynamics between the clock and system from physical states in terms of conditional probabilities: what is the probability of an observable $\hat{f}_S$ associated with the system S giving a particular outcome f, if the measurement of the clock's time observable yields the time τ? If $e_C(\tau)$ and $e_{f_S}(f)$ are the projectors onto the clock reading τ and the system observable $\hat{f}_S$ taking the value f, this conditional probability is postulated in the form

$$\text{Prob}\left(f \text{ when } \tau\right) = \frac{\text{Prob}\left(f \text{ and } \tau\right)}{\text{Prob}\left(\tau\right)} = \frac{\langle\psi_{\text{phys}}|\, e_C(\tau)\otimes e_{f_S}(f)\,|\psi_{\text{phys}}\rangle_{\text{kin}}}{\langle\psi_{\text{phys}}|\, e_C(\tau)\otimes I_S\,|\psi_{\text{phys}}\rangle_{\text{kin}}}. \tag{37}$$

This expression appears at first glimpse to be in violation of the constraints, as it acts with operators on physical states that are not Dirac observables; this is the basis of Kuchař's criticism (b) that the conditional probabilities of the Page-Wootters formalism are incompatible with the constraints [1]. However, for a class of models we have shown in Höhn et al. [7] that the expression Equation (37) is a quantum analog of a gauge-fixed expression of a manifestly gauge-invariant quantity and thus consistent with the constraint. In this section we extend this result to relativistic settings.

Here we shall expand the Page-Wootters formalism to the more general class of Hamiltonian constraints of the form Equation (13) exploiting the covariant clock POVM E_T of section 4.2[15]. On the one hand, this will permit us to prove full equivalence of the so-obtained relational quantum dynamics with the manifestly gauge-invariant formulation in terms of relational Dirac observables on $\mathcal{H}_{\text{phys}}$ of Dynamics (i). As an aside, this will also resolve the normalization issue of physical states appearing in Diaz et al. [50], where the kinematical rather than physical inner product was used to normalize physical states, thus yielding a divergence (when used for equal mass states). On the other hand, the covariant clock POVM will allow us, in section 6, to address the observation by Kuchař [1] that using the Minkowski time observable leads to incorrect localization probabilities for relativistic particles in the Page-Wootters formalism.

The Page-Wootters formalism produces the system state at clock time τ by conditioning physical states on the clock reading τ [39, 40, 134, 135]. Henceforth focusing on the class of models defined by the constraint in Equation (13) and the covariant clock POVM of section 4.2, and given the reducible representation of $\mathcal{H}_{\text{phys}}$, we may additionally condition on the frequency sector thanks to Lemma 5. In extension of Höhn et al. [7], we may use this conditioning to define two reduction maps $\mathcal{R}^{\sigma}_{\mathbf{S}}(\tau): \mathcal{H}_{\text{phys}} \to \mathcal{H}^{\text{phys}}_{S,\sigma}$, one per σ-frequency sector,

$$\mathcal{R}^{\sigma}_{\mathbf{S}}(\tau) := \langle\tau,\sigma|\otimes I_S, \tag{38}$$

where $\mathcal{H}^{\text{phys}}_{S,\sigma}$ is a copy of $\mathcal{H}^{\text{phys}}_S = \Pi_{\sigma_{SC}}(\mathcal{H}_S)$, i.e., the subspace of the system Hilbert space permitted upon solving the constraint, corresponding to the σ-frequency sector. As will become clear shortly, the label **S** on the reduction map stands for "Schrödinger picture" and to distinguish it from the italic S which stands for the system, we henceforth write it in bold face. Due to the decomposition $\mathcal{H}_{\text{phys}} = \mathcal{H}_+ \oplus \mathcal{H}_-$, we equip the two copies $\mathcal{H}^{\text{phys}}_{S,\sigma}$ with the frequency label σ in order to remind ourselves which reduced theory corresponds to which positive or negative frequency mode.

The reduced states (whose normalization factor $1/\sqrt{2}$ will be explained later),

$$\frac{1}{\sqrt{2}}\,|\psi^{\sigma}_S(\tau)\rangle := \mathcal{R}^{\sigma}_{\mathbf{S}}(\tau)\,|\psi_{\text{phys}}\rangle = \int\!\!\!\!\!\!\sum_{E\in\sigma_{SC}} \psi_\sigma(E)\,e^{-i\tau E}\,|E\rangle_S\,, \tag{39}$$

where $\psi_\sigma(E)$ is the Newton-Wigner type wave function defined in Equation (16), satisfy the Schrödinger equation with respect to $\hat{H}_S$:

$$i\frac{d}{d\tau}\,|\psi^{\sigma}_S(\tau)\rangle = \hat{H}_S\,|\psi^{\sigma}_S(\tau)\rangle\,. \tag{40}$$

[15] See also the recent [38] for a complementary approach. Note that it does not employ covariant POVMs for quadratic Hamiltonians, and is thus subject to Kuchař's criticism (a) described in section 1.

We interpret this as the dynamics of S relative to the temporal reference frame C. In particular, this Schrödinger equation looks the same for both the positive and negative frequency sectors because the time defined by the covariant clock POVM E_T runs forward in both sectors. This is clear from Equation (26) and is the quantum analog of the earlier classical observation that the clock function T runs "forward" on both frequency sectors $\mathcal{C}_+$ and $\mathcal{C}_-$ (in contrast to t)[16].

Thanks to Equation (29), the decomposition of the physical states into positive and negative frequency modes, Equation (15) can also be written as follows:

$$|\psi_{\text{phys}}\rangle = \frac{1}{2\sqrt{2}\,\pi} \sum_{\sigma} \int_{\mathbb{R}} dt\ |t,\sigma\rangle\, |\psi_S^{\sigma}(t)\rangle\ . \tag{41}$$

Together with Lemma 1, this implies that the σ-sector left inverse $\mathcal{H}_{S,\sigma}^{\text{phys}} \to \mathcal{H}_{\sigma}$ of the reduction map defined in Equation (38) is given by

$$\begin{aligned} (\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau))^{-1} &= \frac{1}{2\pi} \int_{\mathbb{R}} dt\ |t,\sigma\rangle \otimes U_S(t-\tau) \\ &= \delta(\hat{C}_H)(|\tau,\sigma\rangle \otimes I_S)\,, \end{aligned} \tag{42}$$

where $U_S(t) = \exp(-i\,\hat{H}_S\,t)$, so that

$$\begin{aligned} (\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau))^{-1}\,\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau) &= \delta(\hat{C}_H)(|\tau,\sigma\rangle\,\langle\tau,\sigma| \otimes I_S) \\ &\approx \theta(-\sigma\,\hat{p}_t) \otimes I_S\,, \end{aligned} \tag{43}$$

where $\approx$ is the quantum weak equality, and thus

$$\sum_{\sigma} (\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau))^{-1}\,\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau) \approx I_{\text{phys}}\,. \tag{44}$$

Conversely, we can write the identity on $\mathcal{H}_{S,\sigma}^{\text{phys}}$ in the form

$$\begin{aligned} \mathcal{R}_{\mathbf{S}}^{\sigma}(\tau)\,(\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau))^{-1} &= \langle\tau,\sigma|\,\delta(\hat{C}_H)\,|\tau,\sigma\rangle \\ &= \Pi_{\sigma_{SC}}\,. \end{aligned}$$

The identity in the second line can be checked with the help of Equation (27) (and also follows from the proof of Theorem 2 in [7]). A summary of these maps can be found in **Figure 2**.

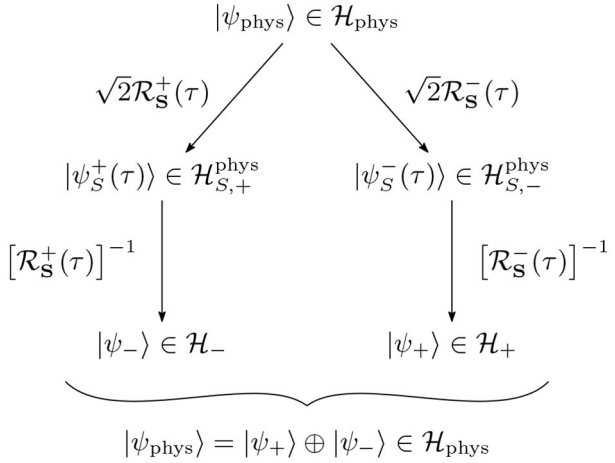

FIGURE 2 | A summary of the Page-Wootters reduction maps and their inverses. The analogous state of affairs holds for the quantum symmetry reduction maps and their inverses.

Using these reduction maps and their inverses, we can define an encoding operation $\mathcal{E}_{\mathbf{S}}^{\tau,\sigma} : \mathcal{L}\left(\mathcal{H}_{S,\sigma}^{\text{phys}}\right) \to \mathcal{L}\left(\mathcal{H}_{\sigma}\right)$, mapping the observables in Equation (36), acting on the physical system Hilbert space $\mathcal{H}_{S,\sigma}^{\text{phys}}$, into Dirac observables on the σ-sector of $\mathcal{H}_{\text{phys}}$:

$$\begin{aligned} \mathcal{E}_{\mathbf{S}}^{\tau,\sigma}\left(\hat{f}_S^{\text{phys}}\right) &:= (\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau))^{-1}\,\hat{f}_S^{\text{phys}}\,\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau) \\ &= \delta(\hat{C}_H)\left(|\tau,\sigma\rangle\langle\tau,\sigma| \otimes \hat{f}_S^{\text{phys}}\right)\,. \end{aligned} \tag{45}$$

These encoded observables turn out to be the σ-sector part of the relational Dirac observables in Equation (31), as articulated in the following theorem.

Theorem 1. *Let $\hat{f}_S \in \mathcal{L}(\mathcal{H}_S)$. The quantum relational Dirac observable $\hat{F}_{f_S,T}(\tau)$ acting on $\mathcal{H}_{\text{phys}}$, Equation (31), reduces under $\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau)$ to the corresponding projected observable on $\mathcal{H}_{S,\sigma}^{\text{phys}}$,*

$$\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau)\ \hat{F}_{f_S,T}(\tau)\,(\mathcal{R}_{\mathbf{S}}^{\sigma}(\tau))^{-1} = \Pi_{\sigma_{SC}}\hat{f}_S\,\Pi_{\sigma_{SC}} \equiv \hat{f}_S^{\text{phys}}.$$

Conversely, let $\hat{f}_S^{\text{phys}} \in \mathcal{L}\left(\mathcal{H}_{S,\sigma}^{\text{phys}}\right)$. The encoding operation in Equation (45) of system observables coincides on the physical Hilbert space $\mathcal{H}_{\text{phys}}$ *with the quantum relational Dirac observables in Equation (31) projected into the σ-sector:*

$$\mathcal{E}_{\mathbf{S}}^{\tau,\sigma}\left(\hat{f}_S^{\text{phys}}\right) \approx \hat{F}^{\sigma}_{\hat{f}_S^{\text{phys}},T}(\tau)\,, \tag{46}$$

where $\hat{F}^{\sigma}_{\hat{f}_S^{\text{phys}},T}(\tau) = \hat{F}_{\hat{f}_S^{\text{phys}},T}(\tau)\,(\theta(-\sigma\,\hat{p}_t)\otimes I_S)$—c.f. Equation (33).

Proof: The proof of Theorem 3 in Höhn et al. [7] applies here per σ-sector. □

[16]Note that we could also define linear combinations of clock states $|\tau\rangle := \sum_{\sigma} c_{\sigma}\ |\tau,\sigma\rangle$. Clearly, then we would also find that

$$\begin{aligned} |\psi_S(\tau)\rangle &:= \langle\tau|\psi_{\text{phys}}\rangle \\ &= \left(\sum_{\sigma} c_{\sigma}^{*} \sum_{E\in\sigma_{SC}} \psi_{\sigma}(E)\,e^{-i\,\tau E}\right)|E\rangle_S \end{aligned}$$

satisfies the same Schrödinger equation (40). However, it is straightforward to check, using Lemma 1, that these new clock states do not give rise to a resolution of the identity, $\int d\tau\ |\tau\rangle\,\langle\tau| \neq I_C$ and so $|\psi_{\text{phys}}\rangle \neq \int d\tau\ |\tau\rangle\,|\psi_S(\tau)\rangle$. In fact, if $c_{\sigma} \neq 0$ for $\sigma = +,-$, then $|\psi_S(\tau)\rangle$ will mix contributions from the positive and negative frequency sectors such that it will become impossible to reconstruct (either of the positive or negative frequency part of) the physical state from it. That is, a reduction map $\langle\tau| \otimes I_S$, which only conditions on the clock time, would not be invertible. This is another consequence of the superselection rule discussed above which entails that superpositions and mixtures across σ-sectors are indistinguishable through Dirac observables. It is also another reason why we condition also on the frequency sector when defining the reduction map in Equation (38).

Note that the relational Dirac observables $\hat{F}_{f_S,T}(\tau)$ commute with the projectors $\theta(-\sigma\,\hat{p}_t)$ due to the reducible representation in Equation (31).

Apart from providing the σ-sector-wise dictionary between the observables on the physical Hilbert space and the physical system Hilbert space, Theorem 1, in conjunction with the weak equivalence in Equation (35), also implies an equivalence between the full sets of relational Dirac observables $\hat{F}^{\sigma}_{f_S^{\rm phys},T}(\tau)$ on $\mathcal{H}_\sigma$ and system observables on $\mathcal{H}^{\rm phys}_{S,\sigma}$.

Crucially, the expectation values of the relational Dirac observables Equation (31) in the physical inner product Equation (18) coincide, σ-sector-wise, with the expectation values of the physically permitted system observables $\hat{f}_S^{\rm phys}$ in the states solving the Schrödinger equation Equation (40) on $\mathcal{H}^{\rm phys}_{S,\sigma}$.

Theorem 2. *Let $\hat{f}_S \in \mathcal{L}(\mathcal{H}_S)$, and denote its associated operator on $\mathcal{H}_S^{\rm phys}$ by $\hat{f}_S^{\rm phys} = \Pi_{\sigma_{SC}}\hat{f}_S\,\Pi_{\sigma_{SC}}$. Then*

$$\langle\psi_{\rm phys}|\hat{F}^{\sigma}_{f_S,T}(\tau)\,|\phi_{\rm phys}\rangle_{\rm phys} = \frac{1}{2}\,\langle\psi^{\sigma}_S(\tau)|\hat{f}_S^{\rm phys}\,|\phi^{\sigma}_S(\tau)\rangle\ ,$$

where $|\psi^{\sigma}_S(\tau)\rangle = \sqrt{2}\,\mathcal{R}^{\sigma}_{\mathbf{S}}(\tau)\ |\psi_{\rm phys}\rangle$ as in Equation (39). Hence,

$$\langle\psi_{\rm phys}|\hat{F}_{f_S,T}(\tau)\,|\phi_{\rm phys}\rangle_{\rm phys} = \frac{1}{2}\sum_{\sigma}\langle\psi^{\sigma}_S(\tau)|\hat{f}_S^{\rm phys}\,|\phi^{\sigma}_S(\tau)\rangle\,.$$

Proof: The proof of Theorem 4 in Höhn et al. [7] applies here per σ-sector. □

Hence, the expectation values in the relational Schrödinger picture (i.e., the Page-Wootters formalism) are equivalent to the gauge-invariant ones of the corresponding relational Dirac observables on $\mathcal{H}_{\rm phys}$. Accordingly, equations, such as Equation (37) (adapted to σ-sectors) are not in violation of the constraint as claimed by Kuchař [1].

This immediately implies that the reduction maps $\mathcal{R}^{\sigma}_{\mathbf{S}}(\tau)$ preserve inner products per σ-sector as follows.

Corollary 2. *Setting $\hat{f}_S = \Pi_{\sigma_{SC}}$ in Theorem 2 yields*

$$\langle\psi_{\rm phys}|\,\theta(-\sigma\,\hat{p}_t)\otimes I_S\,|\phi_{\rm phys}\rangle_{\rm phys} = \frac{1}{2}\,\langle\psi^{\sigma}_S(\tau)|\phi^{\sigma}_S(\tau)\rangle\ ,$$

where $|\psi^{\sigma}_S(\tau)\rangle = \sqrt{2}\,\mathcal{R}^{\sigma}_{\mathbf{S}}(\tau)\ |\psi_{\rm phys}\rangle$. Hence,

$$\begin{aligned}\langle\psi_{\rm phys}|\phi_{\rm phys}\rangle_{\rm phys} &= \frac{1}{2}\sum_{\sigma}\langle\psi^{\sigma}_S(\tau)|\,\phi^{\sigma}_S(\tau)\rangle \qquad (47)\\ &= \sum_{\sigma}\langle\psi_{\rm phys}|(|\tau,\sigma\rangle\,\langle\tau,\sigma|\otimes I_S)|\phi_{\rm phys}\rangle_{\rm kin}\ .\end{aligned}$$

The reason for introducing the normalization factor $1/\sqrt{2}$ in Equation (39) is now clear: it permits us to work with normalized states $\langle\psi^{\sigma}_S(\tau)|\psi^{\sigma}_S(\tau)\rangle = 1 = \langle\psi_{\rm phys}|\psi_{\rm phys}\rangle_{\rm phys}$ in each reduced σ-sector and in the physical Hilbert space simultaneously.

The above results show:

(1) Applying the Page-Wootters reduction map $\mathcal{R}^{\sigma}_{\mathbf{S}}(\tau)$ to the physical Hilbert space $\mathcal{H}_{\rm phys}$ yields a *relational Schrödinger picture* with respect to the clock C on the physical system Hilbert space $\mathcal{H}^{\rm phys}_{S,\sigma}$ corresponding to the σ-frequency sector.
(2) σ-sector wise, the relational quantum dynamics encoded in the relational Dirac observables on the physical Hilbert space is equivalent to the dynamics in the relational Schrödinger picture on the physical system Hilbert space of the Page-Wootters formalism.
(3) Given the invertibility of the reduction map, Theorem 2 formally shows that if $\hat{f}_S^{\rm phys}$ is self-adjoint on $\mathcal{H}^{\rm phys}_{S,\sigma}$, then so is $\hat{F}^{\sigma}_{f_S,T}(\tau)$ on $\mathcal{H}_\sigma$.

We note that the expression in the second line of Equation (47) also defines an inner product on the space of solutions to the Wheeler-DeWitt-type constraint Equation (13), which is equivalent to the physical inner product in Equation (18) obtained through group averaging. These two inner products thus define the same physical Hilbert space $\mathcal{H}_{\rm phys}$. The expression in the second line of Equation (47) is the adaptation of the Page-Wootters inner product introduced in Smith and Ahmadi [44] to the reducible representation of the physical Hilbert space associated to Hamiltonian constraints with quadratic clock Hamiltonians.

5.1.3. Dynamics (iii): Quantum Deparametrization

Classically, one can perform a symmetry reduction of the clock-neutral constraint surface by gauge-fixing the flow of the constraint. In this case, this yields two copies of a gauge-fixed reduced phase space, one for each frequency sector, each equipped with a standard Hamiltonian dynamics, hence yielding a deparametrized theory (see **Supplementary Material**). In contrast to the classical constraint surface, the "quantum constraint surface" $\mathcal{H}_{\rm phys}$ is automatically gauge-invariant since the exponentiation of the symmetry generator $\hat{C}_H$ acts trivially on all physical states and Dirac observables. Hence, there is no more gauge-fixing in the quantum theory after solving the constraint. Nevertheless, following Höhn and Vanrietvelde [30], Höhn [31], and Höhn et al. [7], we now demonstrate the quantum analog of the classical symmetry reduction procedure for the class of models considered in this article. As such it is the quantum analog of deparametrization, which we henceforth refer to as *quantum deparametrization*. This quantum symmetry reduction maps the clock-neutral Dirac quantization to a relational Heisenberg picture relative to clock observable E_T, and involves the following two steps.

1. *Constraint trivialization*: A transformation of the constraint such that it only acts on the chosen reference system (here clock C), fixing its degrees of freedom. The classical analog is a canonical transformation which turns the constraint into a momentum variable and separates gauge-variant from gauge-invariant degrees of freedom [7, 71].

2. *Condition on classical gauge fixing conditions*: A "projection" which removes the now redundant reference frame degrees of freedom[17].

We begin by defining the *constraint trivialization map* $\mathcal{T}_{T,\epsilon}:\mathcal{H}_{\text{phys}} \rightarrow \mathcal{T}_{T,\epsilon}(\mathcal{H}_{\text{phys}})$ relative to the covariant time observable E_T. This map will transform the physical Hilbert space into a new Hilbert space, while preserving inner products and algebraic properties of observables

$$\begin{aligned}\mathcal{T}_{T,\epsilon} &:= \sum_{n=0}^{\infty} \frac{i^n}{n!}\, \hat{T}^{(n)} \otimes \left(\hat{H}_S + s\,\frac{\epsilon^2}{2}\right)^n \\ &= \frac{1}{2\pi}\sum_{\sigma}\int_{\mathbb{R}} dt\ |t,\sigma\rangle\langle t,\sigma| \otimes e^{i\,t\,(\hat{H}_S + s\,\epsilon^2/2)}\,. \end{aligned} \tag{48}$$

In analogy to Höhn and Vanrietvelde [30], Höhn [31], and Höhn et al. [7], we introduce an arbitrary positive parameter $\epsilon > 0$ so that the map becomes invertible. Note that $s\,\epsilon^2/2 \in \text{Spec}(\hat{H}_C)$.

Lemma 6. *The left inverse of the trivialization $\mathcal{T}_{T,\epsilon}$ is given by*

$$\mathcal{T}^{-1}_{T,\epsilon} = \frac{1}{2\pi}\sum_{\sigma}\int_{\mathbb{R}} dt\ |t,\sigma\rangle\langle t,\sigma| \otimes e^{-i\,t\,(\hat{H}_S + s\,\epsilon^2/2)}\,,$$

so that, for any $\epsilon > 0$, $\mathcal{T}^{-1}_{T,\epsilon}:\mathcal{T}_{T,\epsilon}(\mathcal{H}_{\text{phys}}) \rightarrow \mathcal{H}_{\text{phys}}$ and

$$\mathcal{T}^{-1}_{T,\epsilon}\circ\mathcal{T}_{T,\epsilon} \approx I_{\text{phys}}.$$

Proof: The proof of Lemma 2 in Höhn et al. [7] applies here σ-sector wise. □

The main property of the trivialization map is summarized in the following lemma.

Lemma 7. *The map $\mathcal{T}_{T,\epsilon}$ trivializes the constraint in Equation (13) to the clock degrees of freedom*

$$\mathcal{T}_{T,\epsilon}\,\hat{C}_H\,\mathcal{T}^{-1}_{T,\epsilon} \overset{*}{\approx} \frac{s}{2}(\hat{p}_t^2 - \epsilon^2)\otimes I_S, \tag{49}$$

where $\overset{*}{\approx}$ is the quantum weak equality on the trivialized physical Hilbert space $\mathcal{T}_{T,\epsilon}(\mathcal{H}_{\text{phys}})$, and transforms physical states into a sum of two product states with a fixed and redundant clock factor

$$\begin{aligned}\mathcal{T}_{T,\epsilon}\ |\psi_{\text{phys}}\rangle = \frac{1}{\sqrt{\epsilon}}\sum_{\sigma}\ |p_t = -\sigma\epsilon\rangle_C \\ \otimes\ \sum\!\!\!\!\!\!\int_{E\in\sigma_{SC}} \psi_\sigma(E)\ |E\rangle_S\ .\end{aligned} \tag{50}$$

Proof: The proof of Lemma 2 in Höhn et al. [7] applies here σ-sector wise. □

[17]We put projection in quotation marks because it is not a true projection when applied to the physical Hilbert space as it only removes redundancy in the description, i.e., degrees of freedom which are fixed through the constraint. No physical information is lost, which is why this operation will be invertible. It would however be a projection on $\mathcal{H}_{\text{kin}}$.

Hence, per σ-frequency sector, the trivialized physical states are product states with respect to the tensor product decomposition of the kinematical Hilbert space. Recalling the discussion of the superselection rule across σ-sectors, the physical state in Equation (50) is indistinguishable from a separable mixed state when it contains both positive and negative frequency modes. One can therefore also view the trivialization as a σ-sector-wise disentangling operation, given that physical states in Equation (15) appear to be entangled. However, we emphasize that this notion of entanglement is kinematical and not gauge-invariant (see [7] for a detailed discussion of this and how the trivialization can also be used to clarify the role of entanglement in the Page-Wootters formalism).

The clock factors in Equation (50) have become redundant, apart from distinguishing between the positive and negative frequency sector. Indeed, if we had $\epsilon = 0$, then (disregarding the diverging prefactor) both the negative and positive frequency terms in Equation (50) would have a common redundant factor $|p_t = 0\rangle_C$, so that one could no longer distinguish between them at the level of the eigenbases of $\hat{p}_t$ and $\hat{H}_S$ in which the states have been expanded. That illustrates why $\mathcal{T}_{T,\epsilon=0}$ is *not* invertible when acting on physical states. Indeed, $\mathcal{T}^{-1}_{T,\epsilon=0}$ is undefined on states of the form $|p_t = 0\rangle_C\,|\psi\rangle_S$, since $\hat{T}^{(n)}$ is not defined on $|p_t = 0\rangle_C$. This is similar to the construction of the trivialization maps in Höhn and Vanrietvelde [30] and Höhn [31], except that here the decomposition into positive and negative frequency sectors proceeds somewhat differently. This concludes step 1. above.

In order to complete the reduction to the states of the relational Heisenberg picture, and thus also complete step 2. above, we "project" out the redundant clock factor of the trivialized states by projecting onto the classical gauge-fixing condition $T = \tau$ (see **Supplementary Material** for a discussion of the classical gauge-fixing). That is, we now proceed as in the Page-Wootters reduction and condition states in the trivialized physical Hilbert space $\mathcal{T}_{T,\epsilon}(\mathcal{H}_{\text{phys}})$ on the clock reading τ, separating positive and negative frequency modes. Altogether, the quantum symmetry reduction map takes the form

$$\mathcal{R}^{\sigma}_{\mathbf{H}} := e^{-i\,\tau\,s\,\epsilon^2/2}\,(\langle\tau,\sigma|\otimes I_S)\,\mathcal{T}_{T,\epsilon}. \tag{51}$$

Using that

$$\langle\tau,\sigma|p_t = -\sigma'\epsilon\rangle = \delta_{\sigma\sigma'}\sqrt{\epsilon}\,e^{i\,\tau\,s\,\epsilon^2/2}\,, \tag{52}$$

which is another reason why $\epsilon > 0$ is chosen, in Equation (50) one finds τ-independent system states

$$\begin{aligned}\mathcal{R}^{\sigma}_{\mathbf{H}}\ |\psi_{\text{phys}}\rangle\ &= \sum\!\!\!\!\!\!\int_{E\in\sigma_{SC}} \psi_\sigma(E)\ |E\rangle_S \\ &=: \frac{1}{\sqrt{2}}\ |\psi^{\sigma}_S\rangle\ ,\end{aligned} \tag{53}$$

as appropriate for a Heisenberg picture (compare with Equation 39). This is also the reason why, in contrast to the Page-Wootters reduction in Equation (38), we do not label the quantum deparametrization map with a τ, despite the latter

a priori appearing on the right hand side of Equation (51) and why we equip it with an **H** label. The factor $1/\sqrt{2}$ has again been introduced for normalization purposes. Since the wave function

$$\psi_S^\sigma(E) \equiv \sqrt{2}\,\psi_\sigma(E)\,, \tag{54}$$

is square-integrable/summable, it is clear that $|\psi_S^\sigma\rangle$ is an element of the physical system Hilbert space $\mathcal{H}_{S,\sigma}^{\text{phys}}$, corresponding to the σ-sector. We therefore also have $\mathcal{R}_{\mathbf{H}}^\sigma(\tau): \mathcal{H}_{\text{phys}} \to \mathcal{H}_{S,\sigma}^{\text{phys}}$, just as in Page-Wootters reduction. Using Lemmas 6 and 7, it is now also clear how to invert the quantum symmetry reduction—at least per σ-sector:

$$(\mathcal{R}_{\mathbf{H}}^\sigma)^{-1} \coloneqq \mathcal{T}_{T,\epsilon}^{-1}\left(\frac{1}{\sqrt{\epsilon}}\,|p_t = -\sigma\epsilon\rangle_C \otimes I_S\right)$$

defines a map $\mathcal{H}_{S,\sigma}^{\text{phys}} \to \mathcal{H}_\sigma$, so that[18]

$$(\mathcal{R}_{\mathbf{H}}^\sigma)^{-1}\,|\psi_S^\sigma\rangle = \sqrt{2}\,\theta(-\sigma\,\hat{p}_t)\,|\psi_{\text{phys}}\rangle\,. \tag{55}$$

Hence, from the physical system Hilbert space of the positive/negative frequency modes one can only recover the positive/negative frequency sector of the physical Hilbert space. Note that the inverse map is independent of τ in contrast to the Page-Wottters case.

More precisely, the following holds.

Lemma 8. *The quantum symmetry reduction map is weakly equal to the Page-Wootters reduction map and an (inverse) system time evolution*

$$\begin{aligned}\mathcal{R}_{\boldsymbol{H}}^\sigma &\approx \langle\tau,\sigma|\otimes U_S^\dagger(\tau)\,,\\ &= U_S^\dagger(\tau)\,\mathcal{R}_{\boldsymbol{S}}^\sigma(\tau).\end{aligned}$$

Similarly, the inverse of the quantum symmetry reduction is equal to a system time evolution and the inverse of the Page-Wootters reduction:

$$\begin{aligned}(\mathcal{R}_{\boldsymbol{H}}^\sigma)^{-1} &= \delta(\hat{C}_H)\left(|\tau,\sigma\rangle\otimes U_S(\tau)\right)\\ &= (\mathcal{R}_{\boldsymbol{S}}^\sigma(\tau))^{-1}\,U_S(\tau).\end{aligned}$$

Hence

$$(\mathcal{R}_{\boldsymbol{H}}^\sigma)^{-1}\,\mathcal{R}_{\boldsymbol{H}}^\sigma \approx \theta(-\sigma\,\hat{p}_t)\otimes I_S$$

and

$$\mathcal{R}_{\boldsymbol{H}}^\sigma\,(\mathcal{R}_{\boldsymbol{H}}^\sigma)^{-1}\,|\psi_S^\sigma\rangle = |\psi_S^\sigma\rangle\,.$$

[18]This is understood as appending the new clock tensor factor $|p_t = -\sigma\epsilon\rangle_C$ to the reduced system state $|\psi_S^\sigma\rangle$ and then applying the inverse of the trivialization (recall that the conditioning of physical states on clock readings is not a true projection and thus invertible, cf. previous footnote). Note that embedding the reduced system states back into the physical Hilbert space is *a priori* highly ambiguous since the system state alone no longer carries any information about the clock state which had been projected out. However, here it is the physical interpretation of the reduced state as being the description of the system *S* relative to the temporal reference system *C* that singles out the embedding into the clock-neutral (i.e., temporal-reference-system-neutral) $\mathcal{H}_{\text{phys}}$. This physical interpretation is, of course, added information, but it is crucial. For a more detailed discussion of this topic, see Vanrietvelde et al. [71].

Proof: The proof of Lemma 3 in Höhn et al. [7] applies here σ-sector wise. □

Given the Heisenberg-type states in Equation (53), we may consider evolving Heisenberg observables on $\mathcal{H}_{S,\sigma}^{\text{phys}}$

$$\hat{f}_S^{\text{phys}}(\tau) = U_S^\dagger(\tau)\,\hat{f}_S^{\text{phys}}\,U_S(\tau). \tag{56}$$

Indeed, the following theorem shows that these Heisenberg observables are equivalent to the relational Dirac observables on the σ-sector of the physical Hilbert space $\mathcal{H}_\sigma$, thereby demonstrating that the quantum symmetry reduction map yields a relational Heisenberg picture. To this end, we employ these reduction maps and their inverses to define another encoding operation $\mathcal{E}_{\mathbf{H}}^\sigma: \mathcal{L}\left(\mathcal{H}_{S,\sigma}^{\text{phys}}\right) \to \mathcal{L}\left(\mathcal{H}_\sigma\right)$,

$$\mathcal{E}_{\mathbf{H}}^\sigma\left(\hat{f}_S^{\text{phys}}(\tau)\right) \coloneqq (\mathcal{R}_{\mathbf{H}}^\sigma)^{-1}\,\hat{f}_S^{\text{phys}}(\tau)\,\mathcal{R}_{\mathbf{H}}^\sigma. \tag{57}$$

Theorem 3. *Let $\hat{f}_S \in \mathcal{L}(\mathcal{H}_S)$. The quantum relational Dirac observable $\hat{F}_{f_S,T}(\tau)$ acting on $\mathcal{H}_{\text{phys}}$, Equation (31), reduces under $\mathcal{R}_{\boldsymbol{H}}^\sigma$ to the corresponding projected evolving observable in the Heisenberg picture on $\mathcal{H}_{S,\sigma}^{\text{phys}}$, Equation (56), i.e.,*

$$\begin{aligned}\mathcal{R}_{\boldsymbol{H}}^\sigma\,\hat{F}_{f_S,T}(\tau)\,(\mathcal{R}_{\boldsymbol{H}}^\sigma)^{-1} &= \Pi_{\sigma_{SC}}\hat{f}_S(\tau)\,\Pi_{\sigma_{SC}}\\ &\equiv \hat{f}_S^{\text{phys}}(\tau).\end{aligned}$$

Conversely, let $\hat{f}_S^{\text{phys}}(\tau) \in \mathcal{L}\left(\mathcal{H}_{S,\sigma}^{\text{phys}}\right)$ be any evolving Heisenberg observable. The encoding operation in Equation (57) of system observables coincides on the physical Hilbert space $\mathcal{H}_{\text{phys}}$ *with the quantum relational Dirac observables in Equation (31) projected into the σ-sector:*

$$\mathcal{E}_{\boldsymbol{H}}^\sigma\left(\hat{f}_S^{\text{phys}}(\tau)\right) \approx \hat{F}^\sigma_{\hat{f}_S^{\text{phys}},T}(\tau). \tag{58}$$

Proof: The proof of Theorem 5 in Höhn et al. [7] applies here per σ-sector. □

Once more, the theorem establishes an equivalence between the full sets of relational Dirac observables relative to clock E_T on $\mathcal{H}_\sigma$ and evolving Heisenberg observables on the physical system Hilbert space of the σ-modes $\mathcal{H}_{S,\sigma}^{\text{phys}}$. Hence, one can recover the action of the relational Dirac observables only σ-sector wise from the Heisenberg observables.

Lemma 8 and Theorem 2 directly imply that we again have preservation of expectation values per σ-sector, as the following theorem shows.

Theorem 4. *Let $\hat{f}_S \in \mathcal{L}(\mathcal{H}_S)$ and $\hat{f}_S^{\text{phys}}(\tau) = U_S^\dagger(\tau)\,\Pi_{\sigma_{SC}}\hat{f}_S\,\Pi_{\sigma_{SC}}U_S(\tau)$ be its associated evolving Heisenberg operator on $\mathcal{H}_S^{\text{phys}}$. Then*

$$\langle\psi_{\text{phys}}|\hat{F}^\sigma_{f_S,T}(\tau)\,|\phi_{\text{phys}}\rangle_{\text{phys}} = \frac{1}{2}\,\langle\psi_S^\sigma\,|\hat{f}_S^{\text{phys}}(\tau)\,|\phi_S^\sigma\rangle\,,$$

where $|\psi_S^\sigma\rangle = \sqrt{2}\,\mathcal{R}_{\boldsymbol{H}}^\sigma\,|\psi_{\text{phys}}\rangle$.

Proof: The proof of Theorem 6 in Höhn et al. [7] applies here per σ-sector. □

Therefore, the quantum symmetry reduction map $\mathcal{R}_{\mathbf{H}}$ is an isometry, as we state in the following corollary.

Corollary 3. *Setting $\hat{f}_S = \Pi_{\sigma_{SC}}$ in Theorem 4 yields*

$$\langle\psi_{\text{phys}}|\,\theta(-\sigma\,\hat{p}_t)\otimes I_S\,|\phi_{\text{phys}}\rangle_{\text{phys}} = \frac{1}{2}\,\langle\psi_S^\sigma\,|\phi_S^\sigma\rangle\ ,$$

where $|\psi_S^\sigma\rangle = \sqrt{2}\,\mathcal{R}_{\boldsymbol{H}}^\sigma\,|\psi_{\text{phys}}\rangle$. Hence,

$$\langle\psi_{\text{phys}}|\phi_{\text{phys}}\rangle_{\text{phys}} = \frac{1}{2}\sum_\sigma\,\langle\psi_S^\sigma\,|\,\phi_S^\sigma\rangle\ .$$

Accordingly, we can work with normalized states in each reduced σ-sector and in the physical Hilbert space simultaneously.

In conclusion:

(1) Applying the quantum symmetry reduction map $\mathcal{R}_{\mathbf{H}}^\sigma$ to the clock-neutral picture on the physical Hilbert space $\mathcal{H}_{\text{phys}}$ yields a *relational Heisenberg picture* with respect to the clock C on the physical system Hilbert space of the σ-modes, $\mathcal{H}_{S,\sigma}^{\text{phys}}$.
(2) σ-sector wise, the relational quantum dynamics encoded in the relational Dirac observables on the physical Hilbert space is equivalent to the dynamics in the relational Heisenberg picture on the physical system Hilbert space.
(3) Given the invertibility of the reduction map, Theorem 4 formally shows that if $\hat{f}_S^{\text{phys}}(\tau)$ is self-adjoint on $\mathcal{H}_{S,\sigma}^{\text{phys}}$, then so is $\hat{F}_{f_S,T}^\sigma(\tau)$ on $\mathcal{H}_\sigma$.

5.1.4. Equivalence of Dynamics (ii) and (iii)

The previous subsections establish a σ-sector wise equivalence between the relational dynamics, on the one hand, in the clock-neutral picture of Dirac quantization and, on the other, the relational Schrödinger and Heisenberg pictures, obtained through Page-Wootters reduction and quantum deparametrization, respectively. It is thus evident that also the relational Schrödinger and Heisenberg pictures are indeed equivalent up to the unitary $U_S(\tau)$ as they should. This is directly implied by Lemma 8 which shows that the Page-Wotters and quantum symmetry reduction maps are (weakly) related by $U_S(\tau)$.

5.2. Quantum Analogs of Gauge-Fixing and Gauge-Invariant Extensions

In contrast to the classical constraint surface, the "quantum constraint surface" $\mathcal{H}_{\text{phys}}$ is automatically gauge-invariant since the exponentiation of the symmetry generator $\hat{C}_H$ acts trivially on all physical states and Dirac observables. Nevertheless, extending the interpretation established in Höhn et al. [7], we can understand the quantum symmetry reduction map $\mathcal{R}_{\mathbf{H}}$ [and given their unitary relation, also $\mathcal{R}_{\mathbf{S}}^\sigma(\tau)$] as the quantum analog of a classical phase space reduction through gauge-fixing. For completeness, the latter procedure is explained in **Supplementary Material** for the class of models of this article. In particular, we may think of the physical system Hilbert space $\mathcal{H}_{S,\sigma}^{\text{phys}}$ for the σ-sector as the quantum analog of the gauge-fixed reduced phase space obtained by imposing for example the gauge $T = 0$ on the classical σ-frequency sector $\mathcal{C}_\sigma$[19]. Also classically, one obtains two identically looking gauge-fixed reduced phase spaces, one for each frequency sector. Consequently, the relational Schrödinger and Heisenberg pictures can both be understood as the quantum analog of a gauge-fixed formulation of a manifestly gauge-invariant theory.

In this light, Theorems 1 and 3 imply that the encoding operations of system observables in Equations (45) and (57) can be understood as the quantum analog of gauge-invariantly extending a gauge-fixed quantity (see also [7]). Similarly, the alternative physical inner product in the second line of Equation (47) is the quantum analog of a gauge-fixed version of the manifestly gauge-invariant physical inner product obtained through group averaging in Equation (18). Indeed, $\sum_\sigma\,\langle\psi_{\text{phys}}|(|\tau,\sigma\rangle\langle\tau,\sigma|\otimes I_S)|\phi_{\text{phys}}\rangle_{\text{kin}}$ is the (kinematical) expectation value of the 'projector' onto clock time τ in physical states. However, it is clear from the unitarity of the Schrödinger dynamics on $\mathcal{H}_{S,\sigma}^{\text{phys}}$ and Equation (47) that this inner product does not depend on τ (the "gauge"), in line with the interpretation of it being the quantum analog of a gauge-fixed version of a manifestly gauge-invariant quantity.

5.3. Interlude: Alternative Route

As an aside, we mention that there exists an alternative route to establishing a trinity for clock Hamiltonians quadratic in momenta. This again exploits the reducible representation on the physical Hilbert space. The σ-sector of $\mathcal{H}_{\text{phys}}$ is defined by the constraint $\hat{C}_\sigma = \hat{H}'_C + \hat{H}'_S$, where $\hat{H}'_C := \frac{\hat{p}_t}{\sqrt{2}}$ and $\hat{H}'_S := \sigma\sqrt{-s\,\hat{H}_S}$. Clearly, $\hat{H}'_C$ is now a non-degenerate clock Hamiltonian. In Höhn et al. [7] we established the trinity for non-degenerate clock Hamiltonians and the σ-sector defined by $\hat{C}_\sigma = \hat{H}'_C + \hat{H}'_S$ yields a special case of that. This immediately implies a trinity per σ-sector, however, now relative to a clock POVM which is covariant with respect to $\hat{H}'_C$. It is evident that the covariant clock POVM is in this case defined through the eigenstates of $\hat{t}$ which (up to a factor of $\sqrt{2}$) is also the first moment of the POVM. Indeed, the equivalence between the clock-neutral Dirac quantization and the relational Heisenberg picture has previously been established for models with quadratic clock Hamiltonians precisely in this manner in Höhn and Vanrietvelde [30] and Höhn [31] (for a related discussion see also [29, 38]). However, as mentioned in section 4.1, one either has to regularize the relational observables or write them as functions of affine, rather than canonical pairs of evolving degrees of freedom. This is a consequence of the square root nature of $\hat{H}'_S$. None of these extra steps were needed in the trinity construction of this

[19]However, note that the quantization of this classical reduced phase space will in some cases, but not in general, coincide with the quantum theory on $\mathcal{H}_{S,\sigma}^{\text{phys}}$ due to the generic inequivalence between Dirac and reduced quantization (see **Supplementary Material**).

article, which is based on a clock POVM which is covariant with respect to $s\,\frac{\hat{p}_t^2}{2}$, rather than $\hat{p}_t/\sqrt{2}$.

6. RELATIVISTIC LOCALIZATION: ADDRESSING KUCHAŘ'S CRITICISM

In his seminal review on the problem of time, Kuchař raised three criticisms against the Page-Wootters formalism [1]: the Page-Wootters conditional probability in Equation (37) (a) yields the wrong localization probabilities for a relativistic particle, (b) violates the Hamiltonian constraint, and (c) produces incorrect transition probabilities. As mentioned in the introduction, criticisms (b) and (c) have been resolved in Höhn et al. [7]—see Theorem 2 which extends the resolution of (b) to the present class of models.

Here, we shall now also address Kuchař's first criticism (a) on relativistic localization, which is more subtle to resolve. The main reason, as is well-known from the theorems of Perez-Wilde [92] and Malament [93] (see also the discussion in [94, 95]), is that there is no relativistically covariant position-operator-based notion of localization which is compatible with relativistic causality and positivity of energy. This is a key motivation for quantum field theory [90, 94]—and here a challenge for specifying what the "right" localization probability for a relativistic particle should be. Instead, one may resort to an approximate and relativistically non-covariant notion of localization proposed by Newton and Wigner [90, 91]. We will address criticism (a) by demonstrating that our formulation of the Page-Wootters formalism, based on covariant clocks for relativistic models, yields a localization in such an approximate sense.

For the sake of an explicit argument, we shall, just like Kuchař [1], focus solely on the free relativistic particle, whose Hamiltonian constraint reads (cf. **Table 1**)

$$\hat{C}_H = -\hat{p}_t^2 + \hat{\boldsymbol{p}}^2 + m^2\,,$$

where $\hat{\boldsymbol{p}}$ denotes the spatial momentum vector. However, the argument could be extended to the entire class of models considered in this manuscript. It is straightforward to check that the physical inner product Equation (18) reads in this case [31, 104][20]

$$\langle\phi_{\rm phys}|\psi_{\rm phys}\rangle_{\rm phys} = \int_{\mathbb{R}^3}\frac{d^3\boldsymbol{p}}{2\,\varepsilon_p}\Big[\phi^*_{\rm kin}(\varepsilon_p,\boldsymbol{p})\,\psi_{\rm kin}(\varepsilon_p,\boldsymbol{p}) + \phi^*_{\rm kin}(-\varepsilon_p,\boldsymbol{p})\,\psi_{\rm kin}(-\varepsilon_p,\boldsymbol{p})\Big], \tag{59}$$

where $\varepsilon_p = \sqrt{\boldsymbol{p}^2+m^2}$ is the relativistic energy of the particle and the first and second term in the integrand correspond to negative and positive frequency modes, respectively. Fourier transforming to solutions to the Klein-Gordon equation in Minkowski space

$$\psi^\sigma_{\rm phys}(t,\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}}\int_{\mathbb{R}^3}\frac{d^3\boldsymbol{p}}{2\,\varepsilon_p}\,e^{i(\boldsymbol{x}\cdot\boldsymbol{p}-\sigma\,t\,\varepsilon_p)}\,\psi_{\rm kin}(-\sigma\,\varepsilon_p,\boldsymbol{p})\,,$$

one may further check that [31, 104]

$$\langle\phi_{\rm phys}|\psi_{\rm phys}\rangle_{\rm phys} = \left(\phi^+_{\rm phys},\psi^+_{\rm phys}\right)_{\rm KG} - \left(\phi^-_{\rm phys},\psi^-_{\rm phys}\right)_{\rm KG}\,, \tag{60}$$

where

$$\left(\phi^\sigma_{\rm phys},\psi^\sigma_{\rm phys}\right)_{\rm KG} = i\int_{\mathbb{R}^3}d^3\boldsymbol{x}\Big[\left(\phi^\sigma_{\rm phys}(t,\boldsymbol{x})\right)^*\partial_t\,\psi^\sigma_{\rm phys}(t,\boldsymbol{x}) - \left(\partial_t\,\phi^\sigma_{\rm phys}(t,\boldsymbol{x})\right)^*\psi^\sigma_{\rm phys}(t,\boldsymbol{x})\Big], \tag{61}$$

is the Klein-Gordon inner product in which positive frequency modes are positive semi-definite, negative frequency modes are negative semi-definite and positive and negative frequency modes are mutually orthogonal. The physical inner product is thus equivalent to the Klein-Gordon inner product (with correctly inverted sign for the negative frequency modes), which provides the correct and conserved normalization for the free relativistic particle[21]. This raises hopes that the conditional probabilities of the Page-Wootters formalism may yield the correct localization probability for the relativistic particle. Note that so far we have not yet made a choice of time operator.

Suppose now that the Minkowski time operator $\hat{t}$, quantized as a self-adjoint operator on $\mathcal{H}_{\rm kin}$, is used to define the projector onto clock time t as $e_C(t) = |t\rangle\langle t|$ and $e_{\boldsymbol{x}} = |\boldsymbol{x}\rangle\langle\boldsymbol{x}|$ is the projector onto position $\boldsymbol{x}$. This time operator is not covariant with respect to the quadratic clock Hamiltonian. The conditional probability Equation (37) then becomes

$$\text{Prob}\left(\boldsymbol{x}\text{ when }t\right) = \frac{|\psi_{\rm phys}(t,\boldsymbol{x})|^2}{\int_{\mathbb{R}^3}d^3\boldsymbol{x}'\,|\psi_{\rm phys}(t,\boldsymbol{x}')|^2}, \tag{62}$$

where $\psi_{\rm phys}(t,\boldsymbol{x}) = (\langle t|\otimes\langle\boldsymbol{x}|)\,|\psi_{\rm phys}\rangle$ is a general solution to the Klein-Gordon equation. As Kuchař pointed out [1], while this would be the correct localization probability for a non-relativistic particle, it is the wrong result for a relativistic particle. Indeed, apart from not separating positive and negative frequency modes, which is necessary for a probabilistic interpretation e.g., if $\psi_{\rm phys}$ contains both positive and negative frequency modes then the denominator in Equation 62 is not conserved), Equation (62) neither coincides with the charge density of the Klein-Gordon current in Equation (61), nor with the Newton-Wigner approximate localization probability [90, 91]. In particular, one can *not* interpret a solution $\psi_{\rm phys}(t,\boldsymbol{x})$ to the Klein-Gordon equation as a probability amplitude to find the relativistic particle at position $\boldsymbol{x}$ at time t. The reason, as explained in Haag [90], is that the conserved density is the one in Equation (61) and $\psi_{\rm phys}$ and $\partial_t\psi_{\rm phys}$ inside it are not only dependent, but related by a non-local convolution

$$\partial_t\psi_{\rm phys}(t,\boldsymbol{x}) = \int_{\mathbb{R}^3}d^3\boldsymbol{x}'\,\varepsilon(\boldsymbol{x}-\boldsymbol{x}')\,\psi_{\rm phys}(t,\boldsymbol{x}')\,,$$

[20]Note that here we have a doubly degenerate system energy $\hat{H}_S = \hat{\boldsymbol{p}}^2 + m^2$ in contrast to the expression in Equation (18) where we ignored degeneracy.

[21]This also resolves the normalization issue appearing in Diaz et al. [50] where physical states were normalized using the kinematical, rather than physical inner product, thus yielding a divergent normalization (for equal mass states) in contrast to here.

where $\varepsilon(\boldsymbol{x}-\boldsymbol{x}')$ is the Fourier transform of $-i\,\varepsilon_p$.

By contrast, let us now exhibit what form of conditional probabilities the covariant clock POVM E_T of section 4.2 gives rise to. We now insert $e_C(\tau) = \sum_\sigma |\tau,\sigma\rangle\langle\tau,\sigma|$ and, as before, $e_{\boldsymbol{x}}$ into the conditional probability Equation (37). The crucial difference between the covariant clock POVM E_T and the clock operator $\hat{t}$ (which is covariant with respect to $\hat{C}_\sigma$, but not $\hat{C}_H$) is that the denominator of Equation (37) is equal to the physical inner product in the former case (see Corollary 2) but not in the latter[22]. Supposing that we work with normalized physical system states $\langle\psi_S^\sigma(\tau)|\psi_S^\sigma(\tau)\rangle = 1$, Theorem 2 implies

$$\begin{aligned}\text{Prob}\left(\boldsymbol{x}\text{ when }\tau\right) &= \frac{1}{2}\sum_\sigma |\psi_S^\sigma(\tau,\boldsymbol{x})|^2 \\ &= \langle\psi_{\text{phys}}|\hat{F}_{e_{\boldsymbol{x}},T}(\tau)|\psi_{\text{phys}}\rangle_{\text{phys}}\,,\end{aligned} \tag{63}$$

where $\psi_S^\sigma(\tau,\boldsymbol{x}) \coloneqq \sqrt{2}(\langle\tau,\sigma|\otimes\langle\boldsymbol{x}|)\,|\psi_{\text{phys}}\rangle$ and τ is now *not* Minkowski time. For concreteness, let us now focus on positive frequency modes. Using Equations (15) and (25), one obtains

$$\psi_S^+(\tau,\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}}\int_{\mathbb{R}^3}\frac{d^3\boldsymbol{p}}{\sqrt{2\,\varepsilon_p}}e^{i(\boldsymbol{x}\cdot\boldsymbol{p}-\tau\,\varepsilon_p^2)}\,\psi_{\text{kin}}(-\varepsilon_p,\boldsymbol{p})\,. \tag{64}$$

This is *almost* the Newton-Wigner position space wave function for positive frequency modes, which relative to Minkowski time reads [90]

$$\begin{aligned}\psi^+_{\text{NW}}(t,\boldsymbol{x}) &= \int_{\mathbb{R}^3} d^3\boldsymbol{x}\,K(\boldsymbol{x}-\boldsymbol{x}')\,\psi^+_{\text{phys}}(t,\boldsymbol{x}') \\ &= \frac{1}{(2\pi)^{3/2}}\int_{\mathbb{R}^3}\frac{d^3\boldsymbol{p}}{\sqrt{2\,\varepsilon_p}}e^{i(\boldsymbol{x}\cdot\boldsymbol{p}-t\,\varepsilon_p)}\,\psi_{\text{kin}}(-\varepsilon_p,\boldsymbol{p})\,,\end{aligned} \tag{65}$$

where $K(\boldsymbol{x})$ is the Fourier transform of $\sqrt{2\varepsilon_p}$. The key property of $|\psi^+_{\text{NW}}(t,\boldsymbol{x})|^2$ is that, while not relativistically covariant, it *does* admit the interpretation of an approximate localization probability, with accuracy of the order of the Compton wave length, for finding the particle at position $\boldsymbol{x}$ at Minkowski time t [90, 91]. In particular,

$$\left(\phi^+_{\text{phys}},\psi^+_{\text{phys}}\right)_{\text{KG}} = \int_{\mathbb{R}^3} d^3\boldsymbol{x}\left(\phi^+_{\text{NW}}(t,\boldsymbol{x})\right)^*\psi^+_{\text{NW}}(t,\boldsymbol{x})\,,$$

i.e., the Klein-Gordon inner product assumes the usual Schrödinger form for the Newton-Wigner wave function.

Noting that due to Equation (19) we can *heuristically* view τ as t/ε_p, and comparing with Equation (65) we can interpret Equation (64) as a Newton-Wigner wave function as well, but expressed relative to a different time coordinate τ. Indeed, in line with this interpretation, we find that in this case too the physical inner product, Equation (60), for the positive frequency modes assumes the form of the standard Schrödinger theory inner product

$$\left(\phi^+_{\text{phys}},\psi^+_{\text{phys}}\right)_{\text{KG}} = \int_{\mathbb{R}^3} d^3\boldsymbol{x}\left(\phi^+_S(\tau,\boldsymbol{x})\right)^*\psi^+_S(\tau,\boldsymbol{x})\,.$$

The analogous statement is true for the negative frequency modes. In that sense, Equation (63), in contrast to Equation (62), does admit the interpretation as a valid, yet approximate localization probability for the relativistic particle per frequency sector, just like in the standard Newton-Wigner case[23].

Accordingly, computing the conditional probabilities of the Page-Wootters formalism relative to the covariant clock POVM, rather than the non-covariant Minkowski time operator $\hat{t}$, leads to an acceptable localization probability for a relativistic particle, thereby addressing also Kuchař's first criticism (a). Given the equivalence of the Page-Wootters formalism with the clock-neutral and the relational Heisenberg pictures, established through the trinity in section 5, this result also equips the quantum relational Dirac observables $\hat{F}_{\boldsymbol{x},T}(\tau)$ and the evolving Heisenberg observables $\boldsymbol{x}(\tau)$ with the interpretation of providing an approximate, Newton-Wigner type localization in Minkowski space.

7. CHANGING QUANTUM CLOCKS

So far we have worked with a single choice of clock. Let us now showcase how to change from the evolution relative to one choice of clock to that relative to another. Our discussion will apply to both the relational Schrödinger picture of the Page-Wootters formalism and the relational Heisenberg picture obtained through quantum deparametrization.

For concreteness, suppose we are given a Hamiltonian constraint of the form

$$\hat{C}_H = s_1\frac{\hat{p}_1^2}{2} + s_2\frac{\hat{p}_2^2}{2} + \hat{H}_S\,, \tag{66}$$

where $s_i = \pm 1$ and $\hat{p}_i$ denotes the momentum of clock subsystem C_i, $i = 1, 2$ and we have suppressed tensor products with identity operators. In particular, suppose $\hat{H}_S$ does not depend on either of the clock degrees of freedom. We will work with the covariant clock POVM of section 4.2 for both clock choices. For example, the constraints of the relativistic particle, the flat ($k = 0$) FLRW model with a massless scalar field and the Bianchi I and II models

[22] It is instructive to see how this is linked to the (non-)covariance of the clock observable with respect to $\hat{C}_H$. Let e_C be either the covariant $e_C(\tau)$ or non-covariant $e_C(t)$. The denominator of Equation (37) reads

$$\langle\psi_{\text{phys}}|\,e_C\otimes I_S\,|\psi_{\text{phys}}\rangle_{\text{kin}} \;=\; \langle\psi_{\text{kin}}|\,\delta(\hat{C}_H)\,(e_C\otimes I_S)\,|\psi_{\text{phys}}\rangle_{\text{kin}}\,.$$

Equation (44) implies that $\delta(\hat{C}_H)(e_C(\tau)\otimes I_S)\approx I_{\text{phys}}$. This exploits the covariance and immediately shows that the denominator coincides with the physical inner product Equation (18). By contrast, the non-covariance entails $\delta(\hat{C}_H)(e_C(t)\otimes I_S)\neq I_{\text{phys}}$, so that in this case the denominator differs from the physical inner product.

[23] The physical inner product for the positive frequency solutions $\psi^+_{\text{phys}}(t,\boldsymbol{x})$ to the Klein-Gordon equation takes, of course, the standard Klein-Gordon (and not the Schrödinger) form Equation (61). Nevertheless, the kernel $K(\boldsymbol{x}-\boldsymbol{x}')$ in the non-local convolution in the first line in Equation (65) decreases quickly as a function of $m|\boldsymbol{x}-\boldsymbol{x}'|$ [90]. Hence, for massive particles, we can interpret even Equation (62) as providing an approximate localization.

from **Table 1** are of the above form[24]. Our subsequent discussion will thus directly apply to these models.

Since we will exploit the Page-Wootters and symmetry reduction maps as "quantum coordinate maps" from the clock-neutral picture to the given "clock perspective," we will be able to change from the description of the dynamics relative to one clock to that relative to another in close analogy to coordinate changes on a manifold. Due to the shape of the constraint in Equation (66) we now have superselection of Dirac observables and the physical Hilbert space across both the σ_1-frequency sectors of clock C_1 and the σ_2-frequency sectors of clock C_2. The physical Hilbert space takes the form

$$\mathcal{H}_{\text{phys}} = \bigoplus_{\sigma_1,\sigma_2} \mathcal{H}_{\sigma_1,\sigma_2}\,, \tag{67}$$

where $\mathcal{H}_{\sigma_1,\sigma_2} := \mathcal{H}_{\sigma_1} \cap \mathcal{H}_{\sigma_2}$ is the overlap of the σ_1-sector of clock C_1 and the σ_2-sector of clock C_2. As we have seen the reduction maps are only invertible per frequency sector. Hence, we will only be able to change from a given σ_1-sector to the part of the σ_2-frequency sector which is contained in it. In other words, the "quantum coordinate changes" are restricted to each overlap $\mathcal{H}_{\sigma_1,\sigma_2}$.

The method of changing temporal reference frames exhibited below is a direct extension of several previous works: Höhn and Vanrietvelde [30] and Höhn [31] developed the method σ-sector-wise for states and observables in the relational Heisenberg picture for Hamiltonians of the type in Equation (66) for two example models, but used clock operators canonically conjugate to the clock momenta $\hat{p}_i$ (and thus not a clock POVM covariant with respect to the full clock Hamiltonian). The method of transforming relational observables from one clock description to another was demonstrated in Höhn and Vanrietvelde [30] and Höhn [31] for a subset of relational Dirac observables, paying, however, detailed attention to regularization necessities arising from time-of-arrival observables [18, 23, 117–119]. Our previous article [7] developed the method comprehensively for both states and observables for clock Hamiltonians with non-degenerate and continuous spectrum in both the relational Schrödinger and Heisenberg pictures; specifically, the transformation of arbitrary relational observables corresponding to relations between S and the clocks was developed for the corresponding class of models. In Castro-Ruiz et al. [47] the clock change method was exhibited for state transformations in the relational Schrödinger picture for ideal clocks whose Hamiltonian coincides with the clock momentum itself. Our discussion can also be viewed as a full quantum extension of the semiclassical method in Bojowald et al. [32, 33] and Höhn et al. [34] which is equivalent at semiclassical order, however, also applies to clock functions which are non-monotonic, i.e., have turning points, in contrast to the other works mentioned (See [70–72, 74] for related spatial quantum frame changes).

[24]Indeed, the Hamiltonian constraint of the vacuum Bianchi I and II models can be written in the form [116]

$$\hat{C}_H = -\frac{\hat{\bar{p}}_0^2}{2} + \frac{\hat{\bar{p}}_-^2}{2} + \frac{\hat{\bar{p}}_+^2}{2} + k_+\, e^{-4\sqrt{3}\hat{\bar{\beta}}^+}\,.$$

In particular, owing to our focus on covariant clock POVMs, all the results and proofs [7] apply σ-sector-wise to the present case. However, we will also study novel effects, such as the temporal frame dependence of comparing clock readings.

7.1. State Transformations

Denote by $\mathcal{R}_I^{\sigma_i}$, where $I \in \{\mathbf{S}, \mathbf{H}\}$, the Page-Wootters or quantum symmetry reduction map to the σ_i-sector of clock C_i. For notational simplicity, we drop the dependence of the Page-Wootters reduction on the clock reading τ_i of clock C_i for the moment. The temporal frame change (TFC) map $\Lambda_{I\to J}^{\sigma_i\to\sigma_j} : \mathcal{H}^{\text{phys}}_{C_j,\sigma_i}\otimes\mathcal{H}^{\text{phys}}_{S,\sigma_i} \to \mathcal{H}^{\text{phys}}_{C_i,\sigma_j}\otimes\mathcal{H}^{\text{phys}}_{S,\sigma_j}$ from clock C_i's σ_i-sector to clock C_j's σ_j-sector then reads

$$\Lambda_{I\to J}^{\sigma_i\to\sigma_j} := \mathcal{R}_J^{\sigma_j} \circ \left(\mathcal{R}_I^{\sigma_i}\right)^{-1}\,. \tag{68}$$

Here $\mathcal{H}^{\text{phys}}_{C_j,\sigma_i}$ denotes the physical clock C_j Hilbert space corresponding to the σ_i-sector of clock C_i, i.e., the subspace of $\mathcal{H}_{C_j}$ compatible with solutions to the constraint Equation (66) and similarly for the other Hilbert spaces. When $I \neq J$ in Equation (68), then the TFC map changes not only the temporal reference frame, but also between the corresponding relational Heisenberg and Schrödinger pictures. Let us write $\Lambda_I^{\sigma_i\to\sigma_j} := \Lambda_{I\to I}^{\sigma_i\to\sigma_j}$ when no relational picture change takes place.

More explicitly, the TFC map from the σ_1-frequency sector of clock C_1 in the relational Schrödinger picture to the σ_2-frequency sector of clock C_2 in the relational Schrödinger picture takes the form

$$\Lambda_{\mathbf{S}}^{\sigma_1\to\sigma_2} = \left(\langle\tau_2,\sigma_2|\otimes I_{C_1S}\right)\,\delta(\hat{C}_H)\,\left(|\tau_1,\sigma_1\rangle\otimes I_{C_2S}\right)\,.$$

Here we have made use of Equations (38) and (42) and the covariant clock states Equation (25) for both clocks. The reduced states transform under this map as follows:

$$\left(\theta(-\sigma_1\,\hat{p}_1)\otimes I_S\right)\,|\psi^{\sigma_2}_{C_1S|C_2}(\tau_2)\rangle = \Lambda_{\mathbf{S}}^{\sigma_1\to\sigma_2}\;|\psi^{\sigma_1}_{C_2S|C_1}(\tau_1)\rangle\,, \tag{69}$$

where we made use of Equation (43) and $|\psi^{\sigma_i}_{C_jS|C_i}(\tau_i)\rangle$ is the relational Schrödinger picture state of clock C_j and system S relative to clock C_i, which is chosen as the temporal reference frame, in its σ_i-sector. In other words, the Heaviside-function on the l.h.s. highlights that we can only map from the σ_1-sector of clock C_1 to that part of the σ_2-sector of clock C_2 which is contained in the σ_1-sector of clock C_1. This is clear, because any reduction map is only invertible on its associated σ-sector: from the σ_1 relational Schrödinger picture one can only recover the σ_1-sector of the physical Hilbert space. Hence, the subsequent Page-Wootters reduction map to the σ_2-sector of clock C_2 in Equation (68) can then only yield information in the overlap of the σ_1- and σ_2-sectors in the physical Hilbert space (see also [30, 31] for explicit examples of this situation in the relational Heisenberg picture). This is a manifestation of the superselection across both the σ_1- and the σ_2-sectors.

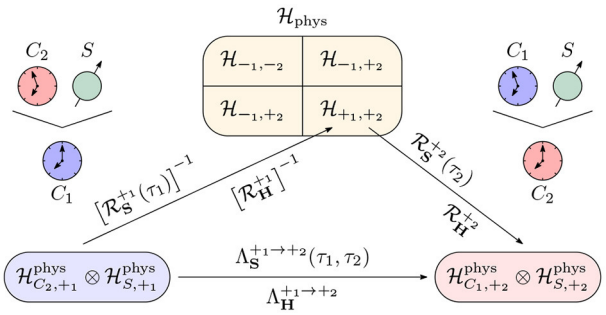

FIGURE 3 | Schematic representation of a temporal frame change, as defined through Equation (68). The figure encompasses both the relational Schrödinger picture of the Page-Wotters formalism and the relational Heisenberg picture of the quantum deparametrization (as well as their mixtures). Viewing the reduction maps $\mathcal{R}_I^{+_I}$ as quantum coordinate maps, any such temporal frame change takes the form of a quantum coordinate transformation from the description relative to clock C_1 to the one relative to clock C_2. Just as coordinate transformation pass through the reference-frame-neutral manifold, the quantum coordinate transformations pass through the clock-neutral physical Hilbert space in line with the general discussion of the clock-neutral structure and quantum general covariance in section 2. Due to the double superselection rule, the quantum coordinate transformations have to preserve the overlaps of the frequency sectors of C_1 and C_2. Here we illustrate the example of the overlap of the positive frequency sectors of both clocks, so that the corresponding frame transformation passes through $\mathcal{H}_{+_1,+_2}$ (cf. Equation 67).

Similarly, the TFC map from the σ_1-frequency sector of clock C_1 in the relational Heisenberg picture to the σ_2-frequency sector of clock C_2 in the relational Heisenberg picture reads

$$\Lambda_{\mathbf{H}}^{\sigma_1\to\sigma_2} = \left(\langle\tau_2,\sigma_2| \otimes U_{C_1S}^{\dagger}(\tau_2)\right) \times \delta(\hat{C}_H)\ \left(|\tau_1,\sigma_1\rangle \otimes U_{C_2S}(\tau_1)\right), \tag{70}$$

where we have made use of Lemma 8. Using the same lemma, the reduced states transform under this map in complete analogy to Equation (69)

$$\left[\theta(-\sigma_1\,\hat{p}_1)\otimes I_S\right]\,|\psi_{C_1S|C_2}^{\sigma_2}\rangle = \Lambda_{\mathbf{S}}^{\sigma_1\to\sigma_2}\,|\psi_{C_2S|C_1}^{\sigma_1}\rangle\,,$$

with $U_{C_jS}(\tau_i) := \exp\left[-i\,\tau_i\,(s_j\frac{\hat{p}_j^2}{2}+\hat{H}_S)\right]$ the evolution operator of the composite system of clock C_j and system S relative to clock C_i.

Note that, interpreting the reduction maps as the "quantum coordinate maps" taking one from the clock-neutral physical Hilbert space to a specific "clock perspective," any such TFC map in Equation (68) takes the same compositional form as coordinate changes on a manifold. In particular, any such temporal frame change proceeds by mapping via the clock-neutral physical Hilbert space in analogy to how coordinate changes always proceed via the manifold (see **Figure 3**). This observation lies at the heart of the perspective-neutral approach to quantum reference frame changes [7, 30, 31, 71, 72]. It is also the reason why we may interpret the physical Hilbert space as a clock-neutral structure, providing a description of the dynamics prior to having chosen a temporal reference frame relative to which the other degrees of freedom evolve. In line with this, in terms of different one-parameter families of relational Dirac observables, the physical Hilbert space contains the complete information about the dynamics relative to all the different possible clock choices at once.

7.2. Observable Transformations

Just as we transformed reduced states from the perspective of one clock to the perspective of another by passing through the gauge-invariant physical Hilbert space (see **Figure 3**), we now transform the description of observables relative to one clock to that relative to the other by passing through the gauge-invariant algebra of Dirac observables on the physical Hilbert space. The observable transformations will thus be dual to the state transformations. The idea is always that we describe the *same* physics, encoded in the gauge-invariant states and observables of the clock-neutral physical Hilbert space, but relative to different temporal frames. Again, we have to pay attention to the two superselection rules on the clock-neutral physical Hilbert space and we will demonstrate the observable transformations separately for the relational Schrödinger and Heisenberg pictures.

7.2.1. Observable Transformations in the Relational Schrödinger Picture

Suppose we are given an observable $\hat{O}_{C_2S|C_1}^{\text{phys}}$ describing certain properties of the composite system C_2S in the relational Schrödinger picture of clock C_1 in either frequency sector of the latter[25]. Owing to Theorem 1, we can write this as a reduction of a corresponding relational Dirac observable on $\mathcal{H}_{\text{phys}}$:

$$\mathcal{R}_{\mathbf{S}}^{\sigma_1}(\tau_1)\ \hat{F}_{O_{C_2S|C_1},T_1}(\tau_1)\,(\mathcal{R}_{\mathbf{S}}^{\sigma_1}(\tau_1))^{-1} = \hat{O}_{C_2S|C_1}^{\text{phys}}.$$

We can now also map the same relational Dirac observable into the σ_2-sector of the relational Schrödinger picture of clock C_2:

$$\hat{O}_{C_1S|C_2}^{\text{phys}}(\tau_1,\tau_2) := \mathcal{R}_{\mathbf{S}}^{\sigma_2}(\tau_2)\ \hat{F}_{O_{C_2S|C_1},T_1}(\tau_1)\,(\mathcal{R}_{\mathbf{S}}^{\sigma_2}(\tau_2))^{-1}. \tag{71}$$

The result will be the image of the original observable $\hat{O}_{C_2S|C_1}^{\text{phys}}$, describing properties of C_2S relative to C_1, in the "perspective" of clock C_2. Hence, if $\hat{O}_{C_2S|C_1}^{\text{phys}}$ depends non-trivially on C_2, an indirect self-reference effect occurs in the last equation [7]. Notice that, while the original observable in the Schrödinger picture of C_1 is independent of the evolution parameter τ_1, the

[25]Recall that the label "phys" highlights that the observable acts on the physical C_2S Hilbert space, i.e., on $\Pi_{\sigma_{C_2SC_1}}\left(\mathcal{H}_{C_2}\otimes\mathcal{H}_S\right)$, where $\Pi_{\sigma_{C_2SC_1}} = \theta\left[-s_1\left(s_2\frac{\hat{p}_2^2}{2}+\hat{H}_S\right)\right]$ is the projector onto the subspace corresponding to the spectrum $\sigma_{C_2SC_1} := \text{Spec}\left(s_2\frac{\hat{p}_2^2}{2}+\hat{H}_S\right)\cap\text{Spec}\left(-s_1\frac{\hat{p}_1^2}{2}\right)$ permitted by the constraint Equation (66) (cf. Equation 17). We can thus also understand this observable as a projection $\hat{O}_{C_2S|C_1}^{\text{phys}} := \Pi_{\sigma_{C_2SC_1}}\,\hat{O}_{C_2S|C_1}\,\Pi_{\sigma_{C_2SC_1}}$ of some observable $\hat{O}_{C_2S|C_1}\in\mathcal{L}(\mathcal{H}_{C_2}\otimes\mathcal{H}_S)$; cf. Equation (36).

description of that same observable in the Schrödinger picture relative to clock C_1 will generally depend on both evolution parameters τ_1, τ_2. The dependence on τ_1 is a consequence of it being the reduction of a relational Dirac observable with evolution parameter τ_1, but into the "perspective" of C_2. The possible τ_2 dependence may arise as a consequence of said indirect self-reference. For example, suppose $\hat{O}^{\text{phys}}_{C_2S|C_1} = \hat{T}_2 \otimes I_S$ so that the relational Dirac observable is $\hat{F}_{O_{C_2S|C_1},T_1}(\tau_1) = \hat{F}_{T_2,T_1}(\tau_1)$. The observable on the l.h.s. in Equation (71) then describes how the first moment operator $\hat{T}_2$ associated with C_2 evolves relative to C_1 from the "perspective" of C_2; this certainly should yield a τ_1 dependence. We will explain this in more detail shortly.

Taking into account the two superselection rules across the σ_1- and σ_2-sectors, these observations imply that observable transformations from the relational Schrödinger picture of the σ_1-sector of clock C_1 into the relational Schrödinger picture of the σ_2-sector of clock C_2 read

$$\begin{aligned}
&\Lambda_{\mathbf{S}}^{\sigma_1\to\sigma_2}\,\hat{O}^{\text{phys}}_{C_2S|C_1}\,\left(\Lambda_{\mathbf{S}}^{\sigma_1\to\sigma_2}\right)^{-1} = \left(\Lambda_{\mathbf{S}}^{\sigma_2\to\sigma_1}\right)^{-1}\,\hat{O}^{\text{phys}}_{C_2S|C_1}\,\Lambda_{\mathbf{S}}^{\sigma_2\to\sigma_1}\\
&= \mathcal{R}^{\sigma_2}_{\mathbf{S}}(\tau_2)\circ\left(\mathcal{R}^{\sigma_1}_{\mathbf{S}}(\tau_1)\right)^{-1}\,\hat{O}^{\text{phys}}_{C_2S|C_1}\,\mathcal{R}^{\sigma_1}_{\mathbf{S}}(\tau_1)\circ\left(\mathcal{R}^{\sigma_2}_{\mathbf{S}}(\tau_2)\right)^{-1}\\
&= \mathcal{R}^{\sigma_2}_{\mathbf{S}}(\tau_2)\;\mathcal{E}^{\tau_1,\sigma_1}_{\mathbf{S}}\left(\hat{O}^{\text{phys}}_{C_2S|C_1}\right)\left(\mathcal{R}^{\sigma_2}_{\mathbf{S}}(\tau_2)\right)^{-1} \qquad (72)\\
&\approx \mathcal{R}^{\sigma_2}_{\mathbf{S}}(\tau_2)\,\hat{F}_{O_{C_2S|C_1},T_1}(\tau_1)\,\left(\theta(-\sigma_1\hat{p}_1)\otimes I_{C_2S}\right)\left(\mathcal{R}^{\sigma_2}_{\mathbf{S}}(\tau_2)\right)^{-1}\\
&= \left(\theta(-\sigma_1\hat{p}_1)\otimes I_S\right)\,\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1,\tau_2)\,.
\end{aligned}$$

In the second line we made use of Equation (45), in the third of Theorem 1, and in the fourth of Equation (71) and the fact that $\theta(-\sigma_1\hat{p}_1)$ commutes with the reduction map of the C_2 clock and with $\hat{F}_{O_{C_2S|C_1},T_1}(\tau_1)$ (see Lemma 5).

Observe that the structure of this transformation shows that reduced observables relative to one clock will transform always via the gauge-invariant Dirac observable algebra to reduced observables relative to another clock.

Using Equations (38), (42), and (45), we can write this transformation as

$$\begin{aligned}
\left(\theta(-\sigma_1\hat{p}_1)\otimes I_S\right)\,\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1,\tau_2) &= \langle\tau_2,\sigma_2|\;\delta(\hat{C}_H)\\
&\left(|\tau_1,\sigma_1\rangle\langle\tau_1,\sigma_1|\otimes\hat{O}^{\text{phys}}_{C_2S|C_1}\right)\delta(\hat{C}_H)\;|\tau_2,\sigma_2\rangle\;. \qquad (73)
\end{aligned}$$

This transformation reveals that expectation values are preserved in the following manner:

$$\begin{aligned}
&\langle\psi^{\sigma_2}_{C_1S|C_2}(\tau_2)|\;\left(\theta(-\sigma_1\hat{p}_1)\otimes I_S\right)\,\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1,\tau_2)\;|\phi^{\sigma_2}_{C_1S|C_2}(\tau_2)\rangle\\
&= \langle\psi^{\sigma_1}_{C_2S|C_1}(\tau_1)|\;\left(\theta(-\sigma_2\hat{p}_2)\otimes I_S\right)\,\hat{O}^{\text{phys}}_{C_2S|C_1}\left(\theta(-\sigma_2\hat{p}_2)\otimes I_S\right)\;|\phi^{\sigma_1}_{C_2S|C_1}(\tau_1)\rangle\;. \qquad (74)
\end{aligned}$$

The projectors onto the σ_2-sector on the r.h.s. appears because the C_2 reduction map in Equation (72) induces such a projection (compare this with the state transformations Equation (69) which are dual). In other words, *only the physical information in the overlap of the σ_1- and σ_2-sector is preserved when changing from the description relative to clock C_1 to one relative to clock C_2, or vice versa*. Once more, this is a direct consequence of the double superselection rule induced by the shape of the constraint Equation (66).

7.2.2. Observable Transformations in the Relational Heisenberg Picture

The argumentation for the relational Heisenberg picture proceeds in complete analogy. We thus just quote the results, which immediately follow from those of the previous subsection through use of Lemma 8. Of course, in this case, the reduced observables have an explicit dependence on the evolution parameter (Equation 56).

The observable transformations from the relational Heisenberg picture of the σ_1-sector of clock C_1 into the relational Heisenberg picture of the σ_2-sector of clock C_2 are given by

$$\begin{aligned}
&\Lambda_{\mathbf{H}}^{\sigma_1\to\sigma_2}\,\hat{O}^{\text{phys}}_{C_2S|C_1}(\tau_1)\,\left(\Lambda_{\mathbf{H}}^{\sigma_1\to\sigma_2}\right)^{-1}\\
&\approx \mathcal{R}^{\sigma_2}_{\mathbf{H}}\,\hat{F}_{O_{C_2S|C_1},T_1}(\tau_1)\,\left(\theta(-\sigma_1\hat{p}_1)\otimes I_{C_2S}\right)\left(\mathcal{R}^{\sigma_2}_{\mathbf{H}}\right)^{-1}\\
&= \left(\theta(-\sigma_1\hat{p}_1)\otimes I_S\right)\,U^{\dagger}_{C_1S}(\tau_2)\,\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1,\tau_2)\,U_{C_1S}(\tau_2) \qquad (75)\\
&=: \left(\theta(-\sigma_1\hat{p}_1)\otimes I_S\right)\,\hat{O}^{\mathbf{H}}_{C_1S|C_2}(\tau_1,\tau_2)\,,
\end{aligned}$$

where $\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1,\tau_2)$ is given by Equations (71) and (73). Thanks to the double superselection rule, this transformation preserves expectation values again per overlap of a σ_1- with a σ_2-sector, in obvious analogy to Equation (74).

7.3. Occurrence of Indirect Clock Self-Reference

Finally, let us now come back to the indirect self-reference effect of clock C_2 alluded to above. The following theorem, which is adapted from Höhn et al. [7] and whose proof applies here per pair of σ_1- and σ_2-sector, reveals the necessary and sufficient conditions for this indirect self-reference to occur:

Theorem 5. *Consider an operator $\hat{O}^{\text{phys}}_{C_2S|C_1} \in \mathcal{L}(\mathcal{H}^{\text{phys}}_{C_2}\otimes\mathcal{H}^{\text{phys}}_{S})$ of the composite system C_2S described from the perspective of clock C_1. From the perspective of clock C_2, this operator is independent of τ_2, so that $\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1,\tau_2) = \hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1) \in \mathcal{L}(\mathcal{H}^{\text{phys}}_{C_1}\otimes\mathcal{H}^{\text{phys}}_{S})$ if and only if*

$$\hat{O}^{\text{phys}}_{C_2S|C_1} = \sum_i \left(\hat{O}^{\text{phys}}_{C_2|C_1}\right)_i \otimes \left(\hat{f}^{\text{phys}}_{S|C_1}\right)_i\,,$$

where $(\hat{f}^{\text{phys}}_{S|C_1})_i$ is an operator on $\mathcal{H}^{\text{phys}}_{S}$ and $(\hat{O}^{\text{phys}}_{C_2|C_1})_i$ is a constant of motion, $\left[(\hat{O}^{\text{phys}}_{C_2|C_1})_i, s_2\frac{\hat{p}_2^2}{2}\right] = 0$. Furthermore, in this case the

transformed observable reads

$$\begin{aligned}&\left[\theta(-\sigma_1\hat{p}_1)\otimes I_S\right]\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1)\\&=\Pi_{\sigma_{C_1SC_2}}\left[\sum_i \mathcal{G}_{C_1S}\left(|\tau_1,\sigma_1\rangle\langle\tau_1,\sigma_1|\otimes\left(\hat{f}^{\text{phys}}_{S|C_1}\right)_i\right)\right.\\&\left.\langle t_2,\sigma_2|\left(\hat{O}^{\text{phys}}_{C_2|C_1}\right)_i\delta(\hat{C}_H)|t_2,\sigma_2\rangle\right]\Pi_{\sigma_{C_1SC_2}},\end{aligned}\tag{76}$$

where $\Pi_{\sigma_{C_1SC_2}}=\theta\left[-s_2\left(s_1\frac{\hat{p}_1^2}{2}+\hat{H}_S\right)\right]$ *is the projector onto the physical subspace of* $\mathcal{H}_{C_1}\otimes\mathcal{H}_S$, $|t_2,\sigma_2\rangle$ *is an arbitrary* σ_2*-sector clock state of* C_2*, and* $\mathcal{G}_{C_1S}$ *is the G-twirl over the group generated by the evolution generator* $s_1\frac{\hat{p}_1^2}{2}+\hat{H}_S$ *of the composite system* C_1S*.*

That is to say, the indirect self-reference effect and thus τ_2-dependence of Equation (72) is absent if and only if the relational Dirac observable encoding how C_2S properties evolve relative to C_1 does not contain any degrees of freedom of clock C_2 that evolve.

When $\hat{O}^{\text{phys}}_{C_2S|C_1}=I_{C_2}\otimes\hat{f}^{\text{phys}}_{S|C_1}$, i.e., only the evolution of system degrees of freedom relative to C_1 is described, Theorem 5 entails that the transformation to the description relative to C_2 simplifies as follows:

$$\begin{aligned}&\left(\theta(-\sigma_1\hat{p}_1)\otimes I_S\right)\hat{O}^{\text{phys}}_{C_1S|C_2}(\tau_1)=\\&\qquad\Pi_{\sigma_{C_1SC_2}}\mathcal{G}_{C_1S}\left(|\tau_1,\sigma_1\rangle\langle\tau_1,\sigma_1|\otimes\hat{f}^{\text{phys}}_{S|C_1}\right)\Pi_{\sigma_{C_1SC_2}}.\end{aligned}$$

In particular, the transformed system observable is perspective independent, i.e., its description relative to C_1 and C_2 coincide if and only if it is a constant of motion (see [7] for the proof of this statement, which again applies here per pair of σ_1- and σ_2-sector):

Corollary 4. *An operator of* C_2S *relative to* C_1

$$\hat{O}^{\text{phys}}_{C_2S|C_1}=I_{C_2}\otimes\hat{f}^{\text{phys}}_{S|C_1}.$$

transforms under a temporal frame change map to the perspective of C_2 *as follows*

$$\hat{O}^{\text{phys}}_{C_1S|C_2}=I_{C_1}\otimes\hat{f}^{\text{phys}}_{S|C_2},$$

where $\hat{f}^{\text{phys}}_{S|C_1}=\hat{f}^{\text{phys}}_{S|C_2}$ *if and only if* $\hat{f}^{\text{phys}}_{S|C_1}$ *is a constant of motion,* $[\hat{f}^{\text{phys}}_{S|C_1},\hat{H}_S]=0$*.*

Theorem 5 translates as follows into the relational Heisenberg picture (see [7] for the proof which applies here per pair of σ_1- and σ_2-sector):

Corollary 5. *Let* $\hat{O}^{\text{phys}}_{C_2S|C_1}(\tau_1)\in\mathcal{L}(\mathcal{H}^{\text{phys}}_{C_2}\otimes\mathcal{H}^{\text{phys}}_S)$ *be an operator describing the dynamics of properties of the composite system* C_2S *relative to* C_1 *in the Heisenberg picture. Under a temporal frame change Equation (75) to the perspective of* C_2*, this operator transforms to an operator* $\hat{O}^{\boldsymbol{H}}_{C_1S|C_2}(\tau_1,\tau_2)$ *that satisfies the Heisenberg equation of motion in clock* C_2 *time* τ_2 *without an explicitly* τ_2 *dependent term,*

$$\frac{d}{d\tau_2}\hat{O}^{\boldsymbol{H}}_{C_1S|C_2}(\tau_1,\tau_2)=i\left[s_2\frac{\hat{p}_2^2}{2}+\hat{H}_S,\hat{O}^{\boldsymbol{H}}_{C_1S|C_2}(\tau_1,\tau_2)\right],$$

if and only if

$$\hat{O}^{\text{phys}}_{C_2S|C_1}(\tau_1)=\sum_i\left(\hat{O}^{\text{phys}}_{C_2|C_1}\right)_i\otimes\left(\hat{f}^{\text{phys}}_{S|C_1}(\tau_1)\right)_i,$$

and $\left(\hat{O}^{\text{phys}}_{C_2|C_1}\right)_i$ *is a constant of motion,* $[s_1\frac{\hat{p}_1^2}{2},\hat{O}^{\text{phys}}_{C_2|C_1}]=0$*.*

The interpretation of the transformations is of course completely analogous to the relational Schrödinger picture.

7.4. Application: Comparing Clock Readings

One application of the temporal frame change method developed above is comparing readings of different clocks. This is also a prerequisite for developing a notion of clock synchronization.

For example, we may wish to compare the evolution of some system property $\hat{f}_S$ relative to clock C_1 with $\hat{f}_S$ relative to clock C_2. These two relational evolutions will be encoded in two one-parameter families of Dirac observable of the form $\hat{F}_{I_{C_2}\otimes f_S,T_1}(\tau_1)$ and $\hat{F}_{I_{C_1}\otimes f_S,T_2}(\tau_2)$. In order to relate these two dynamics, we need a consistent method for relating the different clock readings τ_1,τ_2. While classically, there is an unambiguous way to answer the question "what is the value of the reading τ_2 of clock C_2, when clock C_1 reads τ_1?" namely by setting $\tau_2(\tau_1):=F_{T_2,T_1}(\tau_1)$, this is not so in the quantum theory because both clocks are now described in terms of quantum operators and their relation depends on the quantum state. In fact, we shall argue shortly that comparing clock readings is generally dependent on the choice of temporal frame (here either C_1 or C_2) in the quantum theory.

7.4.1. Three Ways of Comparing Clock Readings

To address this conundrum in the quantum theory, let us recall the conditional probabilities in Equation (37) and ask for the probability that C_2 reads τ_2 when C_1 reads τ_1 (ignoring frequency sectors for simplicity for the moment):

$$\begin{aligned}&P(T_2=\tau_2|T_1=\tau_1)\\&\quad=\langle\psi_{\text{phys}}|\,e_{C_1}(\tau_1)\otimes e_{C_2}(\tau_2)\otimes I_S\,|\psi_{\text{phys}}\rangle_{\text{kin}}\\&\quad=P(T_1=\tau_1|T_2=\tau_2).\end{aligned}\tag{77}$$

Here we have assumed that the physical state is normalized such that by Corollary 2 also the reduced states in the Schrödinger picture of either clock are normalized.

Comparing clock readings. *Given the conditional probabilities Equation (77), we may consider the following three generally distinct options for comparing clock readings.*

(A) *The clock reading of C_2 when C_1 reads τ_1 is defined to be the value of τ_2 that maximizes the conditional probability $P(T_2 = \tau_2|T_1 = \tau_1)$. This assumes the distribution to have a unique maximal peak.*

(B) *The clock reading of C_2 when C_1 reads τ_1 is defined to be the expectation value*

$$\tau_2(\tau_1) := \int_{\mathbb{R}} d\tau'\, \tau'\, P(T_2 = \tau'|T_1 = \tau_1)\,. \tag{78}$$

(C) *The clock reading of C_2 when C_1 reads τ_1 is defined to be $\left(\tau_2^{(n)}(\tau_1)\right)^{1/n}$ for $n > 1$, where*

$$\tau_2^{(n)}(\tau_1) := \int_{\mathbb{R}} d\tau'\, (\tau')^n\, P(T_2 = \tau'|T_1 = \tau_1) \tag{79}$$

is the nth-moment of the conditional probability distribution in Equation (77).

Relating different clock readings in terms of expectation values, as in (B), is arguably the most natural choice and has originally been discussed in Höhn and Vanrietvelde [30], Höhn [31], Bojowald et al. [32, 33], Höhn et al. [34], and Smith and Ahmadi [45]; we expand on this here.

Clearly, the two definitions (A) and (B) only agree when the conditional probability distribution is peaked on the expectation value. Furthermore, all three definitions (A)–(C) agree in the special case that $P(T_2 = \tau'|T_1 = \tau_1) = \delta(\tau' - \tau_1)$, i.e., when there are no fluctuations in the conditional probability distribution.

7.4.2. Comparing Clock Readings for Quadratic Clock Hamiltonians

Let us now explore these definitions in our present class of models defined by Equation (66), taking into account the different frequency sectors again. Minding the double superselection rule, we replace Equation (77) by

$$\begin{aligned} &P_{\sigma_1,\sigma_2}(T_2 = \tau_2|T_1 = \tau_1) \\ &\quad = \langle\psi_{\text{phys}}|\, e^{\sigma_1}_{C_1}(\tau_1) \otimes e^{\sigma_2}_{C_2}(\tau_2) \otimes I_S\, |\psi_{\text{phys}}\rangle_{\text{kin}} \\ &\quad = \langle\psi_{\sigma_1,\sigma_2}|\, e_{C_1}(\tau_1) \otimes e_{C_2}(\tau_2) \otimes I_S\, |\psi_{\sigma_1,\sigma_2}\rangle_{\text{kin}}\,, \end{aligned} \tag{80}$$

where $|\psi_{\sigma_1,\sigma_2}\rangle \in \mathcal{H}_{\sigma_1,\sigma_2}$ lies in the overlap of the σ_1- and σ_2-sectors (see Equation 67) and $e^{\sigma_i}_{C_i}(\tau_i) := \frac{1}{2\pi}\,|\tau_i, \sigma_i\rangle\langle\tau_i, \sigma_i|$, $i = 1, 2$. We can then write the n^{th}-moment of the conditional probability distribution in Equations (78) and (79) for $n \in \mathbb{N}$, thus considering both definitions (B) and (C), as follows:

$$\begin{aligned} \tau_2^{(n)}(\tau_1) &= \int_{\mathbb{R}} d\tau'\, (\tau')^n\, P_{\sigma_1,\sigma_2}(T_2 = \tau'|T_1 = \tau_1) \\ &= \langle\psi^{\sigma_1}_{C_2S|C_1}(\tau_1)\,|\,\hat{T}^{(n)}_{2,\sigma_2} \otimes I_S\,|\psi^{\sigma_1}_{C_2S|C_1}(\tau_1)\rangle \\ &= \langle\psi_{\sigma_1,\sigma_2}\,|\,\hat{F}_{T_2^{(n)}\otimes I_S, T_1}(\tau_1)\,)|\,\psi_{\sigma_1,\sigma_2}\rangle_{\text{phys}}\,, \end{aligned} \tag{81}$$

where by Equation (30)

$$\begin{aligned} \hat{T}^{(n)}_{2,\sigma_2} &= \frac{1}{2\pi}\int_{\mathbb{R}} dt\, t^n\, |t, \sigma_2\rangle\langle t, \sigma_2| \\ &= \theta(-\sigma_2\hat{p}_2)\,\hat{T}^{(n)}_2\,\theta(-\sigma_2\hat{p}_2) \end{aligned}$$

is the σ_2-sector nth-moment of the covariant clock POVM corresponding to C_2. In the second line of Equation (81) we have made use of Equations (38) and (80), while in the third line we invoked Theorem 2. Note that by Equation (74), the expression in Equation (81) defines an expectation value which is preserved during a temporal frame change between C_1 and C_2.

Thanks to Lemmas 3 and 4 we can write the nth-moment in Equation (81) also in the form

$$\begin{aligned} \tau_2^{(n)}(\tau_1) &= \langle\psi^{\sigma_1}_{C_2S|C_1}(\tau_1)\,|\,\hat{T}^n_{2,\sigma_2} \otimes I_S\,|\psi^{\sigma_1}_{C_2S|C_1}(\tau_1)\rangle \\ &= \langle\psi_{\sigma_1,\sigma_2}\,|\,\hat{F}_{T_2^n\otimes I_S, T_1}(\tau_1)\,)|\,\psi_{\sigma_1,\sigma_2}\rangle_{\text{phys}}\,, \end{aligned}$$

as long as $|\psi^{\sigma_1}_{C_2S|C_1}(\tau_1)\rangle \in \left(\mathcal{D}(\hat{T}^n_2) \cap \mathcal{H}^{\text{phys}}_{C_2}\right) \otimes \mathcal{H}^{\text{phys}}_S$. Since $\left(\tau_2^{(n)}(\tau_1)\right)^{1/n} \neq \tau_2(\tau_1)$ for $n > 1$ for general states, definitions (B) and (C) will generically not be equivalent. In the sequel, we shall mostly consider definition (B) in extension of Höhn and Vanrietvelde [30], Höhn [31], Bojowald et al. [32, 33], Höhn et al. [34], and Smith and Ahmadi [45]. This seems to be the physically most appealing one, especially if an ensemble interpretation could be developed for the models under consideration. Definition (A) is only unambiguous when the conditional probability distribution has a single maximal peak and definition (C) is operationally unnatural and convoluted. That is, we set for the value of the reading of clock C_2 when C_1 reads τ_1:

$$\tau_2(\tau_1) := \tau_2^{(1)}(\tau_1)\,. \tag{82}$$

The following discussion, however, qualitatively also applies to definition (C).

7.4.3. Comparing Clock Readings Is Temporal Frame Dependent

Notice that definitions (A)–(C) treat C_2 as the fluctuating subsystem. We can thus interpret them as providing a definition of the clock reading of C_2 relative to the temporal reference frame C_1. Conversely, we can of course switch the roles of C_1 and C_2 above and ask for the clock reading of C_1 relative to C_2. Resorting to definition (B), this would yield

$$\tau_1(\tau_2) = \int_{\mathbb{R}} d\tau'\, \tau'\, P_{\sigma_1,\sigma_2}(T_2 = \tau_2|T_1 = \tau')\,. \tag{83}$$

Dropping the labels of the arguments in Equations (81) and (83), both of which run over all of $\mathbb{R}$, it is important to note that $\tau_1(\tau)$ and $\tau_2(\tau)$ will generally *not* be the same functions of τ. This is because generally $P_{\sigma_1,\sigma_2}(T_2 = \tau'|T_1 = \tau) \neq P_{\sigma_1,\sigma_2}(T_2 = \tau|T_1 = \tau')$ in Equation (80). Said another way, the evolution of C_2 from the perspective of C_1 according to definition (B) may differ from the evolution of C_1 relative to C_2 (for the same physical state).

One might wonder whether the function $\tau_1(\tau_2)$ in Equation (83) is the inversion of $\tau_2(\tau_1)$ in Equation (81), i.e., obtained by solving $\tau_2(\tau_1)$ for τ_1. Classically, this is certainly the case and it would entail that for a fixed clock reading τ_1^* of C_1 one finds $\tau_1(\tau_2(\tau_1^*)) = \tau_1^*$. Physically this would mean

that both temporal reference frames C_1 and C_2 agree that when C_1 reads τ_1^*, C_2 reads the value $\tau_2(\tau_1^*)$. This does occur in a special case when definitions (A)–(C) all coincide, namely when $P_{\sigma_1,\sigma_2}(T_2 = \tau_2(\tau_1^*)|T_1 = \tau') = \delta(\tau' - \tau_1^*)$ in Equation (83) in which case expectation value, most probable value and the value defined through the nth-moment all agree. While this does happen in simple models with a high degree of symmetry between C_1 and C_2 [31], this will in more interesting cases not be the case because the physical state will generically have a different spread along the τ_1 and τ_2 axes [30, 32–34]. In our case this means that the wave function

$$\psi^{\sigma_1,\sigma_2}_{C_2S|C_1}(\tau_1,\tau_2) := (\langle\tau_1,\sigma_1| \otimes \langle\tau_2,\sigma_2| \otimes \langle\phi_S|)\ |\psi_{\text{phys}}\rangle\ ,$$

for some physical system state $|\phi_S\rangle \in \mathcal{H}_S^{\text{phys}}$, which can be viewed as either a wave function in the C_1 or C_2 relational Schrödinger picture, may have a different spread in τ_1 than in τ_2. In such a case we will generally find $\tau_1(\tau_2(\tau_1^*)) \neq \tau_1^*$. This effect will occur in the class of models considered here because physical states need not have the same momentum distribution in p_1 and p_2 (and thus neither in τ_1 or τ_2) due to the presence of the system S. This effect has also been demonstrated in a semiclassical approach in various models in Bojowald et al. [32, 33] and Höhn et al. [34] where one finds discrepancies of the order of $\hbar$ between τ_1^* and $\tau_1(\tau_2^* = \tau_2(\tau_1^*))$.

In conclusion, this effect can be interpreted as a temporal frame dependence of comparing clock readings according to definition (B) [or (C)]: if from the perspective of the temporal reference frame defined by C_1 the clock C_2 reads $\tau_2(\tau_1^*)$ (computed according to Equation 81) when C_1 reads τ_1^*, then conversely from the perspective of the temporal reference frame defined by C_2 the clock C_1 will not in general read τ_1^* when C_2 reads the value $\tau_2(\tau_1^*)$. That is, C_1 and C_2 will generally disagree about the pairings of their clock readings.

Let us now also briefly comment on the notion of quantum clock synchronization. Using the state dependent relation Equation (82), we could ask for which state would yield $\tau_2(\tau_1^*) = \tau_1^*$ so that C_1 and C_2 read the *same* value when C_1 reads the value τ_1^*. Even stronger, we could ask whether there are states for which $\tau_2(\tau_1) = \tau_1 + const$, for all $\tau_1 \in \mathbb{R}$, so that, up to a constant offset, C_1 and C_2 are always synchronized. Equation (81) tells us that this is the case if $P_{\sigma_1,\sigma_2}(T_2 = \tau'|T_1 = \tau_1) = \delta(\tau' - \tau_1 - const)$. Again, while this happens in simple models [31], this will generically not happen for the models in the class which we are studying on account of the above observations concerning the frame dependence of comparing clock readings. Such a notion of synchronization is therefore too strong and can generally not be implemented. It will furthermore generally be frame dependent too.

7.4.4. Comparing a System's Evolution Relative to Two Clocks

Returning to our original ambition, it is thus more useful to employ the more general (frame dependent) clock comparison, according to definition (B), in order to compare the evolutions of S with respect to C_1 and C_2. Working in the relational Schrödinger picture, if $|\psi^{\sigma_1}_{C_2S|C_1}(\tau_1^*)\rangle$ is the initial state of C_2S from the perspective of C_1, then according to Equation (69) the corresponding initial state of C_1S from the perspective of C_2 is

$$\begin{aligned}&\left(\theta(-\sigma_1\hat{p}_1)\otimes I_S\right)|\psi^{\sigma_2}_{C_1S|C_2}(\tau_2(\tau_1^*))\rangle\\ &\quad = \mathcal{R}^{\sigma_2}_{\mathbf{S}}(\tau_2(\tau_1^*))\circ\left(\mathcal{R}^{\sigma_1}_{\mathbf{S}}(\tau_1^*)\right)^{-1}|\psi^{\sigma_1}_{C_2S|C_1}(\tau_1^*)\rangle\ .\end{aligned}\tag{84}$$

We can then evaluate the 'same' reduced system observable $I_{C_i}\otimes\hat{f}_S^{\text{phys}}$ in the two states, where $i = 1$ when evaluated relative to C_2 and vice versa (cf. Corollary 4), in order to compare the evolution of property $\hat{f}_S^{\text{phys}}$ relative to the two clocks in different quantum states (which amount also to quantum states of the clocks). To avoid confusion, we emphasize, that $I_{C_i}\otimes\hat{f}_S^{\text{phys}}$, $i = 1,2$, correspond to two different relational Dirac observables $\hat{F}_{I_{C_2}\otimes f_S,T_1}(\tau_1)$ and $\hat{F}_{I_{C_1}\otimes f_S,T_2}(\tau_2)$ on the clock-neutral physical Hilbert space $\mathcal{H}_{\text{phys}}$; in particular, the two are *not* related by the TFC map $\Lambda_{\mathbf{S}}^{\sigma_1\to\sigma_2}$. Hence, by evaluating these two reduced observables in the relational Schrödinger states related via the TFC map $\Lambda_{\mathbf{S}}^{\sigma_1\to\sigma_2}$ by Equation (84), we can compare two genuinely distinct relational dynamics. The construction in the relational Heisenberg picture is of course completely analogous.

In Höhn et al. [7] and Castro-Ruiz [47] a frame dependent temporal non-locality effect was exhibited for idealized clocks whose Hamiltonian is the unbounded momentum operator. For example, when clock C_2 is seen to be in a superposition of two peaked states and in a product relation with S from the perspective of C_1, then C_1S will generally be entangled as seen from the perspective of C_2 and undergo a superposition of time evolutions. This effect applies here per overlap of the different σ_1- and σ_2-sectors. It will be interesting to study how such a frame dependent temporal locality affects the (potentially frame dependent) comparison and synchronization of the clocks and the comparison of the evolutions of S relative to C_1 and C_2 in different quantum states, corresponding to different choices of the clock-neutral physical states. Such an exploration will appear elsewhere.

Finally, these temporal frame changes and clock synchronizations will be relevant in quantum cosmology. For example, recently it was pointed out that singularity resolution in quantum cosmology depends on the choice of clock which one uses to define a relational dynamics [77]. The different relational dynamics employed in Gielen and Menéndez-Pidal [77] can be interpreted as different choices of reduced dynamics in the sense of our relational Schrödinger/Heisenberg picture. Temporal frame changes as developed here can in principle be used to study the temporal frame dependence of the fate of cosmological singularities more systematically.

8. CONCLUSIONS

In this work we demonstrated the equivalence of three distinct approaches to relational quantum dynamics—relational Dirac observables, the Page-Wootters formalism, and quantum deparametrizations—for models described by a Hamiltonian constraint in which the momentum of the system being employed as a clock appears quadratically. Since this class of models encompasses many relativistic settings, we have thereby extended

our previous results of Höhn et al. [7] into a relativistic context. A crucial ingredient in this extension has been a clock POVM which is covariant with respect to the group generated by the Hamiltonian constraint and is used to describe the temporal reference frame defined by the clock. This choice differs from the usual resort to self-adjoint clock operators in relativistic settings.

Owing to a superselection rule induced by the shape of the Hamiltonian constraint across positive and negative frequency modes, this equivalence, which we refer to as the trinity of relational quantum dynamics, holds frequency sector wise. Moreover, we further develop the method of temporal quantum frame changes [7, 30–34, 47, 69] in this setting to address the multiple choice problem. This method is then used to explore an indirect self-reference phenomenon that arises when transforming between clock perspectives and to reveal the temporal frame and state dependence of comparing or even synchronizing the readings of different quantum clocks. This result adds to the growing list of quantum reference frame dependent physical properties, such as entanglement [70, 72, 74], spin [73], classicality [72] or objectivity [79, 80] of a subsystem, superpositions [70, 72], certain quantum resources [78], measurements [70, 76], causal relations [47, 83], temporal locality [7, 47], and even spacetime singularity resolution [77]. The temporal frame changes may also be employed to extend recent proposals for studying time dilation effects of quantum clocks [45, 136, 137] (see also [137–140]). Furthermore, it would be interesting to expand the temporal frame changes to cosmological perturbation theory to study the temporal frame dependence of power spectra [64, 65].

Importantly, the covariant clock POVM permitted us to resolve Kuchař's criticism that the Page-Wootters formalism does not produce the correct localization probability for a relativistic particle in Minkowski space [1]. Indeed, such incorrect localization probabilities arise when conditioning on times defined by the quantization of an inertial Minkwoski time coordinate. We showed that conditioning instead on the covariant clock POVM surprisingly produces a Newton-Wigner type localization probability, which, while approximate and not fully covariant, is usually regarded as the best possible notion of localization in relativistic quantum mechanics [90, 94]. This result underscores the benefits of covariant clock POVMs in defining a consistent relational quantum dynamics [7, 44, 45, 96, 97].

In conjunction with our previous article [7], we have thus resolved all three criticisms (a)–(c) (see Introduction) that Kuchař raised against the Page-Wootters formalism in Kuchař [1]. The Page-Wootters formalism is therefore a viable approach to relational quantum dynamics. Through the equivalence established by the trinity, it also equips the relational observable formulation and deparametrizations with a consistent conditional probability interpretation. In particular, relational observables describing the evolution of a position operator relative to a covariant clock POVM yield a Newton-Wigner type localization in relativistic settings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors have contributed extensively to the research and writing of the article with the lead input from PH.

FUNDING

PH was grateful for support from the Simons Foundation through an It-from-Qubit Fellowship and the Foundational Questions Institute through Grant number FQXi-RFP-1801A. AS acknowledges support from the Natural Sciences and Engineering Research Council of Canada and the Dartmouth Society of Fellows. ML acknowledges support from the ESQ Discovery Grant of the Austrian Academy of Sciences (ÖAW), as well as from the Austrian Science Fund (FWF) through the START project Y879-N27. This work was supported in part by funding from Okinawa Institute of Science and Technology Graduate University. The initial stages of this project were made possible through the support of a grant from the John Templeton Foundation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2021.587083/full#supplementary-material

REFERENCES

1. Kuchař KV. Time and interpretations of quantum gravity. *Int J Mod Phys D*. (2011) 20:3–86. doi: 10.1142/S02182718110 19347
2. Isham CJ. Canonical quantum gravity and the problem of time. In: Ibort LA, Rodríguez MA, editors. *Integrable Systems, Quantum Groups, and Quantum Field Theories*. Dordrecht: Springer Netherlands (1993). p. 157–287. doi: 10.1007/978-94-011-1980-1_6
3. Anderson E. *The Problem of Time. Vol. 190 of Fundamental Theories of Physics*. Basel: Springer International Publishing (2017).
4. Rovelli C. *Quantum Gravity*. Cambridge: Cambridge University Press (2004).
5. Smolin L. *The Case for Background Independence*. Oxford: Oxford University Press (2006).
6. Smolin L. Temporal relationalism. *arXiv*. (2018) 180512468.
7. Höhn PA, Smith ARH, Lock MPE. The trinity of relational quantum dynamics. *arXiv*. (2019) 191200033.
8. DeWitt BS. Quantum theory of gravity. I. The canonical theory. *Phys Rev*. (1967) 160:1113–48. doi: 10.1103/PhysRev.160.1113
9. Rovelli C. Is there incompatibility between the ways time is treated in general relativity and in standard quantum mechanics. In: Ashtekar A, Stachel J,

editors. *Conceptual Problems of Quantum Gravity*. Basel: Birkhauser (1991). p. 126–40.
10. Ashtekar A. *Lectures on Non-Perturbative Canonical Gravity. Vol. 6 of Physics and Cosmology*. Singapore: World Scientific (1991).
11. Thiemann T. *Modern Canonical Quantum General Relativity*. Cambridge: Cambridge University Press (2008).
12. Rovelli C. Quantum mechanics without time: a model. *Phys Rev D*. (1990) 42:2638–46. doi: 10.1103/PhysRevD.42.2638
13. Rovelli C. Time in quantum gravity: physics beyond the Schrödinger regime. *Phys Rev D*. (1991) 43:442–56. doi: 10.1103/PhysRevD.43.442
14. Rovelli C. What is observable in classical and quantum gravity? *Class Quantum Grav*. (1991) 8:297–316. doi: 10.1088/0264-9381/8/2/011
15. Rovelli C. Quantum reference systems. *Class Quantum Grav*. (1991) 8:317–32. doi: 10.1088/0264-9381/8/2/012
16. Marolf D. Almost ideal clocks in quantum cosmology: a brief derivation of time. *Class Quantum Grav*. (1995) 12:2469–86. doi: 10.1088/0264-9381/12/10/007
17. Marolf D. Quantum observables and recollapsing dynamics. *Class Quantum Grav*. (1995) 12:1199–220. doi: 10.1088/0264-9381/12/5/011
18. Gambini R, Porto RA. Relational time in generally covariant quantum systems: four models. *Phys Rev D*. (2001) 63:105014. doi: 10.1103/PhysRevD.63.105014
19. Tambornino J. Relational observables in gravity: a review. *SIGMA*. (2012) 8:017. doi: 10.3842/SIGMA.2012.017
20. Giddings SB, Marolf D, Hartle JB. Observables in effective gravity. *Phys Rev D*. (2006) 74:064018. doi: 10.1103/PhysRevD.74.064018
21. Dittrich B. Partial and complete observables for Hamiltonian constrained systems. *Gen Relativ Gravit*. (2007) 39:1891–927. doi: 10.1007/s10714-007-0495-2
22. Dittrich B. Partial and complete observables for canonical general relativity. *Class Quantum Grav*. (2006) 23:6155–84. doi: 10.1088/0264-9381/23/22/006
23. Dittrich B, Tambornino J. A perturbative approach to Dirac observables and their space-time algebra. *Class Quantum Grav*. (2007) 24:757–84. doi: 10.1088/0264-9381/24/4/001
24. Dittrich B, Tambornino J. Gauge invariant perturbations around symmetry reduced sectors of general relativity: applications to cosmology. *Class Quantum Grav*. (2007) 24:4543–86. doi: 10.1088/0264-9381/24/18/001
25. Gambini R, Porto RA, Pullin J, Torterolo S. Conditional probabilities with Dirac observables and the problem of time in quantum gravity. *Phys Rev D*. (2009) 79:041501. doi: 10.1103/PhysRevD.79.041501
26. Pons JM, Salisbury DC, Sundermeyer KA. Revisiting observables in generally covariant theories in the light of gauge fixing methods. *Phys Rev D*. (2009) 80:084015. doi: 10.1103/PhysRevD.80.084015
27. Ashtekar A, Pawlowski T, Singh P. Quantum nature of the big bang: an analytical and numerical investigation. I. *Phys Rev D*. (2006) 73:124038. doi: 10.1103/PhysRevD.73.124038
28. Kaminski W, Lewandowski J, Pawlowski T. Physical time and other conceptual issues of QG on the example of LQC. *Class Quantum Grav*. (2009) 26:035012. doi: 10.1088/0264-9381/26/3/035012
29. Kaminski W, Lewandowski J, Pawlowski T. Quantum constraints, Dirac observables and evolution: group averaging versus Schrodinger picture in LQC. *Class Quantum Grav*. (2009) 26:245016. doi: 10.1088/0264-9381/26/24/245016
30. Höhn PA, Vanrietvelde A. How to switch between relational quantum clocks. *New J Phys*. (2020) 22:123048. doi: 10.1088/1367-2630/abd1ac
31. Höhn PA. Switching internal times and a new perspective on the 'wave function of the universe'. *Universe*. (2019) 5:116. doi: 10.3390/universe5050116
32. Bojowald M, Höhn PA, Tsobanjan A. An effective approach to the problem of time. *Class Quantum Grav*. (2011) 28:035006. doi: 10.1088/0264-9381/28/3/035006
33. Bojowald M, Höhn PA, Tsobanjan A. Effective approach to the problem of time: general features and examples. *Phys Rev D*. (2011) 83:125023. doi: 10.1103/PhysRevD.83.125023
34. Höhn PA, Kubalova E, Tsobanjan A. Effective relational dynamics of a nonintegrable cosmological model. *Phys Rev D*. (2012) 86:065014. doi: 10.1103/PhysRevD.86.065014
35. Dittrich B, Höhn PA, Koslowski TA, Nelson MI. Can chaos be observed in quantum gravity? *Phys Lett B*. (2017) 769:554–60. doi: 10.1016/j.physletb.2017.02.038
36. Dittrich B, Höhn PA, Koslowski TA, Nelson MI. Chaos, Dirac observables and constraint quantization. *arXiv*. (2015) 150801947.
37. Chataignier L. On the construction of quantum Dirac observables and the emergence of WKB time. *Phys Rev D*. (2020) 101:086001. doi: 10.1103/PhysRevD.101.086001
38. Chataignier L. Relational observables, reference frames and conditional probabilities. *arXiv*. (2020) 200605526. doi: 10.1103/PhysRevD.103.026013
39. Page DN, Wootters WK. Evolution without evolution: dynamics described by stationary observables. *Phys Rev D*. (1983) 27:2885. doi: 10.1103/PhysRevD.27.2885
40. Wootters WK. "Time" replaced by quantum correlations. *Int J Theor Phys*. (1984) 23:701. doi: 10.1007/BF02214098
41. Gambini R, Pullin J. Relational physics with real rods and clocks and the measurement problem of quantum mechanics. *Found Phys*. (2007) 37:1074–92. doi: 10.1007/s10701-007-9144-6
42. Gambini R, Porto R, Pullin J. Fundamental decoherence from quantum gravity: a pedagogical review. *Gen Relativ Gravit*. (2007) 39:1143–56. doi: 10.1007/s10714-007-0451-1
43. Giovannetti V, Lloyd S, Maccone L. Quantum time. *Phys Rev D*. (2015) 79:945933. doi: 10.1103/PhysRevD.92.045033
44. Smith ARH, Ahmadi M. Quantizing time: interacting clocks and systems. *Quantum*. (2019) 3:160. doi: 10.22331/q-2019-07-08-160
45. Smith ARH, Ahmadi M. Quantum clocks observe classical and quantum time dilation. *Nat Commun*. (2020) 11:5360. doi: 10.1038/s41467-020-18264-4
46. Dolby CE. The conditional probability interpretation of the Hamiltonian constraint. *arXiv*. (2004) 0406034.
47. Castro-Ruiz E, Giacomini F, Belenchia A, Brukner Č. Quantum clocks and the temporal localisability of events in the presence of gravitating quantum systems. *Nat Commun*. (2020) 11:2672. doi: 10.1038/s41467-020-16013-1
48. Boette A, Rossignoli R. History states of systems and operators. *Phys Rev A*. (2018) 98:032108. doi: 10.1103/PhysRevA.98.032108
49. Diaz NL, Rossignoli R. History state formalism for Dirac's theory. *Phys Rev D*. (2019) 99:045008. doi: 10.1103/PhysRevD.99.045008
50. Diaz NL, Matera JM, Rossignoli R. History state formalism for scalar particles. *Phys Rev D*. (2019) 100:125020. doi: 10.1103/PhysRevD.100.125020
51. Singh A. Quantum space, quantum time, and relativistic quantum mechanics. *arXiv*. (2020) 200409139.
52. Leon J, Maccone L. The Pauli objection. *Found Phys*. (2017) 47:1597–608. doi: 10.1007/s10701-017-0115-2
53. Marletto C, Vedral V. Evolution without evolution and without ambiguities. *Phys Rev D*. (2017) 95:043510. doi: 10.1103/PhysRevD.95.043510
54. Nikolova A, Brennen G, Osborne TJ, Milburn G, Stace TM. Relational time in anyonic systems. *Phys Rev A*. (2018) 97:030101. doi: 10.1103/PhysRevA.97.030101
55. Baumann V, Santo FD, Smith ARH, Giacomini F, Castro-Ruiz E, Brukner C. Generalized probability rules from a timeless formulation of Wigner's friend scenarios. *arXiv*. (2019) 191109696.
56. Favalli T, Smerzi A. Hermitian time operator in a timeless universe. *arXiv*. (2020) 200309042. doi: 10.22331/q-2020-10-29-354
57. Foti C, Coppo A, Barni G, Cuccoli A, Verrucchi P. There is only one time. *arXiv*. (2020) 200612103.
58. Thiemann T. Reduced phase space quantization and Dirac observables. *Class Quantum Grav*. (2006) 23:1163–80. doi: 10.1088/0264-9381/23/4/006
59. Bojowald M, Tsobanjan A. Quantization of dynamical symplectic reduction. *arXiv*. (2019) 190604792. doi: 10.1007/s00220-020-03856-4
60. Giesel K, Thiemann T. Algebraic quantum gravity (AQG). IV. Reduced phase space quantisation of loop quantum gravity. *Class Quantum Grav*. (2010) 27:175009. doi: 10.1088/0264-9381/27/17/175009
61. Domagala M, Giesel K, Kaminski W, Lewandowski J. Gravity quantized: loop quantum gravity with a scalar field. *Phys Rev D*. (2010) 82:104038. doi: 10.1103/PhysRevD.82.104038
62. Husain V, Pawlowski T. Time and a physical Hamiltonian for quantum gravity. *Phys Rev Lett*. (2012) 108:141301. doi: 10.1103/PhysRevLett.108.141301

63. Giesel K, Vetter A. Reduced loop quantization with four Klein-Gordon scalar fields as reference matter. *Class Quantum Grav*. (2019) 36:145002. doi: 10.1088/1361-6382/ab26f4
64. Giesel K, Herold L, Li BF, Singh P. Mukhanov-Sasaki equation in manifestly gauge-invariant linearized cosmological perturbation theory with dust reference fields. *Phys Rev D*. (2020) 102:023524. doi: 10.1103/PhysRevD.102.023524
65. Giesel K, Li BF, Singh P. Towards a reduced phase space quantization in loop quantum cosmology with an inflationary potential. *arXiv*. (2020) 200706597. doi: 10.1103/PhysRevD.102.126024
66. Albrecht A, Iglesias A. The clock ambiguity and the emergence of physical laws. *Phys Rev D*. (2008) 77:063506. doi: 10.1103/PhysRevD.77.063506
67. Albrecht A, Iglesias A. The clock ambiguity: implications and new developments. In: Mersini-Houghton L, Vaas R, editors. *The Arrows of Time. Fundamental Theories of Physics*, Vol. 172. Berlin; Heidelberg: Springer (2012). doi: 10.1007/978-3-642-23259-6_4
68. Corbin V, Cornish NJ. Semi-classical limit and minimum decoherence in the conditional probability interpretation of quantum mechanics. *Found Phys*. (2009) 39:474. doi: 10.1007/s10701-009-9298-5
69. Bojowald M, Halnon T. Time in quantum cosmology. *Phys Rev D*. (2018) 98:066001. doi: 10.1103/PhysRevD.98.066001
70. Giacomini F, Castro-Ruiz E, Brukner Č. Quantum mechanics and the covariance of physical laws in quantum reference frames. *Nat Commun*. (2019) 10:494. doi: 10.1038/s41467-018-08155-0
71. Vanrietvelde A, Höhn PA, Giacomini F, Castro-Ruiz E. A change of perspective: switching quantum reference frames via a perspective-neutral framework. *Quantum*. (2020) 4:225. doi: 10.22331/q-2020-01-27-225
72. Vanrietvelde A, Höhn PA, Giacomini F. Switching quantum reference frames in the N-body problem and the absence of global relational perspectives. *arXiv*. (2018) 180905093.
73. Giacomini F, Castro-Ruiz E, Brukner Č. Relativistic quantum reference frames: the operational meaning of spin. *Phys Rev Lett*. (2019) 123:090404. doi: 10.1103/PhysRevLett.123.090404
74. de la Hamette AC, Galley TD. Quantum reference frames for general symmetry groups. *Quantum*. (2020) 4:367. doi: 10.22331/q-2020-11-30-367
75. Loveridge L, Miyadera T, Busch P. Symmetry, reference frames, and relational quantities in quantum mechanics. *Found Phys*. (2018) 48:135–98. doi: 10.1007/s10701-018-0138-3
76. Yang JM. Switching quantum reference frames for quantum measurement. *Quantum*. (2020) 4:283. doi: 10.22331/q-2020-06-18-283
77. Gielen S, Menéndez-Pidal L. Singularity resolution depends on the clock. *arXiv*. (2020) 200505357. doi: 10.1088/1361-6382/abb14f
78. Savi MF, Angelo RM. Quantum resources covariance. *arXiv*. (2020) 200509612. doi: 10.1103/PhysRevA.103.022220
79. Le TP, Mironowicz P, Horodecki P. Blurred quantum Darwinism across quantum reference frames. *arXiv*. (2020) 200606364. doi: 10.1103/PhysRevA.102.062420
80. Tuziemski J. Decoherence and information encoding in quantum reference frames. *arXiv*. (2020) 200607298.
81. Hardy L. The construction interpretation: conceptual roads to quantum gravity. *arXiv*. (2018) 180710980.
82. Hardy L. Implementation of the quantum equivalence principle. In: Finster F, Giulini D, Kleiner J, Tolksdorf J, editors. *Progress and Visions in Quantum Theory in View of Gravity: Bridging Foundations of Physics and Mathematics*. Basel: Birkhäuser (2020). p. 189–220. doi: 10.1007/978-3-030-38941-3_8
83. Guérin PA, Brukner Č. Observer-dependent locality of quantum events. *New J Phys*. (2018) 20:103031. doi: 10.1088/1367-2630/aae742
84. Zych M, Costa F, Ralph TC. Relativity of quantum superpositions. *arXiv*. (2018) 180904999.
85. Barbado LC, Castro-Ruiz E, Apadula L, Brukner Č. Unruh effect for detectors in superposition of accelerations. *arXiv*. (2020) 200312603. doi: 10.1103/PhysRevD.102.045002
86. Holevo AS. *Probabilistic and Statistical Aspects of Quantum Theory. Vol. 1 of Statistics and Probability*. Amsterdam: North-Holland Publishing Company (1982).
87. Busch P, Grabowski M, Lahti PJ. *Operational Quantum Physics. Vol. 31 of Lecture Notes in Physics Monographs*. Berlin; Heidelberg: Springer-Verlag (1995).
88. Busch P, Grabowski M, Lahti PJ. Time observables in quantum theory. *Phys Lett A*. (1994) 191:357–61. doi: 10.1016/0375-9601(94)90785-4
89. Braunstein SL, Caves CM, Milburn GJ. Generalized uncertainty relations: theory, examples, and Lorentz invariance. *Ann Phys*. (1996) 247:135–73. doi: 10.1006/aphy.1996.0040
90. Haag R. *Local Quantum Physics: Fields, Particles, Algebras*. Berlin; Heidelberg: Springer-Verlag (2012).
91. Newton TD, Wigner EP. Localized states for elementary systems. *Rev Mod Phys*. (1949) 21:400–6. doi: 10.1103/RevModPhys.21.400
92. Fernando Perez J, Wilde IF. Localization and causality in relativistic quantum mechanics. *Phys Rev D*. (1977) 16:315. doi: 10.1103/PhysRevD.16.315
93. Malament DB. In defense of dogma: why there cannot be a relativistic quantum mechanics of (Localizable) particles. In: Clifton R, editor. *Perspectives on Quantum Reality. The University of Western Ontario Series in Philosophy of Science (A Series of Books in Philosophy of Science, Methodology, Epistemology, Logic, History of Science, and Related Fields)*, Vol. 57. Dordrecht: Springer (1996). p. 1–10. doi: 10.1007/978-94-015-8656-6_1
94. Yngvason J. Localization and entanglement in relativistic quantum physics. *Lect Notes Phys*. (2015) 899:325–48. doi: 10.1007/978-3-662-46422-9_15
95. Papageorgiou M, Pye J. Impact of relativity on particle localizability and ground state entanglement. *J Phys A*. (2019) 52:375304. doi: 10.1088/1751-8121/ab3593
96. Brunetti R, Fredenhagen K, Hoge M. Time in quantum physics: from an external parameter to an intrinsic observable. *Found Phys*. (2010) 40:1368–78. doi: 10.1007/s10701-009-9400-z
97. Loveridge L, Miyadera T. Relative quantum time. *Found Phys*. (2019) 49:549–60. doi: 10.1007/s10701-019-00268-w
98. Höhn PA. Reflections on the information paradigm in quantum and gravitational physics. *J Phys Conf Ser*. (2017) 880:012014. doi: 10.1088/1742-6596/880/1/012014
99. Dalla Chiara ML. Logical self reference, set theoretical paradoxes and the measurement problem in quantum mechanics. *J Philos Logic*. (1977) 6:331–47. doi: 10.1007/BF00262066
100. Breuer T. The impossibility of accurate state self-measurements. *Philos Sci*. (1995) 62:197–214. doi: 10.1086/289852
101. Dirac PAM. *Lectures on Quantum Mechanics*. New York, NY: Belfer Graduate School of Sciencem Yeshiva University (1964).
102. Henneaux M, Teitelboim C. *Quantization of Gauge Systems*. Princeton, NJ: Princeton University Press (1992).
103. Marolf D. Refined algebraic quantization: systems with a single constraint. *arXiv*. (1995) 9508015.
104. Hartle JB, Marolf D. Comparing formulations of generalized quantum mechanics for reparametrization–invariant systems. *Phys Rev D*. (1997) 56:6247–57. doi: 10.1103/PhysRevD.56.6247
105. Giulini D, Marolf D. On the generality of refined algebraic quantization. *Class Quantum Grav*. (1999) 16:2479–88. doi: 10.1088/0264-9381/16/7/321
106. Giulini D, Marolf D. A Uniqueness theorem for constraint quantization. *Class Quantum Grav*. (1999) 16:2489–505. doi: 10.1088/0264-9381/16/7/322
107. Marolf D. Group averaging and refined algebraic quantization: where are we now? *arXiv*. (2000) 0011112.
108. Pailas T. "Time"-covariant Schrödinger equation and the canonical quantization of the Reissner-Nordström black hole. *arXiv*. (2020). doi: 10.3390/quantum2030029
109. Blyth WF, Isham CJ. Quantization of a Friedmann universe filled with a scalar field. *Phys Rev D*. (1975) 11:768–78. doi: 10.1103/PhysRevD.11.768
110. Hawking SW. Quantum cosmology. In: DeWitt BS, Stora R, editors. *Relativity, Groups and Topology II, Les Houches Summer School, 1983*. Amsterdam: North-Holland Publishing Company (1984). p. 333.
111. Hájíček P. Origin of nonunitarity in quantum gravity. *Phys Rev D*. (1986) 34:1040. doi: 10.1103/PhysRevD.34.1040
112. Kiefer C. Wave packets in minisuperspace. *Phys Rev D*. (1988) 38:1761. doi: 10.1103/PhysRevD.38.1761
113. Ashtekar A, Singh P. Loop quantum cosmology: a status report. *Class Quantum Grav*. (2011) 28:213001. doi: 10.1088/0264-9381/28/21/213001
114. Ashtekar A, Corichi A, Singh P. Robustness of key features of loop quantum cosmology. *Phys Rev D*. (2008) 77:024046. doi: 10.1103/PhysRevD.77.024046
115. Bojowald M. *Canonical Gravity and Applications: Cosmology, Black Holes and Quantum Gravity*. Cambridge: Cambridge University Press (2010).

116. Ashtekar A, Tate R, Uggla C. Minisuperspaces: observables and quantization. *Int J Mod Phys D*. (1993) 2:15–50. doi: 10.1142/S0218271893000039
117. Grot N, Rovelli C, Tate RS. Time-of-arrival in quantum mechanics. *Phys Rev A*. (1996) 54:4676. doi: 10.1103/PhysRevA.54.4676
118. Aharonov Y, Oppenheim J, Popescu S, Reznik B, Unruh W. Measurement of time of arrival in quantum mechanics. *Phys Rev A*. (1998) 57:4130. doi: 10.1103/PhysRevA.57.4130
119. Muga JG, Leavens CR. Arrival time in quantum mechanics. *Phys Rep*. (2000) 338:353–438. doi: 10.1016/S0370-1573(00)00047-8
120. Aharonov Y, Bohm D. Time in the quantum theory and the uncertainty relation for time and energy. *Phys Rev*. (1961) 122:1649. doi: 10.1103/PhysRev.122.1649
121. Peres A. Measurement of time by quantum clocks. *Am J Phys*. (1980) 48:552–7. doi: 10.1119/1.12061
122. Pauli W. Die allgemeinen prinzipien der wellenmechanik. *Handbuch der Physik*. (1958) 5:1–168. doi: 10.1007/978-3-642-80539-4_1
123. Rovelli C. Partial observables. *Phys Rev D*. (2002) 65:124013. doi: 10.1103/PhysRevD.65.124013
124. Ashtekar A, Horowitz G. On the canonical approach to quantum gravity. *Phys Rev D*. (1982) 26:3342–53. doi: 10.1103/PhysRevD.26.3342
125. Kuchař KV. Covariant factor ordering of gauge systems. *Phys Rev D*. (1986) 34:3044–57. doi: 10.1103/PhysRevD.34.3044
126. Schleich K. Is reduced phase space quantization equivalent to Dirac quantization? *Class Quantum Grav*. (1990) 7:1529–38. doi: 10.1088/0264-9381/7/8/028
127. Romano JD, Tate RS. Dirac versus reduced space quantization of simple constrained systems. *Class Quantum Grav*. (1989) 6:1487. doi: 10.1088/0264-9381/6/10/017
128. Loll R. Noncommutativity of constraining and quantizing: a U(1) gauge model. *Phys Rev D*. (1990) 41:3785–91. doi: 10.1103/PhysRevD.41.3785
129. Kunstatter G. Dirac versus reduced quantization: a geometrical approach. *Class Quantum Grav*. (1992) 9:1469–86. doi: 10.1088/0264-9381/9/6/005
130. Bartlett SD, Rudolph T, Spekkens RW. Reference frames, superselection rules, and quantum information. *Rev Mod Phys*. (2007) 79:555–609. doi: 10.1103/RevModPhys.79.555
131. Smith ARH. Communicating without shared reference frames. *Phys Rev A*. (2019) 99:052315. doi: 10.1103/PhysRevA.99.052315
132. Smith ARH, Piani M, Mann RB. Quantum reference frames associated with noncompact groups: the case of translations and boosts and the role of mass. *Phys Rev A*. (2016) 94:012333. doi: 10.1103/PhysRevA.94.012333
133. Banerjee K, Calcagni G, Martín-Benito M. Introduction to loop quantum cosmology. *SIGMA*. (2012) 8:016. doi: 10.3842/SIGMA.2012.016
134. Page DN. *Time as an Inaccessible Observable*. NSF-ITP-89-18 (1989).
135. Page DN. Clock time and entropy. In: Halliwell JJ, Pérez-Mercader J, Zurek WH, editors. *Physical Origins of Time Asymmetry*. Cambridge: Cambridge University Press (1994). p. 287–98.
136. Smith ARH. Quantum time dilation: a new test of relativistic quantum theory. *arXiv*. (2020) 200410810.
137. Grochowski PT, Smith ARH, Dragan A, Dębski K. Quantum time dilation in atomic spectra. *Physic Rev Res*. (2021) 3:023053. doi: 10.1103/PhysRevResearch.3.023053
138. Zych M, Costa F, Pikovski I, Brukner Č. Quantum interferometric visibility as a witness of general relativistic proper time. *Nat Commun*. (2011) 2:505. doi: 10.1038/ncomms1498
139. Khandelwal S, Lock MPE, Woods MP. Universal quantum modifications to general relativistic time dilation in delocalised clocks. *Quantum*. (2020) 4:309. doi: 10.22331/q-2020-08-14-309
140. Paige AJ, Plato ADK, Kim MS. Classical and nonclassical time dilation for quantum clocks. *Phys Rev Lett*. (2020) 124:160602. doi: 10.1103/PhysRevLett.124.160602

Disclaimer: The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

Conflict of Interest: 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.

frontiers in Physics

ORIGINAL RESEARCH
published: 28 July 2021
doi: 10.3389/fphy.2021.534265

Functorial Evolution of Quantum Fields

Stefano Gogioso, Maria E. Stasinou and Bob Coecke*

Department of Computer Science, University of Oxford, Oxford, United Kingdom

We present a compositional algebraic framework to describe the evolution of quantum fields in discretised spacetimes. We show how familiar notions from Relativity and quantum causality can be recovered in a purely order-theoretic way from the causal order of events in spacetime, with no direct mention of analysis or topology. We formulate theory-independent notions of fields over causal orders in a compositional, functorial way. We draw a strong connection to Algebraic Quantum Field Theory (AQFT), using a sheaf-theoretical approach in our definition of spaces of states over regions of spacetime. We introduce notions of symmetry and cellular automata, which we show to subsume existing definitions of Quantum Cellular Automata (QCA) from previous literature. Given the extreme flexibility of our constructions, we propose that our framework be used as the starting point for new developments in AQFT, QCA and more generally Quantum Field Theory.

Keywords: causality, quantum field theory, relativity, algebraic quantum field theory, quantum cellular automata

OPEN ACCESS

Edited by:
Caslav Brukner,
University of Vienna, Austria

Reviewed by:
Pablo Arrighi,
Aix-Marseille Université, France
Sayantan Choudhury,
National Institute of Science Education and Research (NISER), India

****Correspondence:***
Stefano Gogioso
stefano.gogioso@cs.ox.ac.uk

Specialty section:
This article was submitted to High-Energy and Astroparticle Physics, a section of the journal Frontiers in Physics

Received: *11 February 2020*
Accepted: *12 January 2021*
Published: *28 July 2021*

Citation:
Gogioso S, Stasinou ME and Coecke B (2021) Functorial Evolution of Quantum Fields. Front. Phys. 9:534265. doi: 10.3389/fphy.2021.534265

1 INTRODUCTION

Like much of classical physics, the study of Relativity and quantum field theory has deep roots in topology and geometry. However, recent years have seen a steady shift from the traditional approaches to a more abstract algebraic perspective, based on the identification of spacetime structure with causal order.

This new way of looking at causality finds its origin in a much-celebrated result by Malament [1], itself based on previous work by Kronheimer, Penrose, Hawking, King and McCarthy [2, 3]. If M is a Lorentzian manifold, we say that M is future- (resp. past-) distinguishing iff two events $x, y \in M$ (i.e. two spacetime points) having the same exact causal future (resp. past) are necessarily identical[1]. Given a Lorentzian manifold M, we can define a partial order $\leq_M$ between its events—the causal order—by setting $x \leq_M y$ iff x causally precedes y in M, i.e. iff there exists a future-directed causal curve—a smooth curve in M with everywhere future-directed time-like or light-like tangent vector—from x to y. The 1977 result by Malament [1] can then be stated as follows.

Theorem 1: *Let M and M' be two Lorentzian manifolds, both manifolds being future-and-past–distinguishing. The associated causal orders $(M, \leq_M)$ and $(M', \leq_{M'})$ are order-isomorphic if and only if M and M' are conformally equivalent.*

While the result by Malament guarantees that future-and-past–distinguishing manifolds (up to conformal equivalence) can be identified with their causal orders, it does not provide a characterization of which partial orders arise as causal orders on manifolds (or restrictions thereof to manifold subsets). This lack of exact correspondence between topology and order is the motivation behind many past and current lines of enquiry. Notable mention in this regard is deserved by the work of [4, 5], which aims to formulate causal order in terms of partial orders and domain theory. Within that framework, a complete characterization of which partial orders arise as the causal orders of Lorentzian manifolds is still an open question.

[1]The requirement for a manifold M to be future- and past-distinguishing is essentially one of well-behaviour, e.g. excluding causal violations such as closed timelike curves (all points of which necessarily have the same causal past and future).

A different approach to the order-theoretic study of spacetime is given by the causal sets research program (cf [6, 7]). A causal set is a poset which is locally finite, i.e. such that for every $x, y \in C$ the subset $\{z \in C | x \leq z \leq y\}$ is finite.[2] Causal sets arise as discrete subsets of Lorentzian manifolds (under the causal order inherited by restriction) and a fundamental pursuit for the community is a characterisation of the large-scale properties of spacetime as emergent from a discrete small-scale structure. In particular, the question whether a causal set can always be (suitably) embedded as a discrete subset of a Lorentzian manifold is central to the programme and—as far as we are aware—one which is still to be completely answered [7].

When it comes to incorporating quantum fields into the spacetimes, efforts have mostly been focused in three directions: algebraic approaches, topological approaches and quantum cellular automata.

The algebraic approaches take a functorial and sheaf-theoretic view of quantum fields, studying the local structure of fields through the algebras of observables—usually C*-algebras or von Neumann algebras—over the regions of spacetime. Prominent examples include Algebraic Quantum Field Theory (AQFT) [9, 10] and the topos-theoretic programmes [11, 12]. Presheaves are special functors used to associate (field) data to spacetime regions, in a way which is guaranteed to respects causality and locality constraints imposed by space-time topology. We will take a deeper look at this approach in **Section 5**.

The topological approaches focus instead on global aspects of relativistic quantum fields, foregoing any possibility of studying local structure by requiring that field theories be topological, i.e. invariant under large scale deformations of spacetime. The resulting Topological Quantum Field Theories (TQFTs) [13–15] have achieved enormous success in fields such as condensed matter theory and quantum error correction. Like AQFT, TQFTs have a categorical formulation as functors from a category of spacetime "pieces" to categories of Vector spaces and algebras. The difference is in the nature of those "pieces": in AQFT a spacetime is given and the order structure of its regions is considered; in TQFT, on the other hand, (equivalence classes of) basic topological manifolds are given, which can be combined together to form myriad different spacetimes.

The approaches based on Quantum Cellular Automata (QCA) [16–18], finally, attempt to tame the issues with the formulation of quantum field theory by positing that full-fledged quantum fields in spacetime can be understood as the continuous limit of much-more-manageable theories, dealing with quantum fields living on discrete lattices and subject to discrete time evolution (known as Quantum Cellular Automata).

In this work, we propose to use tools from category theory to unify key aspects of the approaches above under a single generalized framework. Specifically, our work is part of an effort to gain an operational, process-theoretic understanding of the relationship between quantum theory and Relativistic causality [8, 19–21]. Our key contribution, across the next four sections, will be the formulation of a functorial and theory-independent notion of field theory based solely on the order-theoretic structure of causality. To exemplify the flexibility of our construction, in **Section 5** we will build a strong connection to Algebraic Quantum Field Theory, based on a sheaf-theoretic formulation of states over regions. In **Section 6**, finally, we will formulate a notion of cellular automaton which encompasses and greatly generalizes notions of QCA from existing literature.

2 CAUSAL ORDERS

In this work, we will consider posets as an abstract model of causally well-behaved spacetimes. This means that we will be working in the category Pos of posets and monotone maps between them, with Malament's result [1] showing that future-and-past–distinguishing conformal Lorentzian manifolds embed into Pos. To highlight the intended relationship to spacetimes, we will refer to partial orders as causal orders for the remainder of this work.

Definition 2: *By a causal order we mean a poset $\Omega = (|\Omega|, \leq)$, i.e. a set $|\Omega|$ equipped with a partial order $\leq$ on it. We refer to the elements of Ω as events. Given two events $x, y \in \Omega$ we say that x causally precedes* y *(equivalently that* y *causally follows* x*) iff $x \leq y$. We say that x and y are causally related iff $x \leq y$ or $y \leq x$. A causal sub-order Ω' of a causal order Ω is a subset $|\Omega'| \subseteq |\Omega|$ endowed with the structure of a poset by restriction.*[3]

As we now proceed to demonstrate, several familiar concepts from Relativity can be defined in a purely combinatorial manner on partial orders.

2.1 Causal Paths

Definition 3: *Let Ω be a causal order and let $x, y \in \Omega$ be two events. A causal path from* x *to* y *is a maximal totally ordered subset $\gamma \subseteq \Omega$ such that $x = \min \gamma$ and $y = \max \gamma$. Maximality of the subset $\gamma \subseteq \Omega$ here means that there is no total order $\gamma' \subseteq \Omega$ strictly containing gamma and such that $x = \min \gamma'$ and $y = \max \gamma'$. We write $\gamma : x \rightsquigarrow y$ to denote that γ is a causal path from x to y.*

The causal diamond from x to y in a causal order Ω is the union of all causal paths $x \rightsquigarrow y$ in Ω. Furthermore, causal paths in Ω can be naturally organized into a category as follows:

- The objects are the events $x \in \Omega$;
- The morphisms from x to y are the paths $x \rightsquigarrow y$;
- The identity morphism on x is the singleton path $\{x\} : x \rightsquigarrow x$;
- Composition of two paths $\gamma : x \rightsquigarrow y$ and $\xi : y \rightsquigarrow z$ is the set-theoretic union of the subsets $\gamma, \xi \subseteq \Omega$:

$$\xi \circ \gamma := (\xi \cup \gamma) : x \rightsquigarrow z. \tag{1}$$

[2]The local finiteness condition for a causal set can equivalently be stated as the requirement that the partial order arises as the reflexive-transitive closure of a non-transitive directed graph, its Hasse diagram (see e.g. [8]).

[3]I.e. such that for all $x, y \in |\Omega'|$ we have that $x \leq y$ in Ω' if and only if $x \leq y$ in Ω.

Definition 4: *Let Ω be a causal order and let $x \in \Omega$ be an event. The causal future $J^+(x)$ of x is the set of all events y which causally follow it:*

$$J^+(x) := \{y \in \Omega \,|\, x \leq y\}. \tag{2}$$

Similarly, the causal past $J^-(x)$ of x is the set of all events y which causally precede it:

$$J^-(x) := \{y \in \Omega \,|\, y \leq x\}. \tag{3}$$

We also define causal future and past for arbitrary subsets $A \subseteq \Omega$ by union:

$$J^+(A) := \bigcup_{x \in A} J^+(x) \quad J^-(A) := \bigcup_{x \in A} J^-(x). \tag{4}$$

Remark 5: *A causal order Ω is automatically future-and-past–distinguishing. To see this, assume that $J^+(x) = J^+(y)$ for some $x, y \in \Omega$: then both $x \in J^+(x) = J^+(y)$, implying $y \leq x$, and $y \in J^+(y) = J^+(x)$, implying $x \leq y$, so that $x = y$ by antisymmetry of the partial order $\leq$. The assumption that $J^-(x) = J^-(y)$ analogously implies that $x = y$.*

Definition 6: *Let Ω be a causal order and let $x \in \Omega$ be an event. By a causal path $\gamma : x \rightsquigarrow +\infty$ (resp. $\gamma : -\infty \rightsquigarrow x$) we denote a maximal totally ordered subset $\gamma \subseteq \Omega$ such that $x = \min \gamma$ (resp. $x = \max \gamma$). If Ω has a global maximum (resp. global minimum), then we denote it by $+\infty$ (resp. $-\infty$) for consistency with our previous definition of causal paths, otherwise the symbol $+\infty$ (resp. $-\infty$) is never used to denote an actual element of* C.

The causal future (resp. causal past) of an event x is the union of all causal paths $x \rightsquigarrow +\infty$ (resp. $-\infty \rightsquigarrow x$).

2.2 Space-Like Slices

Definition 7: *Let Ω be a causal order and let $A \subseteq \Omega$ be any subset. The future domain of dependence $D^+(A)$ of* A *is the subset of all events $x \in \Omega$ which "necessarily causally follow* A*," in the sense that every causal path $-\infty \rightsquigarrow x$ intersects A:*

$$D^+(A) := \{x \in \Omega \,|\, \forall \gamma : -\infty \rightsquigarrow x.\ \gamma \cap A \neq \emptyset\}. \tag{5}$$

The past domain of dependence $D^-(A)$ of A is the subset of all events $x \in \Omega$ which "necessarily causally precede A", in the sense that every causal path $x \rightsquigarrow +\infty$ intersects A:

$$D^-(A) := \{x \in \Omega \,|\, \forall \gamma : x \rightsquigarrow +\infty.\ \gamma \cap A \neq \emptyset\}. \tag{6}$$

The domains of dependence of a subset A are related to its past and future by the following two Propositions.

Proposition 8: *Let Ω be a causal order and let $A \subseteq \Omega$ be any subset. Then $D^+(A) \subseteq J^+(A)$ and $D^-(A) \subseteq J^-(A)$.*

Proof: Let $x \in D^+(A)$ be any event in the future domain of dependence of A. The set of causal paths $-\infty \rightsquigarrow x$ is necessarily non-empty, because there must be at least one such path extending the singleton path $\{x\} : x \rightsquigarrow x$. Let $\gamma : -\infty \rightsquigarrow x$ be one such path. Because $x \in D^+(A)$, γ must intersect A at some point $y \leq x$, and we define $\gamma' := \gamma \cap J^+(\{y\}) \neq \emptyset$. By definition, $y = \min \gamma'$. Because $J^+(\{y\})$ is upward-closed, $x = \max \gamma'$ and $\gamma' : y \rightsquigarrow x$ is such that $\gamma' \subseteq J^+(\{y\}) \subseteq J^+(A)$, so we conclude that $x \in J^+(A)$. The proof that $D^-(A) \subseteq J^-(A)$ is analogous. □

Proposition 9: *Let Ω be a causal order and let $A \subseteq \Omega$ be any subset. If $B \subseteq D^+(A)$ then $J^+(B) \subseteq J^+(A)$ and $J^-(B) \subseteq J^-(A) \cup J^+(A)$. Dually, if $B \subseteq D^-(A)$ then $J^-(B) \subseteq J^-(A)$ and $J^+(B) \subseteq J^-(A) \cup J^+(A)$.*

Proof: Without loss of generality, assume $B \subseteq D^+(A)$—the case $B \subseteq D^-(A)$ is proven analogously. From **Proposition 8** we have that $B \subseteq D^+(A) \subseteq J^+(A)$, so we conclude that $J^+(B) \subseteq J^+(A)$ by upward-closure of $J^+(A)$. Now consider $x \in J^-(B)$. Let $\gamma : x \rightsquigarrow y$ be any path with $y \in B$ and let $\gamma' : -\infty \rightsquigarrow y$ be any path extending γ. Because $B \subseteq D^+(A)$, the intersection $\gamma' \cap A$ contains at least some point z. Because γ' is totally ordered, we have two possible cases: $z \leq x$ and $z \geq x$. If $z \leq x$, then $\gamma' \cap J^+(\{z\}) \cap J^-(\{x\}) : z \rightsquigarrow x$ shows that $x \in J^+(A)$. If $z \geq x$, then $\gamma' \cap J^-(\{z\}) \cap J^+(\{x\}) : x \rightsquigarrow z$ shows that $x \in J^-(A)$. □

Definition 10: *Let Ω be a causal order. We say that two events x, y are space-like separated if they are not causally related, i.e. if neither $x \leq y$ nor $y \leq x$. Consequently, we define a (space-like) slice Σ in Ω to be an antichain, i.e. a subset $\Sigma \subseteq \Omega$ such that $(\Sigma, \leq)$ is a discrete partial order (equivalently, any two distinct $x, y \in \Sigma$ are space-like separated).*

Definition 11: *Let Ω be a causal order and let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ be a collection of subsets of Ω. We say that the subsets in $\mathcal{A}$ are* space-like separated *if the following conditions holds for all distinct $A, B \in \mathcal{A}$:*

$$A \cap (J^+(B) \cup J^-(B)) = \emptyset. \tag{7}$$

In particular, a space-like slice is the union of a collection of space-like separated singleton subsets. See **Figure 1** for examples.

More than diamonds or paths, slices are the focus of this work. Space-like slices are a generalization of space-like surfaces from Relativity: the term "slice" is used here in place of "surface" because the latter traditionally implies some topological conditions.

Definition 12: *Let Ω be a causal order. The category of all slices on Ω, denoted by $\mathit{Slices}(\Omega)$, is the strict partially monoidal category [22] defined as follows.*

- *Objects of $\mathit{Slices}(\Omega)$ are the slices of Ω.*
- *The category is a poset and the unique morphism from a space-like slice Σ to another space-like slice Γ is denoted $\Sigma \twoheadrightarrow \Gamma$ if it exists. Specifically, we say that $\Sigma \twoheadrightarrow \Gamma$ if and only if $\Gamma \subseteq D^+(\Sigma)$, i.e. iff Γ lies entirely into the future domain of dependence of Σ. See **Figures 2**, **3** for examples.*
- *The monoidal product on objects $\Sigma \otimes \Gamma$ is only defined when Σ and Γ are space-like separated, in which case it is the disjoint union $\Sigma \sqcup \Gamma$.*
- *The unit for the monoidal product is the empty space-like slice $\emptyset \subseteq \Omega$.*
- *The partial monoidal product on objects extends to morphisms because whenever $\Sigma' \subseteq D^+(\Sigma)$ and $\Gamma' \subseteq D^+(\Gamma)$—i.e. whenever $\Sigma \twoheadrightarrow \Sigma'$ and $\Gamma \twoheadrightarrow \Gamma'$—we necessarily have:*

$$\Sigma' \sqcup \Gamma' \subseteq D^+(\Sigma) \cup D^+(\Gamma) \subseteq D^+(\Sigma \sqcup \Gamma), \quad \text{i.e. } \Sigma \otimes \Gamma \twoheadrightarrow \Sigma' \otimes \Gamma'. \tag{8}$$

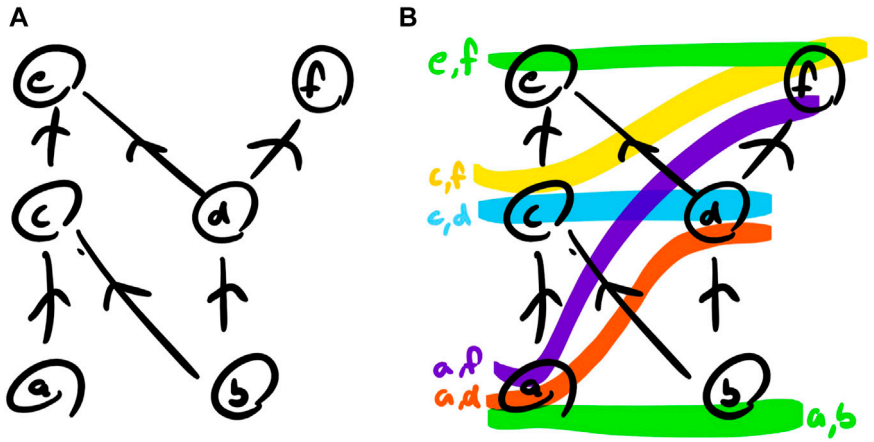

FIGURE 1 | (A): the Hasse diagram for a causal order on six events $\{a, b, c, d, e, f\}$. **(B)**: the maximal slices for the causal order highlighted (all other slices can be obtained as subsets of the maximal slices).

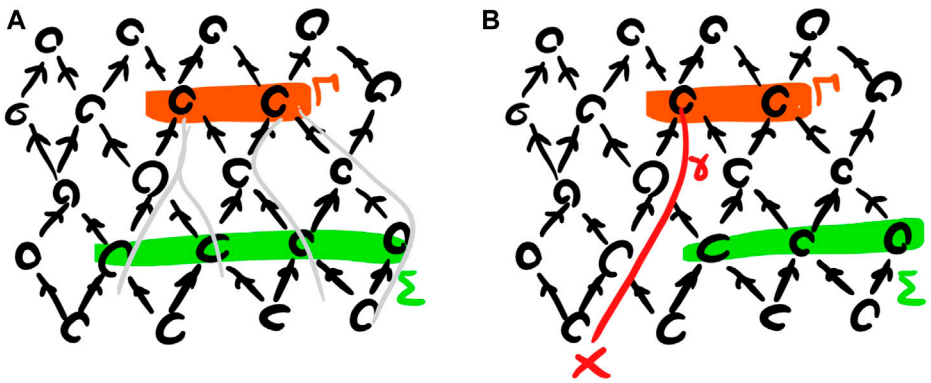

FIGURE 2 | (A): two slices Σ, Γ such that $\Sigma \twoheadrightarrow \Gamma$. **(B)**: two slices Σ, Γ such that $\Sigma \twoheadrightarrow$ &*doublehyphen*; 9*pt*/ Γ, highlighting a past-directed path γ starting from an event of Γ and not intersecting Σ at any point.

The partial monoidal product is strict, i.e. strictly associative and unital[4] *when all products are defined. The partial monoidal product is also commutative, i.e. it is symmetric (wherever defined) with an identity* $\Sigma \otimes \Gamma = \Gamma \otimes \Sigma$ *as the symmetry isomorphism.*

The order relation $\Sigma \twoheadrightarrow \Gamma$ on slices has been defined in such a way as to ensure that the field state local to the codomain slice Γ will be entirely determined by evolution and marginalization of the field state on the domain slice Σ. In particular, the definition is such that any sub-slice $\Sigma' \subseteq \Sigma$ necessarily satisfies $\Sigma \twoheadrightarrow \Sigma'$, since the field state on Σ' can be obtained from the field state on Σ by marginalization/discarding. The connection to marginalisation will be discussed in further detail in **Section 4.3** below.

2.3 Diamonds and Regions

Let Ω *be a causal order. If* x, y *are two events in* Ω*, the causal diamond from* x *to* y *in* Ω *is the causal sub-order* $(\Diamond_{x,y}, \leq) \hookrightarrow \Omega$ *defined as follows:*

$$\Diamond_{x,y} := \{z \in \Omega \,|\, x \leq z \leq y\} = \bigcup_{\gamma : x \rightsquigarrow y} \gamma. \tag{9}$$

Definition 14: *Let* Ω *be a causal order. A region in* Ω *is a causal sub-order* $(R, \leq) \hookrightarrow \Omega$ *which is convex, i.e. one such that for all events* $x, y \in R$ *the causal diamond from* x *to* y *in* Ω *is a subset of* R *(i.e.* R *contains all paths* $\gamma : x \rightsquigarrow y$ *in* Ω*).*

Definition 14 is the order-theoretic incarnation of the requirement that causal diamonds generate the topology of Lorentzian manifolds: we could have equivalently stated it as saying that regions in Ω are all the possibly unions of causal diamonds in Ω (including the empty one). A special case of region of particular interest is the region between two slices $\Sigma \twoheadrightarrow \Gamma$.

Definition 15: *Let* Ω *be a causal order and consider two slices* $\Sigma \twoheadrightarrow \Gamma$*. We define the region between* Σ *and* Γ *as follows:*

$$\Diamond_{\Sigma,\Gamma} := \bigcup_{x \in \Sigma} \bigcup_{y \in \Gamma} \Diamond_{x,y} \tag{10}$$

In particular, a causal diamond $\Diamond_{x,y}$ *is the region between the slices* $\{x\}$ *and* $\{y\}$*. More generally, a region between slices* Σ *and* Γ *is the intersection* $\Diamond_{\Sigma,\Gamma} = J^+(\Sigma) \cap J^-(\Gamma)$ *of their future and past respectively.*

[4]I.e. we have that $\Sigma \otimes \varnothing = \Sigma = \varnothing \otimes \Sigma$, for all slices Σ.

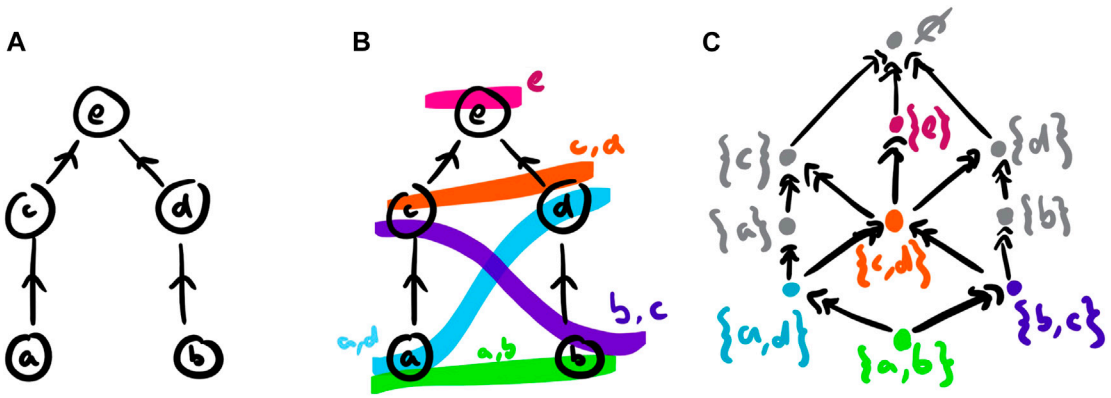

FIGURE 3 | (A): the Hasse diagram for a causal order. **(B)**: the maximal slices for the causal order highlighted. **(C)**: the category of all slices for the causal order.

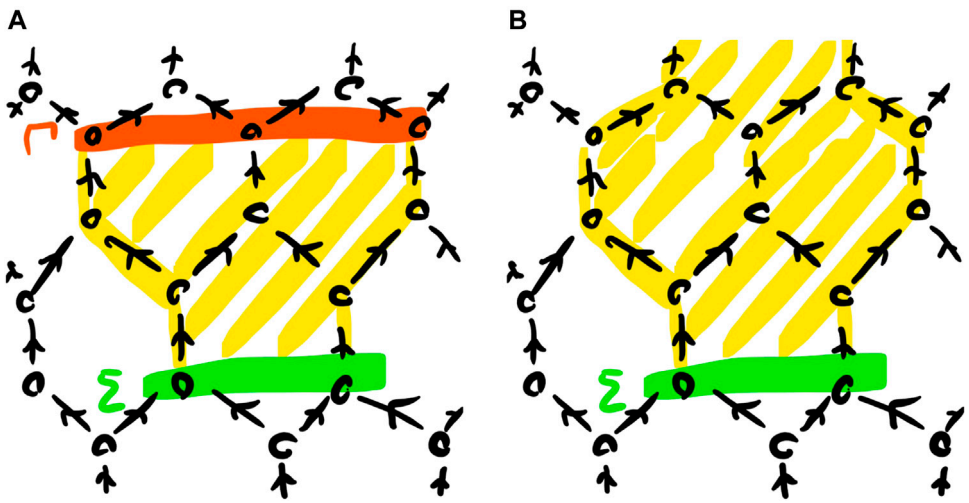

FIGURE 4 | (A): the region between two slices on the honeycomb lattice. **(B)**: an unbounded (necessarily infinite) region on the honeycomb lattice.

The slices Σ and Γ bounding the region $\Diamond_{\Sigma,\Gamma}$ can be obtained respectively as the sets of its minima $\Sigma = \min \Diamond_{\Sigma,\Gamma}$ and of its maxima $\Gamma = \max \Diamond_{\Sigma,\Gamma}$. As a special case, a slice Σ is the region between Σ and Σ. Conversely, every closed bounded region R—and in particular every finite region—is in the form $R = \Diamond_{\min R, \max R}$. See **Figure 4** for examples.

3 CATEGORIES OF SLICES

Because we didn't impose any topological constraints on the slices, it is possible that the category $\mathrm{Slices}(\Omega)$ will, in practice, contain objects which are too irregular or exotic for physical fields to be defined over (such as fractal slices with low topological dimension). To obviate this issue, we consider more general categories of slices on a given causal order: this will allow us to restrict our attention to slices with any properties we desire, as long as we retain enough slices to reconstruct the structure of the causal order Ω, both 1) globally and 2) locally.

No requirement is made for all products that exist in $\mathrm{Slices}(\Omega)$ to also exist on members of a more general category of slices: it is the case that certain properties desirable in practice may not be closed under arbitrary union of space-like separated slices themselves satisfying the property.[5] However, we impose the requirement 3) that these more general categories of slices be partially monoidal sub-categories of $\mathrm{Slices}(\Omega)$.

Definition 16: *Let Ω be a causal order. A category of slices on Ω is the full sub-category $\mathcal{C}$ of $\mathrm{Slices}(\Omega)$ defined by a given set*

[5]An example of this phenomenon is given by constant-time partial Cauchy slices in Minkowski spacetime: the union of two disjoint constant-time partial Cauchy slices having the same time parameter yields another constant-time partial Cauchy slice, but the union of two space-like separated constant-time partial Cauchy slices having different time parameters does not yield a constant-time partial Cauchy slice as a result.

obj(*C*) *of slices chosen in such a way that the following three conditions hold.*

(1) *For any two events* $x, y \in \Omega$ *with* $x \leq y$*, there exist slices* $\Sigma, \Gamma \in obj(\mathcal{C})$ *such that* $x \in \Sigma$, $y \in \Gamma$ and $\Sigma \twoheadrightarrow \Gamma$.
(2) *If* Σ, Γ *and* Δ *are three slices in* $\mathcal{C}$*, then the restriction* $(\Delta \cap \diamondsuit_{\Sigma,\Gamma})$ *of* Δ *to the region* $\diamondsuit_{\Sigma,\Gamma}$ *is also a slice in* $\mathcal{C}$.
(3) *The category of slices* $\mathcal{C}$ *is a partially monoidal subcategory of Slices*(Ω)*. In particular,* $\emptyset \in obj(\mathcal{C})$ *and whenever* $\Sigma \otimes \Gamma$ *exists in* $\mathcal{C}$ *for some* $\Sigma, \Gamma \in obj(\mathcal{C})$ *then* $\Sigma \otimes \Gamma$ *also exists in Slices*(Ω) *(Associativity and unitality of* $\otimes$ *are strict in* $\mathcal{C}$ *as they are in Slices*(Ω)*).*

In particular, Slices(Ω) *is itself a category of slices on* Ω.

Condition (2) in the definition above tells us that we can talk about regions directly within a given category $\mathcal{C}$ of slices, without first having to reconstruct the causal order Ω: this will form the basis of the connection to AQFT in **Section 5** below.

As an example of particularly well-behaved slices, we define a notion of Cauchy slices—akin to that of Cauchy surfaces from Relativity—and remark that any "foliation" of a causal order in terms of such slices gives rise to what is arguably the simplest non-trivial example of category of slices.

Definition 17: *A slice* Σ *on* Ω *is a Cauchy slice if every causal path* $\gamma : -\infty \rightsquigarrow +\infty$ *in* Ω *intersects* Σ *at some (necessarily unique) event. Cauchy slices are in particular maximal slices. A category of Cauchy slices on* Ω *is a category* $\mathcal{C}$ *of slices on* Ω *such that every slice* $\Sigma \in obj(\mathcal{C})$ *is a subset* $\Sigma \subseteq \Gamma$ *of some* Cauchy slice $\Gamma \in obj(\mathcal{C})$.

Proposition 18: *A foliation on a causal order* Ω *is a set* $\mathcal{F}$ *of Cauchy slices on* Ω *such that:*

(1) *The slices in* $\mathcal{F}$ *are totally ordered according to* $\twoheadrightarrow$*;*
(2) *Every event* $x \in \Omega$ *is contained in some slice* $\Sigma \in \mathcal{F}$*;*
(3) *The slices in* $\mathcal{F}$ *are pairwise disjoint.*

If $\mathcal{F}$ *is a foliation, write CauchySlices*$(\mathcal{F})$ *for the full sub-category of Slices*(Ω) *generated by all slices which are subsets of some Cauchy slice in* $\mathcal{F}$*. Then CauchySlices*$(\mathcal{F})$ *is a category of Cauchy slices on* Ω.

Proof: Let CauchySlices$(\mathcal{F})$ denote the full sub-category of Slices(Ω) generated by all slices which are subsets of some Cauchy slice.

For any two events $x \leq y$ in Ω, let $\Sigma, \Gamma \in \text{obj}(\text{CauchySlices}(\mathcal{F}))$ be two Cauchy slices such that $x \in \Sigma$ and $y \in \Gamma$, the existence of such slices guaranteed by the definition of foliation. Because the foliation is totally ordered, we have that $\Sigma \twoheadrightarrow \Gamma$ or $\Gamma \twoheadrightarrow \Sigma$ (or both, if $\Sigma = \Gamma$ and $x = y$). If $x = y$, either works, while if $x < y$ then necessarily $\Sigma \twoheadrightarrow \Gamma$. Either way, condition (1) for CauchySlices$(\mathcal{F})$ to be a category of slices is satisfied. □

Let Σ', Γ' and Δ' be three slices, respectively contained in three Cauchy slices Σ, Γ and Δ inside the foliation. Because of total ordering and disjointness of slices in $\mathcal{F}$, the only instance in which $\Delta \cap \diamondsuit_{\Sigma,\Gamma} \neq \emptyset$ is when $\Sigma \twoheadrightarrow \Delta \twoheadrightarrow \Gamma$. In this case, $\Delta \cap \diamondsuit_{\Sigma,\Gamma} = \Delta \in \text{obj}(\text{Cauchy Slices}(\mathcal{F}))$. Otherwise, $\Delta \cap \diamondsuit_{\Sigma,\Gamma} = \emptyset \in \text{obj}(\text{Cauchy Slices}(\mathcal{F}))$. Either way, condition (2) for Cauchy Slices$(\mathcal{F})$ to be a category of slices is satisfied when Σ, Γ and Δ are Cauchy slices. This result immediately generalises to Σ', Γ' and Δ': we have that $\Delta' \cap \diamondsuit_{\Sigma',\Gamma'} \subseteq \Delta \cap \diamondsuit_{\Sigma,\Gamma} \subseteq \Delta$, so that $\Delta' \in \text{obj}(\text{Cauchy Slices}(\mathcal{F}))$ and condition (2) for Cauchy Slices$(\mathcal{F})$ to be a category of slices is satisfied.

Finally, if Σ, Γ are two slices such that $\Sigma \otimes \Gamma$ is defined in Cauchy Slices$(\mathcal{F})$, then Σ, Γ are necessarily disjoint subsets of the same Cauchy slice Δ. It is then immediate to conclude that condition (3) for Cauchy Slices$(\mathcal{F})$ to be a category of slices is satisfied.

3.1 The category of Causal Orders

As objects, causal orders have been defined simply as posets. However, causal orders are note simply posets, and this should be reflected in the kind of morphisms that can be used to related them to one another. Malament's result [1] may seem at first to indicate that order-preserving maps are the correct choice, but upon closer inspection one realises that the result itself only talks about order-preserving isomorphisms, giving no indication about other maps.

A prototypical example of the behavior we wish to avoid is that where $\Omega' \hookrightarrow \Omega$ is a sub-poset such that $x \leq y$ in Ω for some $x, y \in \Omega'$ but $x \not\preceq y$ in Ω'. The issue above is the reason behind the rather specific formulation of the notion of causal sub-order in **Definition 2**, prompting us to choose a special subclass of order-preserving maps as morphisms between causal orders.

Definition 19: *The category* CausOrd *of causal orders is the symmetric monoidal category defined as follows:*

- *Objects of* CausOrd *are causal orders, i.e. posets.*
- *Morphisms* $\Omega \to \Theta$ *in* CausOrd *are the order-preserving functions* $f : \Omega \to \Theta$ *such that we have* $x < y$ *in* Ω *whenever we have* $f(x) < f(y)$ *in* Θ.
- *The monoidal product on objects* $\Omega \otimes \Theta$ *is the (forcedly) disjoint union* $\Omega \sqcup \Theta := \Omega \times \{0\} \cup \Theta \times \{1\}$.
- *The unit for the monoidal product is the empty causal order* $\emptyset$.
- *The monoidal product extends to the disjoint union of morphisms. If* $f : \Omega \to \Omega'$ *and* $g : \Theta \to \Theta'$*, then the monoidal product* $f \otimes g : \Omega \otimes \Theta \to \Omega' \otimes \Theta'$ *is defined as follows:*

$$f \otimes g := f \sqcup g = (x, i) \mapsto \begin{cases} (f(x), 0) \in \Omega' \times \{0\} & \text{if } i = 0 \\ (g(x), 1) \in \Theta' \times \{1\} & \text{if } i = 1 \end{cases}. \tag{11}$$

The monoidal product is not strict nor commutative, but symmetric under the symmetry isomorphisms $s : \Omega \sqcup \Theta \to \Theta \sqcup \Omega$ *defined by* $s(x, i) = (x, 1 - i)$.

It is easy to check that the causal sub-orders Ω' of a causal order Ω according to **Definition 2** are all sub-objects $\Omega' \hookrightarrow \Omega$ in the category CausOrd, so that the notion of causal sub-order is consistent with the usual notion of categorical sub-object. As discussed above, the regions in a causal order Ω are examples of causal sub-orders, but not all sub-orders are regions: e.g. paths are always sub-orders but not necessarily regions. In general, if we

have $\Omega' \hookrightarrow \Omega$ then it is not necessary for Ω' to be convex, i.e. it is not necessary for Ω' to contain all paths $x \rightsquigarrow y$ in Ω for any two events $x, y \in \Omega'$: in the sense, the causal sub-order Ω' can "coarsen" the causal order Ω by dropping events "in between" events of the latter. As the following proposition shows, this "coarsening" of causal orders is the only other case we need to consider when talking about causal sub-orders.

Definition 20: *Let Ω be a causal order and let the morphism $\mathrm{i} : \Omega' \hookrightarrow \Omega$ be a causal sub-order of Ω. We say that the morphism $\mathrm{i} : \Omega' \hookrightarrow \Omega$ is a region if the image $i(\Omega') \subseteq \Omega$ is a region in Ω. We say that the morphism $\mathrm{i} : \Omega' \hookrightarrow \Omega$ is a coarsening if the image $i(\Omega') \subseteq \Omega$ is such that for all $x \leq y \in \Omega$ there exist $x', y' \in \Omega'$ with $i(x') \leq x \leq y \leq i(y')$.*

Proposition 21: *Let Ω be a causal order and let $\mathrm{i} : \Omega' \hookrightarrow \Omega$ be a causal sub-order of Ω. Then i factors (essentially) uniquely as $i = r \circ f$ for some region $\mathrm{r} : \Theta \hookrightarrow \Omega$ and some coarsening $\mathrm{f} : \Omega' \hookrightarrow \Theta$.*

Proof: Let Θ be the region of Ω obtained as the union of the causal diamonds $\Diamond_{x,y}$ for all $x, y \in i(\Omega')$. Let $r : \Theta \hookrightarrow \Omega$ be the injection of Θ into Ω as a sub-poset and let $f : \Omega' \hookrightarrow \Theta$ be the restriction of the codomain of i to Θ: clearly $i = r \circ f$, r is a region and f is a coarsening (because of how Θ was constructed).

Now let Θ' be such that $r' : \Theta' \hookrightarrow \Omega$ is a region and $f' : \Omega' \hookrightarrow \Theta'$ is a coarsening with $i = r' \circ f'$: to prove essential uniqueness, we want to show that there is some isomorphism $\theta : \Theta' \to \Theta$ such that $f = \theta \circ f'$ and $r' = r \circ \theta$. Because $r' \circ f' = i$, the image $r'(\Theta')$ is the region Θ itself, so that the restriction $\theta : \Theta' \to \Theta$ of the codomain of r' to Θ is an isomorphism with $r' = r \circ \theta$. Now we have $r \circ f = i = r' \circ f' = r \circ (\theta \circ f')$: but r is a monomorphism (i.e. it is injective), so necessarily $f = \theta \circ f'$. □

The category CausOrd also has epi-mono factorization, i.e. every morphism $f : \Omega' \to \Omega$ can be factorised (essentially) uniquely as an epimorphism (i.e. a surjective map) $q : \Omega' \to \Theta$ and a monomorphism (i.e. an injective map) $i : \Theta \hookrightarrow \Omega$. We have already adopted the nomenclature of causal suborder for the latter form of morphism, while we will henceforth use causal quotient to refer to the former. Causal quotients are surjective morphisms which "collapse" several events into one, in a way which respects the causal order: as an example, a snippet of the causal quotient $q : H \to D$ from the (infinite) honeycomb lattice to the (infinite) diamond lattice is shown in **Figure 5**.

If Σ is a slice in Ω, we can define its pullback $f^*(\Sigma)$ to be the causal suborder of Θ generated by $\{x \in \Theta \,|\, f(x) \in \Sigma\}$, i.e. largest causal sub-order of Θ mapped onto Σ. The pullback of a slice Σ has a rather simple structure: the slices Γ in the pullback $f^*(\Sigma)$ are exactly the disjoint unions $\Gamma := \sqcup_{x \in \Sigma} \Gamma_x$ for all possible choices $(\Gamma_x)_{x \in \Sigma}$ of slice sections of f over the individual events x of Σ.

$$(\Gamma_x)_{x \in \Sigma} \in \coprod_{x \in \Sigma} \mathrm{obj}\left(\mathrm{Slices}\left(f^*(\{x\})\right)\right) \tag{12}$$

A depiction of the pullback under the causal quotient $q : H \to D$ described above can be seen in **Figure 6**.

If $\mathcal{C}$ is a category of slices on Ω, we can define its pullback along f to be the full sub-category $f^*(\mathcal{C})$ of $\mathrm{Slices}(\Theta)$ spanned by all slices Γ in Θ such that $\Gamma \in \mathrm{obj}\left(\mathrm{Slices}\left(f^*(\Sigma)\right)\right)$ for some $\Sigma \in \mathrm{obj}(\mathcal{C})$. The relationship $\twoheadrightarrow$ between slices in pullbacks is a little complicated and its full characterization is left to future work.

Remark 22: *The two notions of pullback defined above—for slices and for categories of slices—are related by the observation*

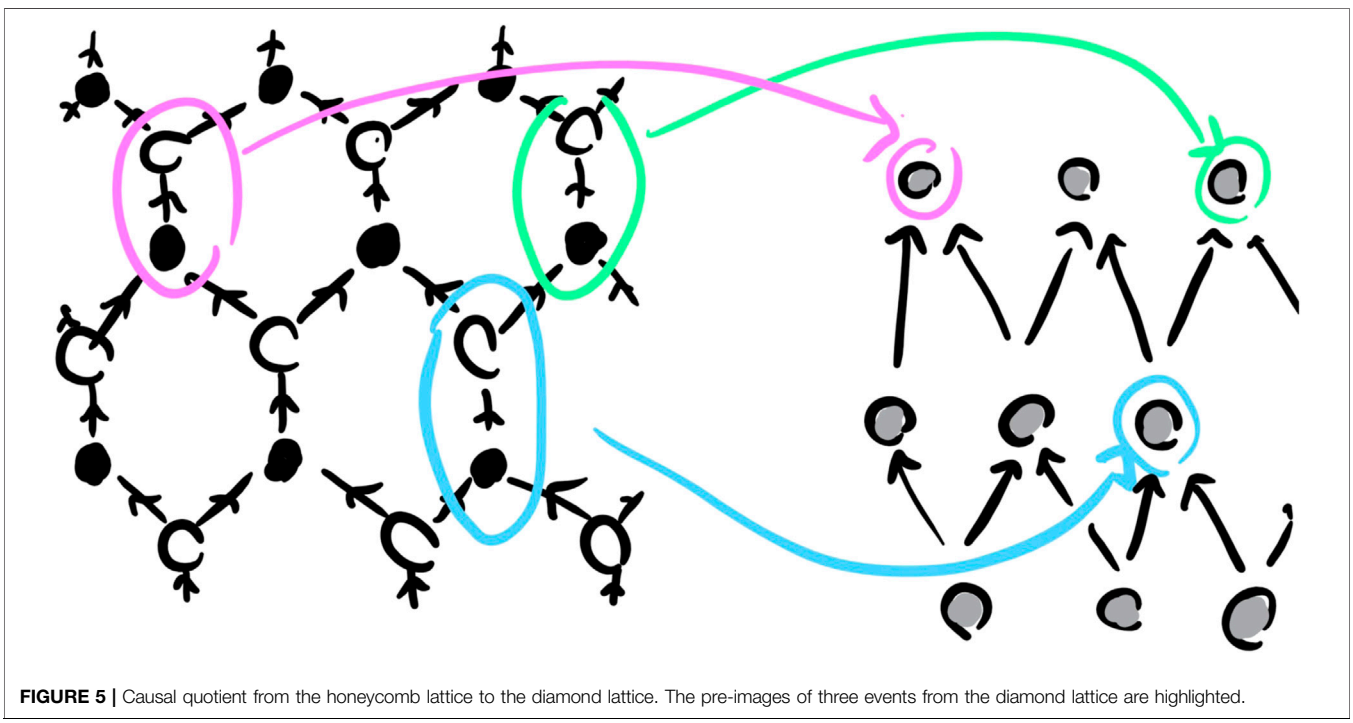

FIGURE 5 | Causal quotient from the honeycomb lattice to the diamond lattice. The pre-images of three events from the diamond lattice are highlighted.

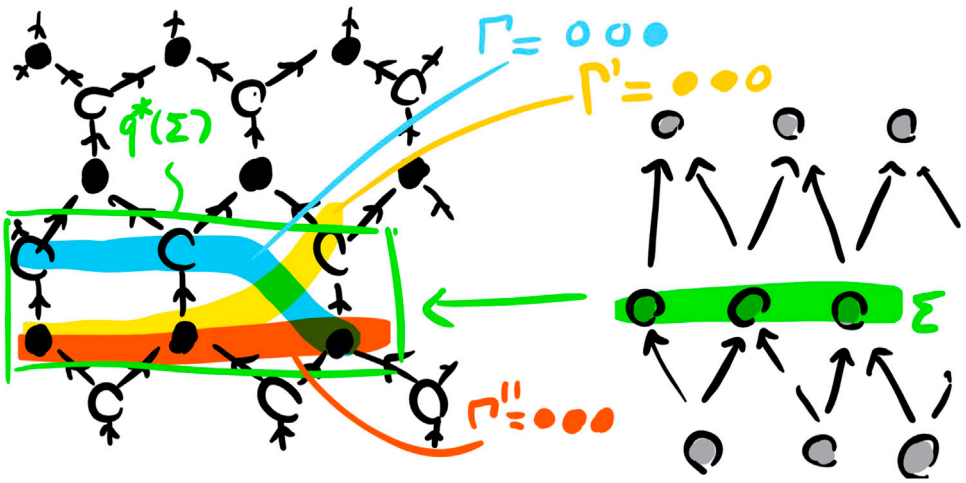

FIGURE 6 | A slice Σ on the diamond lattice and three maximal slices Γ, Γ′ and Γ′′ in its pullback $q^*(\Sigma)$ on the honeycomb lattice.

that $Slices(f^*(\Sigma)) = f^*(Slices(\Sigma))$ *for any slice* Σ *of* Ω *(which can equivalently be seen as a causal sub-order* $\Sigma \hookrightarrow \Omega$*).*

4 CASUAL FIELD THEORIES

In the previous Section, we have defined several commonplace notions from Relativity in the more abstract context of causal orders. In this Section, we endow our causal order with fields, living in an appropriate symmetric monoidal category.

4.1 Categories for Quantum Fields

Depending on the specific applications, there are many symmetric monoidal categories available to model quantum fields.

- If the context is finite-dimensional, quantum fields can be taken to live in the category CPM[fHilb] of finite-dimensional Hilbert spaces and completely positive maps between them.
- If the context is finite-dimensional and super-selected systems are of interest, quantum fields can be taken to live in the category CP*[fHilb] ≅ fC*alg of finite-dimensional C*-algebras and completely positive maps between them.
- If the context is finite-dimensional, an even richer playground available for quantum fields is the category Split[CPM[fHilb]]: this is the Karoubi envelope of the category CPM[fHilb], containing CP*[fHilb] and a number of other systems of operational interest (such as fixed-state systems and constrained systems, see e.g. [23]).
- If the context is infinite-dimensional, e.g. in the case of AQFT [10, 11], the categories usually considered for quantum fields are the category Hilb of Hilbert spaces and bounded linear maps, the category C*alg of C*-algebras and its subcategories W*alg of W*-algebras (sometimes known as "abstract" von Neumann algebras) and vNA of (concrete) von Neumann algebras.
- The categories Hilb, C*alg, W*alg and vNA have some annoying limitations, so in an infinite-dimensional context one can alternatively work with hyperfinite quantum systems [24], which incorporate infinities and infinitesimals to offer additional features—such as duals, traces and unital Frobenius algebras—over plain Hilbert spaces and C*-algebras.

The framework we present here is agnostic to the specific choice of process theory (aka symmetric monoidal category) for quantum fields. In fact, it is agnostic to the specific physical theory considered for the fields: any causal process theory can be considered.

4.2 Causal Field Theories

Let Ω *be a causal order. A causal field theory* Ψ *on* Ω *is a monoidal functor* $\Psi : \mathcal{C} \to \mathcal{D}$ *from a category* $\mathcal{C}$ *of slices on* Ω *to some symmetric monoidal category* $\mathcal{D}$*, which we refer to as the field category.*

Remark 24: *It may sometimes be desirable to add a requirement of injectivity on objects for the functor* Ψ*. This has two main motivations, one of physical character and one of mathematical character. Physically, injectivity means that the field spaces corresponding to distinct events have distinct identities (although they can be isomorphic). Mathematically, injectivity means that the image of the functor is itself a sub-category of* $\mathcal{D}$*, matching the style used by other works on compositional causality [8, 19, 20, 23]. While we do not require this as part of our definition, we will take care for the constructions hereafter to be sufficiently general to accommodate the possibility that such a requirement be imposed.*

We now ask ourselves: what physical information does the functor Ψ encode? On objects, Ψ associates each space-like slice Σ to the space $\Psi(\Sigma)$ of fields over that slice: every point in $\Psi(\Sigma)$ is a valid initial condition for field evolution in the future domain of dependence for Σ.

Remark 25: *If* Σ *is finite and the singleton slices* $\{x\}$ *for the individual events* $x \in \Sigma$ *are all in the chosen category* $\mathcal{C}$ *of slices,*

then the action of Ψ on Σ always factorizes into the tensor product of its action on the individual events:

$$\Psi(\Sigma) = \bigotimes_{x \in \Sigma} \Psi(\{x\}) \tag{13}$$

On morphisms, Ψ associates $\Sigma \twoheadrightarrow \Gamma$ to a morphism $\Psi(\Sigma) \to \Psi(\Gamma)$: this is a specification of how the field evolves from Σ to Γ, i.e. this defines the map sending a field state $|\phi\rangle$ over the initial slice $\Psi(\Sigma)$ to the evolved field state $\Psi(\Sigma \twoheadrightarrow \Gamma)|\phi\rangle$ over the final slice $\Psi(\Gamma)$. This identification of functorial action with field evolution is the core idea of our work. In particular, it explains our specific definition of morphisms in $\mathrm{Slices}(\Omega)$, and hence in all categories of slices: $\Sigma \twoheadrightarrow \Gamma$ if and only if the field data on Σ is sufficient to derive the field data on Γ, assuming causal field evolution. Monoidality of the functor on objects says that the space of fields on the union of disjoint slices is the monoidal product—the tensor product, when working in the familiar linear settings of Hilbert spaces, C*-algebras, von Neumann algebras, etc.—of the spaces of fields on the individual slices. Note that this requirement is stronger than the requirement imposed by AQFT, where field algebras over space-like separated diamonds are only required to commute as sub-algebras of the global field algebra, not necessarily to take the form of a tensor product sub-algebra.

Functoriality and monoidality on morphisms have some interesting consequences, which we now discuss in detail. Let $\Sigma \twoheadrightarrow \Sigma'$ and $\Gamma \twoheadrightarrow \Gamma'$ for a pair of space-like separated slices Σ and Γ and another pair of space-like separated slices Σ' and Γ'. Consider the field evolution between the two disjoint unions of slices:

$$\Psi\big((\Sigma \otimes \Gamma) \twoheadrightarrow (\Sigma' \otimes \Gamma')\big) : \Psi(\Sigma) \otimes \Psi(\Gamma) \to \Psi(\Sigma') \otimes \Psi(\Gamma') \tag{14}$$

Monoidality on morphisms implies that the field evolution above factors as the product of the individual field evolutions $\Psi(\Sigma) \to \Psi(\Sigma')$ and $\Psi(\Gamma) \to \Psi(\Gamma')$:

$$\Psi\big((\Sigma \otimes \Gamma) \twoheadrightarrow (\Sigma' \otimes \Gamma')\big) = \Psi(\Sigma \twoheadrightarrow \Sigma') \otimes \Psi(\Gamma \twoheadrightarrow \Gamma') \tag{15}$$

This may look surprising at first, but it becomes entirely natural upon observing the following.

Proposition 26: *Let Ω be a causal order. If Σ and Γ are space-like separated slices in Ω and $\Sigma \twoheadrightarrow \Sigma'$, then Σ' and Γ are also space-like separated slices.*

Proof: If Σ and Γ are space-like separated, then $\Gamma \cap (J^+(\Sigma) \cup J^-(\Sigma)) = \phi$. Because $\Sigma \twoheadrightarrow \Sigma'$, furthermore, **Proposition 9** tells us that $J^+(\Sigma') \cup J^-(\Sigma') \subseteq J^+(\Sigma) \cup J^-(\Sigma)$. We conclude that $\Gamma \cap (J^+(\Sigma') \cup J^-(\Sigma')) = \emptyset$, i.e. that Σ' and Γ are also space-like separated. □

Proposition 26 above tells us that in our factorization scenario the entire region between Σ and Σ' on one side and the entire region between Γ and Γ' on the other side are space-like separated. Thus any causal field evolution from $\Sigma \otimes \Gamma$ to $\Sigma' \otimes \Gamma'$ would physically be expected to factor: this can be seen as a manifestation of the principle of locality for field theories, sometimes also known as "clustering."

Remark 27: *Please note that the principle of locality obtained above only implies that the evolution of fields must factorize over space-like separated regions. This imposes no constraints on the field state, which can be any state of the space of fields. In particular, if the field category has entanglement (e.g., categories of Hilbert spaces with the usual tensor product) then the field state can entangle space-like separated regions, while field evolution cannot.*

4.3 Causality and No-Signalling

Because any category $\mathcal{C}$ of slices on a causal order Ω is a partially monoidal subcategory of $\mathrm{Slices}(\Omega)$, in particular it necessarily contains the empty slice (the monoidal unit). We define the following family of effects, indexed by all slices $\Sigma \in \mathrm{obj}(\mathcal{C})$:

$$\discard_\Sigma := \Psi(\Sigma \twoheadrightarrow \emptyset) \tag{16}$$

By monoidality we have that $\discard_\emptyset$ is some scalar in the field category $\mathcal{D}$ and that the family respects the partial monoidal structure:

$$\discard_{\Sigma\otimes\Gamma} = \Psi\big((\Sigma\otimes\Gamma) \twoheadrightarrow \emptyset\big) = \Psi\big((\Sigma \twoheadrightarrow \emptyset)\otimes(\Gamma \twoheadrightarrow \emptyset)\big) = \Psi(\Sigma \twoheadrightarrow \emptyset)\otimes\Psi(\Gamma \twoheadrightarrow \emptyset) = \discard_\Sigma \otimes \discard_\Gamma \tag{17}$$

By functoriality, furthermore, the family of effects above is respected by the image of the functor:

$$\discard_\Gamma \circ \Psi(\Sigma \twoheadrightarrow \Gamma) = \Psi(\Gamma \twoheadrightarrow \emptyset) \circ \Psi(\Sigma \twoheadrightarrow \Gamma) = \Psi(\Sigma \twoheadrightarrow \emptyset) = \discard_\Sigma \tag{18}$$

This means that the family of effects $(\discard_\Sigma)_{\Sigma \in \mathrm{obj}(\mathcal{C})}$ defined above is an environment structure and that—as long as injectivity of Ψ is imposed—the image of the causal field theory Ψ is a causal category [19, 20].[6] Physically, this means that the field evolution happens in a no-signalling way: if the effects $(\discard_\Sigma)_{\Sigma \in \mathrm{obj}(\mathcal{C})}$ are used as discarding maps—generalizing the partial traces of quantum theory—then the field state over a given slice Σ does not depend on the field state over slices which are in the future of Σ or are space-like separated from Σ.

This emergence of causality and no-signalling from functoriality is in fact a consequence of a breaking of time symmetry which happened in the very definition of the ordering between slices. Indeed, consider the "time-reversed" causal order Ω^{rev}, obtained by reversing all causal relations in Ω (i.e. $y \leq x$ in Ω^{rev} if and only if $x \leq y$ in Ω). The slices for Ω^{rev} are exactly the slices for Ω, i.e. the categories of all slices $\mathrm{Slices}(\Omega^{rev})$ and $\mathrm{Slices}(\Omega)$ have the same objects. If time symmetry were to hold, we would expect the arrows in $\mathrm{Slices}(\Omega^{rev})$ to be exactly the reverse of the arrows in $\mathrm{Slices}(\Omega)$. However, the conditions defining the arrows in both categories are as follows:

- $\Sigma \twoheadrightarrow \Gamma$ in $\mathrm{Slices}(\Omega)$ iff $\Gamma \subseteq D^+(\Sigma)$ in Ω;
- $\Gamma \twoheadrightarrow \Sigma$ in $\mathrm{Slices}(\Omega^{rev})$ iff $\Sigma \subseteq D^+(\Gamma)$ in Ω^{rev}, i.e. iff $\Sigma \subseteq D^-(\Gamma)$ in Ω.

The two conditions that $\Gamma \subseteq D^+(\Sigma)$ and $\Sigma \subseteq D^-(\Gamma)$, both in Ω, are not in general equivalent: this shows that time symmetry is

[6] We have taken the liberty to extend the definition of environment structures to partially monoidal categories, such as the image of a Ψ injective on objects under the partial monoidal product induced by the partial monoidal product of the domain category $\mathcal{C}$.

broken by our definition of the relationship between slices, ultimately leading to the emergence of causality and no-signalling constraints on functorial evolution of quantum fields.

5 CONNECTION WITH ALGEBRAIC QUANTUM FIELD THEORY

The definition of causal field theories looks somewhat similar to that of Topological Quantum Field Theories (TQFTs) as functors from categories of cobordism to categories of vector spaces. The big difference between the causal field theories we defined above and TQFTs is that the latter take the basic building blocks for field theories to be defined over arbitrary topological spacetimes, while the former define the evolution over a single given spacetime. This difference is an aspect of a general abstract duality between compositionality and decompositionality.

In compositionality, larger objects are created by composing together given elementary building blocks in all possible ways: this is the approach behind an ever growing zoo of process theories (e.g., see [25] and references therein). In decompositionality, on the other hand, larger objects are given as a whole and subsequently decomposed into smaller constituents, with composition of the latter constrained by the context in which they live: this approach, based on partially monoidal structure, was recently introduced by [22] as a way to talk about compositionality in physical theories where a universe is fixed beforehand. While TQFTs are compositional [13, 26], causal field theories are more naturally understood from the decompositional perspective.

In fact, decompositionality is the key ingredient in a completely different family of approaches to quantum theory, including Algebraic Quantum Field Theory (AQFT) [10] and the topos-theoretic approaches [11, 12]. In AQFT, the relationship between fields and the topology of spacetime is encapsulated into the structure of a presheaf, having as its domain the poset formed by causal diamonds in Minkowski space under inclusion and as its codomain a category of C*-algebras and *-homomorphisms. Specifically, each region (causal diamond) of Minkowski spacetime is mapped to the C*-algebra of "local" quantum observables (categorically: effects) on that region. From this perspective, locality and causality are formulated as the requirement that algebras of local observables over space-like separated regions commute within the algebra of global observables (that is, local effects cannot be entangling over space-like separated regions).

To understand the decompositional character of causal field theories, we draw inspiration from the AQFT approach and turn our functors, defined on slices, into presheafs defined on "regions" (generalising unions of causal diamonds in AQFT). However, our approach differs from the AQFT approach in a number of ways:

- We dispense of the algebras themselves: as mentioned earlier in **Section 4.1**, our approach is independent of the specific process theory chosen for the fields.
- Instead of looking at the space of local observables/effects, we take the (equivalent) dual perspective and work with the space of local states.
- Local states can be entangling, so the formulation of locality and causality as "commutativity" is no longer applicable, even in the case where the field category is a category of C*-algebras. Instead, locality and causality arise as a consequence of factorization of field evolution over space-like separated slices.

We begin by showing that categories of slices can be restricted to regions, as long as we take care to define regions in such a way as to respect the restrictions imposed by a specific choice of category of slices.

Definition 28: *A bounded region in a category of slices $\mathcal{C}$ on a causal order Ω is a region on Ω in the form $\Diamond_{\Sigma,\Gamma}$ for some $\Sigma, \Gamma \in obj(\mathcal{C})$. Bounded regions in $\mathcal{C}$ form a poset $Regions_{bnd}(\mathcal{C})$ under inclusion.*

Definition 29: *A region in a category of slices $\mathcal{C}$ is a region* R *on Ω which can be obtained as a union $R = \cup_{\lambda\in\Lambda}\Diamond_{\Sigma_\lambda,\Gamma_\lambda}$ of a family $(\Diamond_{\Sigma_\lambda,\Gamma_\lambda})_{\lambda\in\Lambda}$, closed under finite unions, of bounded regions in $\mathcal{C}$. Regions in $\mathcal{C}$ also form a poset $Regions(\mathcal{C})$ under inclusion, with $Regions_{bnd}(\mathcal{C})$ as a sub-poset.*

Note that if $\mathcal{C} = \text{Slices}(\Omega)$ then the regions in $\mathcal{C}$ are exactly the regions on Ω: by definition, a region R on Ω is the union of the bounded regions $\Diamond_{x,y}$ for all $x, y \in R$.

Proposition 30: *Let $\mathcal{C}$ be a category of slices and* R *be a region in it. The restriction $\mathcal{C}|_R$ of $\mathcal{C}$ to the region* R*, defined as the full sub-category of $\mathcal{C}$ spanned by the slices $\Delta \in obj(\mathcal{C})$ such that $\Delta \subseteq R$, is itself a category of slices.*

Proof: If $R = \Diamond_{\Sigma,\Gamma}$ is a bounded region in $\mathcal{C}$, then the statement is an immediate consequence of requirement (2) for categories of slices. Now assume that $R = \cup_{\lambda\in\Lambda}\Diamond_{\Sigma_\lambda,\Gamma_\lambda}$ is a union of bounded regions in $\mathcal{C}$.

If $x \leq y$ are two events in R, then it must be that $x \in \Diamond_{\Sigma_{\lambda_x},\Gamma_{\lambda_x}}$ and $y \in \Diamond_{\Sigma_{\lambda_y},\Gamma_{\lambda_y}}$ for some $\lambda_x, \lambda_y \in \Lambda$: closure under union of the family $(\Diamond_{\Sigma_\lambda,\Gamma_\lambda})_{\lambda\in\Lambda}$ then guarantees that there exists some $\lambda_{x,y} \in \Lambda$ with $x, y \in \Diamond_{\Sigma_{\lambda_{x,y}},\Gamma_{\lambda_{x,y}}}$. Because $\mathcal{C}$ is a category of slices, we can find two slices $\Delta_x \twoheadrightarrow \Delta_y$ in $\mathcal{C}$ such that $x \in \Delta_x$ and $y \in \Delta_y$. Then the restrictions $(\Delta_x \cap \Diamond_{\Sigma_{\lambda_{x,y}},\Gamma_{\lambda_{x,y}}})$ and $(\Delta_y \cap \Diamond_{\Sigma_{\lambda_{x,y}},\Gamma_{\lambda_{x,y}}})$ satisfy requirement (1) for $\mathcal{C}|_R$ to be a category of slices.

If Σ, Γ and Δ are three slices in $\mathcal{C}|_R$, then in particular the diamond $\Diamond_{\Sigma,\Gamma}$ is a subset of R (the latter is a region) and so is the intersection $\Delta \cap \Diamond_{\Sigma,\Gamma}$, which exists in $\mathcal{C}$ because the latter is a category of slices. Hence requirement (2) for $\mathcal{C}|_R$ to be a category of slices is satisfied.

Requirement (3) for $\mathcal{C}|_R$ to be a category of slices is satisfied, because if $\Sigma, \Gamma \subseteq R$ then also $\Sigma \otimes \Gamma \subseteq R$ whenever the latter is defined. □

Given a causal field theory $\Psi : \mathcal{C} \to \mathcal{D}$, the restrictions $\Psi|_R : \mathcal{C}|_R \to \mathcal{D}$ are again causal field theories. To match the spirit of AQFT, we need two more ingredients: the definition of a space of states $\text{States}_\Psi(R)$ over a region R and the definition of restrictions $\text{States}_\Psi(R) \to \text{States}_\Psi(R')$ between spaces of states associated with inclusions $R' \subseteq R$ of regions.

Definition 31: *Given a region* R *in a category of slices $\mathcal{C}$, the space of states $\text{States}_\Psi(R)$ over the region is defined to be the set*

comprising all families ρ of states over the slices in $\mathcal{C}|_R$ which are stable under the action of Ψ, i.e. comprising all the families

$$\rho \in \prod_{\Delta \in \mathrm{obj}(\mathcal{C}|_R)} \mathrm{States}_{\mathcal{D}}(\Psi(\Delta)), \tag{19}$$

such that for all $\Delta, \Delta' \in obj(\mathcal{C}|_R)$ with $\Delta \twoheadrightarrow \Delta'$ the following condition is satisfied:

$$\Psi\left(\Delta \twoheadrightarrow \Delta'\right) \circ \rho_{\Delta} = \rho_{\Delta'}. \tag{20}$$

By $\mathrm{States}_{\mathcal{D}}(\Psi(\Delta))$ we have denoted the states on the object $\Psi(\Delta)$ of the symmetric monoidal category $\mathcal{D}$, i.e. the homset $\mathrm{Hom}_{\mathcal{D}}[I, \Psi(\Delta)]$ where I is the monoidal unit of $\mathcal{D}$.

Proposition 32: *Given a causal field theory $\Psi : \mathcal{C} \to \mathcal{D}$, we can construct a presheaf $\mathrm{States}_\Psi : \mathrm{Regions}(\mathcal{C})^{op} \to \mathrm{Set}$ by associating each region $R \in \mathrm{obj}(\mathrm{Regions}(\mathcal{C}))$ to the space of states $\mathrm{States}_\Psi(R)$ over the region, and each inclusion $i : R' \subseteq R$ to the restriction function $\mathrm{States}_\Psi(R) \to \mathrm{States}_\Psi(R')$ defined by sending a family $\rho \in \mathrm{States}_\Psi(R)$ to the family $\mathrm{States}_\Psi(i)(\rho) \in \mathrm{States}_\Psi(R')$ given as follows:*

$$\mathrm{States}_\Psi(i)(\rho)_{\Delta'} = \rho_{i(\Delta')}. \tag{21}$$

We refer to States_Ψ as the presheaf of states over regions of $\mathcal{C}$.

Proof: The only thing to show is functoriality of States_Ψ. If $i = id_R : R \subseteq R$ is the identity on a region R, then we have:

$$\mathrm{States}_\Psi(i)(\rho)_{\Delta} = \rho_{i(\Delta)} = \rho_{\Delta}, \tag{22}$$

i.e. $\mathrm{States}_\Psi(i) = id_{\mathrm{States}_\Psi(R)}$ is the identity on the space of states over the region. If now $j : R'' \subseteq R'$ and $i : R' \subseteq R$, then $i \circ j : R'' \subseteq R$ and we have:

$$\begin{aligned} \mathrm{States}_\Psi(j)\left(\mathrm{States}_\Psi(i)(\rho)\right)_{\Delta''} &= \mathrm{States}_\Psi(i)(\rho)_{j(\Delta'')} = \rho_{i(j(\Delta''))} \\ &= \mathrm{States}_\Psi(i \circ j)(\rho)_{\Delta''}. \end{aligned} \tag{23}$$

Hence States_Ψ is a presheaf $\mathrm{States}_\Psi : \mathrm{Regions}(\mathcal{C})^{op} \to \mathrm{Set}$. □

Definition 33: *A global state ρ for a causal field theory $\Psi : \mathcal{C} \to \mathcal{D}$ is a global compatible family for $States_\Psi$, i.e. a family $\rho = (\rho^{(R)})_{R \in Regions(\mathcal{C})}$ such that $States_\Psi(i)(\rho^{(R)}) = \rho^{(R')}$ for all inclusions $i : R' \subseteq R$ in $Regions(\mathcal{C})$. We refer to the set of all global states as the space of global states.*

Remark 34: *If Ω is a region in C, i.e. if $\Omega \in Regions(\mathcal{C})$, then the states in $States_\Psi(\Omega)$ (Ω as a region) are in bijection with the global states as follows:*

$$\begin{cases} \rho \in \mathrm{States}_\Psi(\Omega) \mapsto \left(\mathrm{States}_\Psi(R \subseteq \Omega)(\rho)\right)_{R \in \mathrm{Regions}(\mathcal{C})} \\ \left(\rho^{(R)}\right)_{R \in \mathrm{Regions}(\mathcal{C})} \mapsto \rho^{(\Omega)} \end{cases}. \tag{24}$$

Because of this, we consistently adopt the notation $States_\Psi(\Omega)$ to denote the space of global states. If R *is a region in C, we also adopt the notation $States_\Psi(R \subseteq \Omega)$ for the map $States_\Psi(\Omega) \to States_\Psi(R)$ sending a global state $\rho = (\rho^{(R)})_{R \in Regions(\mathcal{C})}$ to its component $\rho^{(R)}$ over the region* R*. To unify notation, we will also adopt ρ_Σ to denote $(\rho^{(R)})_\Sigma$, taking the same value for any region* R *containing* Σ (e.g., for $R = \Sigma$).

Remark 35: *If the field category $\mathcal{D}$ is suitably enriched (e.g., in a category with all limits), then a natural choice is for the the space of states to be defined by a presheaf valued in the enrichment category. For example, quantum theory is enriched over positive cones, i.e. $\mathbb{R}^+$-modules, and the $\mathbb{R}^+$-linear structure of states in quantum theory extends to a $\mathbb{R}^+$-linear structure on the spaces of states of causal field theories having quantum theory as their field category. We will not consider such enrichment in this work, though all constructions we present can be readily extended to such a setting.*

Spaces of states according to **Definition 31** encode a lot of redundant information, because we don't want to look into the specific structure of regions. However, there are certain special cases in which an equivalent description of the space of states over a region can be given.

To start with, consider two slices $\Sigma \twoheadrightarrow \Gamma$ and note that the state on any slice $\Delta \subseteq \Diamond_{\Sigma,\Gamma}$ in a bounded region $\Diamond_{\Sigma,\Gamma}$ is uniquely determined by applying $\Psi(\Sigma \twoheadrightarrow \Delta)$ to the state on Σ:

$$\rho_\Delta = \Psi(\Sigma \twoheadrightarrow \Delta)(\rho_\Sigma). \tag{25}$$

This is, for example, the case for all bounded regions between Cauchy slices in a category of slices $\mathrm{CauchySlices}(\mathcal{F})$ generated by some foliation $\mathcal{F}$. If the foliation $\mathcal{F}$ has a minimum Σ_0—an initial Cauchy slice—then any global state $\rho \in \mathrm{States}_\Psi(\Omega)$ is entirely determined by its component ρ_{Σ_0} over the initial slice Σ_0:

$$\rho_\Delta = \Psi(\Sigma_0 \twoheadrightarrow \Delta) \circ \rho_{\Sigma_0}, \tag{26}$$

for any $\Delta \in \mathcal{F}$ and any region R in $\mathrm{CauchySlices}(\mathcal{F})$ such that $\Delta \subseteq R$. This extends to all slices in $\mathrm{CauchySlices}(\mathcal{F})$ by restriction.

Inspired by Relativity, we would like the state on any Cauchy slice in the foliation to determine the global state, not only that on an initial Cauchy slice (which may not exist). For this to happen, we need to strengthen our requirements on the causal field theory, which needs to be causally reversible.

Definition 36: *Let Ω be any causal order. By the causal reverse of Ω we mean the causal order Ω^{rev} on the same events as Ω and such that $x \leq y$ in Ω^{rev} if and only if $x \geq y$ in Ω.*

Definition 37: *A category of slices $\mathcal{C}$ on a causal order Ω is said to be causally reversible if the full sub-category of $Slices(\Omega^{rev})$ spanned by $obj(\mathcal{C})$ is a category of slices on the causal reverse Ω^{rev}. If this is the case, we write $\mathcal{C}^{rev}$ for said category of slices over Ω^{rev} and refer to it as the causal reverse of $\mathcal{C}$. We write $\overset{rev}{\twoheadrightarrow}$ for the morphisms of $\mathcal{C}^{rev}$.*

Definition 38: *Let $\Psi : \mathcal{C} \to \mathcal{D}$ be a causal field theory on a causal order Ω. If $\mathcal{C}$ is causally reversible, a causal reversal of Ψ is a causal field theory $\Phi : \mathcal{C}^{rev} \to \mathcal{D}$ such that:*

(1) *The functors Ψ and Φ agree on objects, i.e. for all $\Sigma \in obj(\mathcal{C})$ we have that $\Psi(\Sigma) = \Phi(\Sigma)$;*
(2) *Whenever we have two chains of alternating morphisms in $\mathcal{C}$ and $\mathcal{C}^{rev}$ which start and end at the same slices Σ, Γ, say in the form*

$$\begin{aligned}&\Sigma \twoheadrightarrow \Delta_1 \overset{rev}{\twoheadrightarrow} \Delta_2 \twoheadrightarrow \ldots \Delta_{2n} \twoheadrightarrow \Gamma,\\ &\Sigma \twoheadrightarrow \Delta'_1 \overset{rev}{\twoheadrightarrow} \Delta'_2 \twoheadrightarrow \ldots \Delta'_{3m} \twoheadrightarrow \Gamma,\end{aligned} \tag{27}$$

for some $n, m \geq 0$, the composition of the images of the morphisms under Ψ and Φ always yield the same morphism $\Psi(\Sigma) \to \Psi(\Gamma)$:

$$\begin{aligned}&\Psi(\Delta_{2n} \twoheadrightarrow \Gamma) \circ \ldots \circ \Phi\left(\Delta_1 \overset{rev}{\twoheadrightarrow} \Delta_2\right) \circ \Psi(\Sigma \twoheadrightarrow \Delta_1)\\ &= \Psi\left(\Delta'_{2m} \twoheadrightarrow \Gamma\right) \circ \ldots \circ \Phi\left(\Delta'_1 \overset{rev}{\twoheadrightarrow} \Delta'_2\right) \circ \Psi\left(\Sigma \twoheadrightarrow \Delta'_1\right).\end{aligned} \tag{28}$$

We say that $\Psi : \mathcal{C} \to \mathcal{D}$ is causally reversible—or simply reversible—if $\mathcal{C}$ is causally reversible and Ψ admits a causal reversal.

Proposition 39: *Let* $\mathrm{CauchySlices}(\mathcal{F})$ *be the category of slices on a causal order Ω generated by some foliation $\mathcal{F}$. Then* $\mathrm{CauchySlices}(\mathcal{F})$ *is always causally reversible and for any two* Cauchy slices Δ, Σ *we have that $\Delta \twoheadrightarrow \Sigma$ if and only if $\Sigma \overset{rev}{\twoheadrightarrow} \Delta$. Furthermore, if a causal field theory $\Psi : \mathrm{CauchySlices}(\mathcal{F}) \to \mathcal{D}$ is reversible, then a global state ρ is entirely determined by the state ρ_Σ on any* Cauchy slice $\Sigma \in \mathcal{F}$ *as follows:*

$$\rho_\Delta = \begin{cases} \Psi(\Sigma \twoheadrightarrow \Delta) \circ \rho_\Sigma & \text{if } \Sigma \twoheadrightarrow \Delta \\ \Phi\left(\Sigma \overset{rev}{\twoheadrightarrow} \Delta\right) \circ \rho_\Sigma & \text{if } \Delta \twoheadrightarrow \Sigma \end{cases}, \tag{29}$$

where $\Phi : \mathrm{CauchySlices}(\mathcal{F})^{rev} \to \mathcal{D}$ is any causal reversal of Ψ.

Proof: The main observation behind this result is as follows: if Σ, Δ are two Cauchy slices, then the conditions $\Delta \subseteq D^+(\Sigma)$ and $\Sigma \subseteq D^-(\Delta)$ are equivalent. Hence $\mathrm{CauchySlices}(\mathcal{F})$ is always causally reversible and $\Delta \twoheadrightarrow \Sigma$ if and only if $\Sigma \overset{rev}{\twoheadrightarrow} \Delta$ for any two Cauchy slices Δ, Σ.

Now let Ψ be causally reversible, let $\Sigma \in \mathcal{F}$ be a Cauchy slice in the foliation and consider any global state ρ. If $\Sigma \twoheadrightarrow \Delta$ for some other Cauchy slice $\Delta \in \mathcal{F}$, then the definition of a global state implies that $\rho_\Delta = \Psi(\Sigma \twoheadrightarrow \Delta) \circ \rho_\Sigma$. If instead $\Delta \twoheadrightarrow \Sigma$, then $\Sigma \overset{rev}{\twoheadrightarrow} \Delta$ and the definition of a global state implies that $\rho_\Sigma = \Psi(\Delta \twoheadrightarrow \Sigma) \circ \rho_\Delta$. But the definition of a causal reverse also implies that:

$$\begin{aligned}\Phi\left(\Sigma \overset{rev}{\twoheadrightarrow} \Delta\right) \circ \rho_\Sigma &= \Phi\left(\Sigma \overset{rev}{\twoheadrightarrow} \Delta\right) \circ \Psi(\Delta \twoheadrightarrow \Sigma) \circ \rho_\Delta = \Psi(\Sigma \twoheadrightarrow \Sigma) \circ \rho_\Delta\\ &= id_{\Psi(\Sigma)} \circ \rho_\Delta = \rho_\Delta.\end{aligned} \tag{30}$$

Hence the value ρ_Σ completely determines the global state ρ (since the value on all other slices in $\mathrm{CauchySlices}(\mathcal{F})$ is determined by restriction from the value on a corresponding Cauchy slice). □

It is an easy check that not only the global states $\rho \in \mathrm{States}_\Psi(\Omega)$ are determined—under the conditions of **Proposition 39**—by their component $\rho_\Sigma \in \mathrm{States}_\mathcal{D}(\Psi(\Sigma))$ over any Cauchy slice Σ in the foliation, but also that **Eq. 29** can be used—under the same conditions—to construct a global state $\rho \in \mathrm{States}_\Psi(\Omega)$ from a state $\rho_\Sigma \in \mathrm{States}_\mathcal{D}(\Psi(\Sigma))$ on any Cauchy slice Σ in the foliation.

Before concluding this Section, we would like to remark that a succinct description of spaces of states over regions can be obtained in settings much more general than those of foliations: for example, in all those cases where the every region admits a suitable Cauchy slice and the causal field theory is reversible. The careful formulation of this more general setting is key to the further development of the connection between causal field theory and AQFT and it is left to future work.

6 CONNECTION TO QUANTUM CELLULAR AUTOMATA

The idea of a cellular automaton was first introduced by von Neumann, aimed at designing a self replicating machine [18]. A Cellular Automaton (CA) over some finite alphabet A has its state stored as a d-dimensional lattice of values in A, i.e. as a function $\psi : \mathbb{Z}^d \to A$. The state is updated at discrete time steps, each step updated as $\psi^{(t+1)} := F(\psi^{(t)})$ according to some fixed function $F : (\mathbb{Z}^d \to A) \to (\mathbb{Z}^d \to A)$. The function F acts locally and homogeneously: there is some fixed finite subset $\mathcal{N} \subset \mathbb{Z}^d$ (typically a neighborhood of $\underline{0} \in \mathbb{Z}^d$) and some function $f : \mathcal{N} \to A$ such that the value of each lattice site $\underline{x}$ at time step $t + 1$ only depends on the finitely many values in the subset $\underline{x} + \mathcal{N}$ at time t:

$$F(\psi) := \underline{x} \mapsto f\left(\psi\big|_{\underline{x}+\mathcal{N}}\right). \tag{31}$$

A Quantum Cellular Automaton (QCA) is a generalization of a CA where the lattice states $\psi : \mathbb{Z}^d \to A$ are replaces by (pure) states in the tensor product of Hilbert spaces $\otimes_{\underline{x} \in \mathbb{Z}^d} \mathcal{H}_x$ (all $\mathcal{H}_{\underline{x}}$ finite-dimensional and isomorphic) and the function F is replaced by a unitary $U : \otimes_{\underline{x} \in \mathbb{Z}^d} \mathcal{H}_{\underline{x}} \to \otimes_{\underline{x} \in \mathbb{Z}^d} \mathcal{H}_{\underline{x}}$, with requirements of locality and homogeneity.

Remark 40: *There are several slightly different formulation of the infinite tensor product above that can be used, each with its own advantages and disadvantages: though it is not going to be a concern for this work, the authors are partial to the construction by von Neumann [27].*

An early formulation of the notion of QCA is due to Richard Feynman, in the context of simulations of physics using quantum computers [28]. More recent work on quantum information and quantum causality has shown that the evolution of certain free quantum fields can be recovered as the continuous limit of certain quantum cellular automata (cf [16, 17]. and references therein). In the final section of this work, we show that our framework is well-suited to capture notions of QCA such as those appearing in the literature. Specifically, our construction encompasses and greatly generalises that presented in [17].

6.1 Causal Cellular Automata

The first requirement in the definition of a QCA is that of homogeneity—called "translation invariance" in [17]—i.e. the requirement that the automaton act the same way at all points of spacetime. Because presentations of QCAs are usually given in terms of discrete updates of states on a lattice by means of a unitary U, only the requirement of homogeneity *in space* is usually mentioned. However, such presentations also have homogeneity in time as an implicit requirement, namely in

the assumption that the same unitary U be used to update the state at all times.

Instead of updating the state time-step by time-step in a compositional fashion, our formulation of quantum cellular automata will see the entirety of spacetime at once, with states over slices and regions recovered in a decompositional approach. Nevertheless, the requirement of homogeneity for a QCA can still be formulated as a requirement of invariance under certain symmetries of spacetime, so we begin by formulating such a notion of invariance for causal field theories.

Definition 41: *A symmetry on a causal order Ω is an action of a group G on Ω by automorphisms of causal orders, i.e. a group homomorphism $G \to \mathrm{Aut}_{CausOrd}[\Omega]$. If $\mathcal{C}$ is a category of slices on Ω, a symmetry on $\mathcal{C}$ is a symmetry on Ω which extends to an action on $\mathcal{C}$ by partially monoidal functors, i.e. one such that the following conditions are satisfied:*

(1) *for all $g \in G$, if $\Sigma \in obj(\mathcal{C})$ then $g(\Sigma) \in obj(\mathcal{C})$;*
(2) *for all $g \in G$ and all $\Sigma, \Gamma \in obj(\mathcal{C})$, if $\Sigma \twoheadrightarrow \Gamma$ then $g(\Sigma) \twoheadrightarrow g(\Gamma)$;*
(3) *for all $g \in G$ and all $\Sigma, \Gamma \in obj(\mathcal{C})$, if $\Sigma \otimes \Gamma$ is defined in $\mathcal{C}$* then *$g(\Sigma \otimes \Gamma) = g(\Sigma) \otimes g(\Gamma)$ is also defined in $\mathcal{C}$.*
Note, for all $g \in G$, that $g(\emptyset) = \emptyset$ and that $g(\Sigma)$ is automatically a slice whenever Σ is a slice.

Definition 42: *Let $\mathcal{C}$ be a category of slices with a symmetry action of a group G. A G-invariant (or simply symmetry-invariant) causal field theory on $\mathcal{C}$ is a causal field theory $\Psi : \mathcal{C} \to \mathcal{D}$ equipped with a family of natural isomorphisms $\alpha_g : \Psi \Rightarrow \Psi \circ g$ such that $\alpha_{h \cdot g} = \alpha_h g \circ \alpha_g$, where we have identified elements $g \in G$ with their action as partially monoidal functors $g : \mathcal{C} \to \mathcal{C}$.*

Remark 43: *The spirit behind the definition of symmetry-invariant causal field theories is that the functors Ψ (sending slices $\mapsto$ fields) and $\Psi \circ g$ (sending slices $\mapsto$ g-translated slices $\mapsto$ fields) should be the same. However, we have remarked when first defining causal field theories that—be it for ease of physical interpretation or for conformity with existing literature on causal categories—it may sometimes be desirable that the images $\Psi(\Sigma)$ of different slices be different. Not being able to impose the equality $\Psi = \Psi \circ g$ in such a setting, the next best thing is to ask for natural isomorphism $\Psi \cong \Psi \circ g$.*

Because we are dealing with symmetries, however, it is sensible to require for the natural isomorphisms α_g themselves to respect the group structure. Again, the first instinct might be to require something in the form $\alpha_{h \cdot g} = \alpha_h \circ \alpha_g$, but this expressions does not type-check: we have a natural transformation $\alpha_{h \cdot g} : \Psi \Rightarrow \Psi \circ h \circ g$, a natural transformation $\alpha_g : \Psi \Rightarrow \Psi \circ g$ and a natural transformation $\alpha_h : \Psi \Rightarrow \Psi \circ h$. In order to compose α_h and α_g we instead have to take the action of α_h translated to $\Psi \circ g$:

$$\alpha_h g : \Psi \circ g \Rightarrow (\Psi \circ h) \circ g. \tag{32}$$

Explicitly, the natural transformation $\alpha_h g$ is defined by $(\alpha_h g)(\Sigma) := \alpha_h(g(\Sigma))$.

The second requirement in the definition of a QCA is that of locality (or causality). When quantum cellular automata are considered in a relativistic context—e.g. as discrete models of quantum field theories—the requirement of locality is meant to capture the idea that the action of the automaton should respect the causal structure of spacetime (so that the state on a point $\underline{x}$ at time $t + \Delta t$ should not depend on the state at the previous time t on points $\underline{y}$ which are "too far away", i.e. such that $(\underline{x}, t + \Delta t)$ and $(\underline{y}, t)$ are space-like separated).

In [17], the requirement of locality is formulated as the requirement that the output state of the automaton over a point $\underline{x}$ of the lattice at time $t + 1$ only depend on the state over a finite neighborhood $\underline{x} + \mathcal{N}$ at time t: this is both in terms of local state (causality) and in the stronger sense that the field evolution should factor into a product of local maps (localisability). In our framework, on the other hand, causality and localisability are both automatically enforced: the field evolution always factors over space-like separated regions, as a consequence of monoidality, and the local state over a slice never depends on the state on any other slice which is space-like separated from it (as a consequence of factorisation).

Remark 44: *The causal order Ω which captures the causality requirement from [17] with finite neighborhood $\mathcal{N} \subset \mathbb{Z}^d$ can be constructed by endowing the set $|\Omega| := \mathbb{Z}^d \times \mathbb{Z}$ with the reflexive-transitive closure of the relation $(\underline{y}, t) \leq (\underline{x}, t + 1)$ for all times $t \in \mathbb{Z}$, for all points of the lattice $\underline{x} \in \mathbb{Z}^d$ and for all points $\underline{y} \in \underline{x} + \mathcal{N}$ in the neighborhood of $\underline{x}$.*

The third and final requirement in the definition of a QCA is that of unitarity. In our framework, this is a problem for two (mostly unrelated) reasons.

- Our formulation of causal field theories aims to be agnostic to the choice of process theory. On the other hand, unitarity is a strongly quantum-like feature, the formulation of which would require a significant amount of additional structure on the field category.
- The usual formulation of quantum cellular automata only considers global evolution, never directly dealing with restrictions—situations e.g., in which the state is evolved unitarily but part of the output state is discarded as environment. Our framework instead treats such restrictions as an integral part of evolution.

Luckily, unitarity per se is not necessary from an abstract foundational standpoint: the real feature of interest is reversibility, a feature of causal field theories which we have already explored. For the sake of generality, we will not include reversibility in the definition below, leaving it as an explicit desideratum.

Remark 45: *In categories of Hilbert spaces and completely positive maps, it is legitimate to imagine that causality and reversibility would jointly imply that the cellular automata also be unitary. This is indeed the case under the conditions of* ***Proposition 39****: because the state on any Cauchy slice automatically determines the state on all the other slices—and because that state on a single slice is arbitrary—evolution between Cauchy slices must be unitary.*

Definition 46: *A Causal Cellular Automaton (CCA) consists of the following ingredients.*

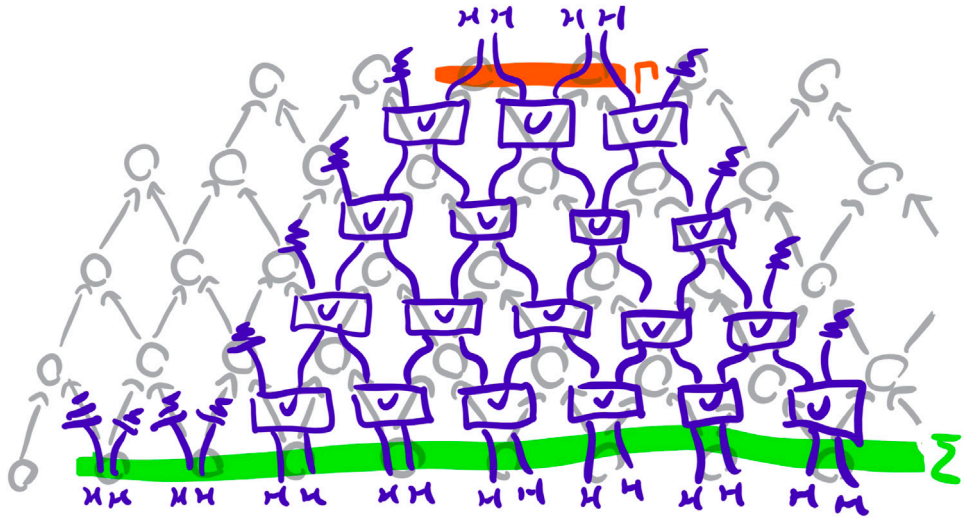

FIGURE 7 | Action of a partitioned causal cellular automaton over a complicated morphism $\Sigma \twoheadrightarrow \Gamma$ in the $(1+1)$-dimensional example of the diamond lattice. Here $\mathcal{N} = \{\pm 1\}$, so each event in the causal order is associated to a copy of $\mathcal{H}^{\otimes \mathcal{N}} \cong \mathcal{H} \otimes \mathcal{H}$. The restriction action of the CCA (**Eq. 45**) can be seen on the two events at the bottom left. The pure evolution action of the CCA (**Eq. 46**) can be seen on the central pyramid of ten events, as the application of U without discarding. The evolution + marginalisation action of the CCA (**Eq. 48**) can be seen on the eight events at the sides of the central pyramid, as the application of U followed by discarding of one of the two outputs. The input of the morphism depicted consists of eight copies of $\mathcal{H} \otimes \mathcal{H}$, one for each event of Σ, while the output of the morphism depicted consists of two copies of $\mathcal{H} \otimes \mathcal{H}$, one for each event of Γ.

(1) *A foliation $\mathcal{F}$ on a causal order Ω.*
(2) *A category of Cauchy slices $\mathcal{C}$ such that each slice in $\mathcal{C}$ is a subset of some Cauchy slice in $\mathcal{F}$.*[7]
(3) *A symmetry action of a group G on $\mathcal{C}$, inducing—via the G-action on Ω—a transitive action of G on the Cauchy slices in the foliation $\mathcal{F}$.*
(4) *A G-invariant causal field theory $\Psi : \mathrm{CauchySlices}(\mathcal{F}) \to \mathcal{D}$.*

A reversible CCA is one where the causal field theory Ψ is reversible.

Definition 46 is much more general than the definition of QCA from [17] and hence captures more sophisticated examples. However, its ingredients are directly analogous to those appearing in that definition of a QCA.

- The foliation $\mathcal{F}$ on Ω generalizes the discrete time steps in the definition of a QCA.
- The slices in $\mathcal{C}$ generalize the equal-time hyper-surfaces which support the state of a QCA at fixed time.
- The symmetry action of G on $\mathrm{CauchySlices}(\mathcal{F})$ and its transitivity on the foliation $\mathcal{F}$ generalise homogeneity in both space and time of the lattices supporting a QCA.
- The G-invariance of the causal field theory Ψ generalizes both the translation symmetry in space and the time-translation symmetry of a QCA.

A different approach to QCAs in non-homogeneous spacetimes appears in [29, 30], in terms of graph dynamics. The graph dynamics models and the models described in this work present a significant overlap—in the specific case of quantum theory—but are ultimately incomparable: on the one end, graph dynamics impose certain structural requirements on spacetime slices for the foliation, requirements which are not necessary in this work; on the other end, quantum graph dynamics allow a superposition of graphs at each slice of the foliation, a possibility which is not considered in this work.

[7]Each Cauchy slice Σ in $\mathcal{F}$ is then automatically the union of all slices $\Delta \in \mathrm{obj}(\mathcal{C})$ such that $\Delta \subseteq \Sigma$.

6.2 Partitioned Causal Cellular Automata

We now proceed to construct a large family of examples of CCAs based on the partitioned QCAs of [17]. In doing so, we generalise the scattering unitaries to arbitrary processes and allow for the definition of state restriction to non-Cauchy equal-time surfaces. We refer to the resulting CCA as partitioned CCA.

6.2.1 Causal Order

As our causal order Ω we consider the following subset of $(1+d)$-dimensional Minkowski spacetime (setting the constant c for the speed of light to $c = \sqrt{d}$):

$$\Omega := \left\{ (t, \underline{x}) \,\middle|\, t \in \mathbb{Z}, \underline{x} \in (t, \ldots, t) + 2\mathbb{Z}^d \right\}. \tag{33}$$

where $(t, \ldots, t) + 2\mathbb{Z}^d$ is the set of all $\underline{x} \in \mathbb{Z}^d$ such that $x_i = t \pmod 2$. For $d = 1$ we get the $(1+1)$-dimensional diamond lattice discussed before. In general, the immediate causal predecessors of a point $(t, \underline{x})$ are the following 2^d points:

$$(t-1, \underline{x} - \mathcal{N}) = \{(t-1, \underline{x} - \underline{\delta}) | \underline{\delta} \in \mathcal{N}\}, \tag{34}$$

where we defined the "neighborhood" $\mathcal{N} := \{\pm 1\}^d$. Similarly, the immediate successors of $(t, \underline{x})$ are the following 2^d points:

$$(t+1, \underline{x}+\mathcal{N}) = \{(t+1, \underline{x}+\underline{\delta}) | \underline{\delta} \in \mathcal{N}\}. \tag{35}$$

6.2.2 Foliation and Category of Slices

The causal order Ω admits a foliation $\mathcal{F}$ where each slice is a constant-time Cauchy slice Σ_t for some $t \in \mathbb{Z}$:

$$\Sigma_t := \left\{ (t, \underline{x}) \,\middle|\, \underline{x} \in (t, \ldots, t) + 2\mathbb{Z}^d \right\}. \tag{36}$$

A suitable category of slices $\mathcal{C}$ to associate to this foliation is given by taking as slices all the finite sets $\Sigma_{t,\mathcal{X}} \subset \Sigma_t$ of events having the same time coordinate t:

$$\Sigma_{t,\mathcal{X}} = \{(t, \underline{x}) | \underline{x} \in \mathcal{X}\}, \tag{37}$$

where $\mathcal{X} \subset (t, \ldots, t) + 2\mathbb{Z}^d$ is some finite subset. The morphisms $\twoheadrightarrow$ of $\mathcal{C}$ are given as follows for $k \geq 0$:

$$\Sigma_{t,\mathcal{X}} \twoheadrightarrow \Sigma_{t+k,\mathcal{Y}} \quad \text{if and only if} \quad \bigcup_{\underline{y} \in \mathcal{Y}} \left(\left(t, \underline{y} + \mathcal{N}^{(k)} \right) \right) \subseteq \mathcal{X}, \tag{38}$$

where the "iterated neighborhood" $\mathcal{N}^{(k)}$ is defined as $\mathcal{N} + \ldots + \mathcal{N}$ by adding together $k \geq 0$ copies of $\mathcal{N}$ (and we set $\mathcal{N}^{(0)} := \{0\}$). Explicitly we have:

$$\mathcal{N}^{(k)} := \begin{cases} \{-k, -k+2, \ldots -1, +1, \ldots, k-2, k\} & \text{if } k \text{ odd} \\ \{-k, -k+2, \ldots -2, 0, +2, \ldots, k-2, k\} & \text{if } k \text{ even} \end{cases}. \tag{39}$$

It is easy to check (by a $t \mapsto -t$ symmetry argument) that $\mathcal{C}$ is reversible.

6.2.3 Symmetry

The category $\mathcal{C}$ admits a symmetry action of the group $G := \mathbb{Z}^{\mathcal{N}} \cong \mathbb{Z}^{2^d}$. We index the coordinates of vectors in $\mathbb{Z}^{\mathcal{N}}$ by the 2^d points $\underline{\delta} \in \mathcal{N} = \{\pm 1\}^d$. We denote by $\tau_{\underline{\delta}}$ the vector in $\mathbb{Z}^{\mathcal{N}}$ which is 1 at the coordinate labeled by $\underline{\delta}$ and 0 at all other coordinates. The action is then specified by setting:

$$\tau_{\underline{\delta}}(t, \underline{x}) := (t+1, \underline{x} - \underline{\delta}), \tag{40}$$

that is, the 2^d generators of $\mathbb{Z}^{\mathcal{N}}$ send a generic event $(t, \underline{x})$ to each of its 2^d immediate causal successors in Ω, one for each possible choice of sign ± 1 along each of the d directions of the space lattice $\mathbb{Z}^d$.[8] Each generator $\tau_{\underline{\delta}}$ for the symmetry action sends a Cauchy slice Σ_t in the foliation to the next Cauchy slice Σ_{t+1}, so the action of G on the foliation is transitive.

6.2.4 Causal Field Theory—Field Over Slices

As our field category we consider a generic causal process theory $\mathcal{D}$, i.e. a symmetric monoidal category equipped with a family of discarding maps $\overset{\blacksquare}{\top}_{\mathcal{H}} : \mathcal{H} \to I$ for all objects $\mathcal{H} \in \text{obj}(\mathcal{D})$, respecting the tensor product $\otimes$ and tensor unit I of $\mathcal{D}$: $\overset{\blacksquare}{\top}_{\mathcal{H}\otimes\mathcal{K}} = \overset{\blacksquare}{\top}_{\mathcal{H}} \otimes \overset{\blacksquare}{\top}_{\mathcal{K}}$ and $\overset{\blacksquare}{\top}_I = 1$. Discarding maps generalize the partial trace of quantum theory: normalized states $\rho : I \to \mathcal{H}$—generalizing density matrices—are defined to be those such that $\overset{\blacksquare}{\top}_{\mathcal{H}} \circ \rho = 1$ and normalized morphisms $U : \mathcal{H} \to \mathcal{K}$—generalizing CPTP maps—are defined to be those such that $\overset{\blacksquare}{\top}_{\mathcal{K}} \circ U = \overset{\blacksquare}{\top}_{\mathcal{H}}$. See e.g. [20, 23, 25] for more information.

To create a G-invariant causal field theory Ψ, we consider some object $\mathcal{H} \in \text{obj}(\mathcal{D})$ together with some endomorphism $U : \mathcal{H}^{\otimes 2^d} \to \mathcal{H}^{\otimes 2^d}$, which we will refer to as the scattering map. For reasons that will soon become clear, it is more convenient to index the factors of $\mathcal{H}^{\otimes 2^d}$ by the 2^d points in the neighborhood $\mathcal{N}$, hence writing $U : \mathcal{H}^{\otimes \mathcal{N}} \to \mathcal{H}^{\otimes \mathcal{N}}$.

We define the action of Ψ on the slices in $\mathcal{C}$ as follows:

$$\Psi(\Sigma_{t,\mathcal{X}}) := \left(\mathcal{H}^{\otimes \mathcal{N}}\right)^{\otimes \mathcal{X}} = \mathcal{H}^{\otimes (\mathcal{N} \times \mathcal{X})}. \tag{41}$$

The tensor product is well-defined in all symmetric monoidal categories, since $\mathcal{X}$ is always finite. Physically, the field takes values in a copy of $\mathcal{H}^{\otimes \mathcal{N}}$ over each event $(t, \underline{x})$ of spacetime, each individual $\mathcal{H}$ factor of $\mathcal{H}^{\otimes \mathcal{N}}$ encoding the contribution to the field state at $(t, \underline{x})$ from the field state at each of its immediate causal predecessors in $(t-1, \underline{x} + \mathcal{N})$.

6.2.5 Causal Field Theory - Restriction and Evolution

From their definition in **Eq. 38**, it is easy to see that morphisms $\Sigma_{t,\mathcal{X}_0} \twoheadrightarrow \Sigma_{t+k,\mathcal{X}_k}$ on $\mathcal{C}$ can always be factored in the following way:

$$\Sigma_{t,\mathcal{X}_0} \twoheadrightarrow \Sigma_{t,\mathcal{Y}_0} \twoheadrightarrow \Sigma_{t+1,\mathcal{X}_1} \twoheadrightarrow \Sigma_{t+1,\mathcal{Y}_1} \twoheadrightarrow \ldots \twoheadrightarrow \Sigma_{t+k,\mathcal{X}_k}, \tag{42}$$

where $\mathcal{Y}_i \subseteq \mathcal{X}_i$ for all $i = 0, \ldots, k-1$ and the following holds for each $i = 1, \ldots, k$:

$$\mathcal{Y}_{i-1} = \bigcup_{\underline{x} \in \mathcal{X}_i} \{(t+i-1, \underline{x} + \underline{\delta}) | \underline{\delta} \in \mathcal{N}\}. \tag{43}$$

This means that we only need to care about the action of Ψ on two kinds of morphisms:

- The restrictions $\Sigma_{t,\mathcal{X}} \twoheadrightarrow \Sigma_{t,\mathcal{Y}}$, where $\mathcal{Y} \subseteq \mathcal{X}$;
- The 1-step evolutions $\Sigma_{t,\mathcal{Y}} \twoheadrightarrow \Sigma_{t+1,\mathcal{X}}$, where $\mathcal{Y} = \bigcup_{\underline{x} \in \mathcal{X}} \{(t, \underline{x} + \underline{\delta}) | \underline{\delta} \in \mathcal{N}\}$.

The existence of the factorisation above can be proven by induction, observing that any morphism $\Sigma_{t,\mathcal{X}_0} \twoheadrightarrow \Sigma_{t+1,\mathcal{X}_1}$ factors into the product:

$$\left(\Sigma_{t,\mathcal{Y}_0} \twoheadrightarrow \Sigma_{t+1,\mathcal{X}_1}\right) \otimes \left(\Sigma_{t,\mathcal{X}_0 \setminus \mathcal{Y}_0} \twoheadrightarrow \emptyset\right), \tag{44}$$

where $\mathcal{Y}_0$ is defined as before so that $\Sigma_{t,\mathcal{Y}_0}$ is exactly the set of immediate causal predecessors of the codomain $\Sigma_{t+1,\mathcal{X}_1}$.

On restrictions $\Sigma_{t,\mathcal{X}} \twoheadrightarrow \Sigma_{t,\mathcal{Y}}$, where $\mathcal{Y} \subseteq \mathcal{X}$, the functor Ψ is defined to act by marginalization, discarding the field state over all those events in the larger slice $\Sigma_{t,\mathcal{X}}$ which don't belong to the smaller slice $\Sigma_{t,\mathcal{Y}}$:

$$\Psi(\Sigma_{t,\mathcal{X}} \twoheadrightarrow \Sigma_{t,\mathcal{Y}}) := \bigotimes_{\underline{x} \in \mathcal{X}} F_{\underline{x}} \quad \text{where} \quad F_{\underline{x}} := \begin{cases} id_{\mathcal{H}^{\otimes \mathcal{N}}} & \text{if } \underline{x} \in \mathcal{Y} \\ \overset{\blacksquare}{\top}_{\mathcal{H}^{\otimes \mathcal{N}}} & \text{if } \underline{x} \notin \mathcal{Y} \end{cases} \tag{45}$$

On 1-step evolutions $\Sigma_{t,\mathcal{Y}} \twoheadrightarrow \Sigma_{t+1,\mathcal{X}}$, where $\mathcal{Y} = \cup_{\underline{x} \in \mathcal{X}} \{(t, \underline{x} + \underline{\delta}) | \underline{\delta} \in \mathcal{N}\}$, the functor Ψ is defined to act by a combination of evolution by U and marginalization. The evolution component is simply an application of U to the state at each event of $\mathcal{Y}$:

[8] The reason for the negative sign in $\underline{x} - \underline{\delta}$ is that $\mathcal{N}$ was originally defined to be the neighborhood in the past.

$$U^{\otimes\mathcal{Y}} : \mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})} \to \mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})}. \tag{46}$$

The marginalization component then needs to go from the codomain $\mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})}$ of the map above to the desired codomain $\mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{X})}$. To do this, we recall that the $\mathcal{H}$ factor of $\mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{X})}$ corresponding to a given $\underline{\delta} \in \mathcal{N}$ and a given $\underline{x} \in \mathcal{X}$ is intended to encode the component of the state at $(t+1, \underline{x})$ coming from $(t, \underline{x} + \underline{\delta})$. Analogously, the $\mathcal{H}$ factor of $\mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})}$ corresponding to a given $\underline{\delta} \in \mathcal{N}$ and a given $\underline{y} \in \mathcal{Y}$ is intended to encode the component of the evolved state going to $\left(t+1, \underline{y} - \underline{\delta}\right)$. Hence to go from $\mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})}$ to $\mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{X})}$ we need to discard all factors in $\mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})}$ corresponding to components of the evolved state which are not going to some $\left(\underline{y} - \underline{\delta}\right) \in \mathcal{X}$:

$$\left(\bigotimes_{(\underline{\delta},\underline{y})\in\mathcal{N}\times\mathcal{Y}} F_{\underline{\delta},\underline{y}}\right) : \mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})} \to \mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{X})} \quad \text{where} \quad F_{\underline{\delta},\underline{y}} := \begin{cases} id_{\mathcal{H}} & \text{if } (\underline{y}-\underline{\delta}) \in \mathcal{X} \\ \dagger_{\mathcal{H}} & \text{if } (\underline{y}-\underline{\delta}) \notin \mathcal{X} \end{cases} \tag{47}$$

Putting the evolution and marginalization components together we get the action of Ψ on 1-step evolutions:

$$\Psi\left(\Sigma_{t,\mathcal{Y}} \twoheadrightarrow \Sigma_{t+1,\mathcal{X}}\right) := \left(\left(\bigotimes_{(\underline{\delta},\underline{y})\in\mathcal{N}\times\mathcal{Y}} F_{\underline{\delta},\underline{y}}\right) \circ U^{\otimes\mathcal{Y}}\right) : \mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{Y})} \to \mathcal{H}^{\otimes(\mathcal{N}\times\mathcal{X})}. \tag{48}$$

By construction, the above is a G-invariant causal field theory, completing the definition of our partitioned causal cellular automaton. If U is an isomorphism, the same construction on $\mathcal{C}^{rev}$ using U^{-1} provides a causal reversal for Ψ, showing that the partitioned causal cellular automaton above is reversible under those circumstances. Finally, **Figure 7** below depicts an example of action on morphisms for a $(1+1)$-dimensional partitioned causal cellular automaton.

6.3 Sketch of the Continuous Limit for the Dirac Quantum Cellular Automata

To conclude, we note how in [17] it is argued that the Dirac equation for free propagation of an electron can be recovered in the continuous limit of a specific $(1+1)$-dimensional partitioned QCA. The original argument could not be made fully rigorous, because the QCAs defined therein were discrete and no setting was available to the author in which to make proper sense of the infinite tensor product arising from the limiting construction. A rigorous analysis of the limit is presented in [31], but the limit itself exists outside of the QCA framework.

Our definition of CCA, on the other hand, has no requirement of discreteness. Furthermore, the freedom left in the choice of field category for a CCA allows us to benefit from the full power of the non-standard approach to categorical quantum mechanics [24, 32]. As a consequence, we are able to sketch below a formalization within our framework of the continuous limit for the Dirac QCA, following the same lines as the construction of a $(1+d)$-dimensional partitioned CCA above.

The key to obtain a continuous limit for the Dirac QCA is to rescale the discrete lattice Ω to one with infinitesimal mesh ε:

$$\varepsilon\Omega := \left\{\varepsilon\left(t, \underline{x}\right) \,\middle|\, t \in {}^{\star}\mathbb{Z},\ \underline{x} \in (t, \ldots, t) + 2^{\star}\mathbb{Z}^d\right\}, \tag{49}$$

where ${}^{\star}\mathbb{Z}$ are the non-standard integer numbers. The slices are now allowed to contain an infinite number of points and can be used to approximate all equal-time partial Cauchy hypersurfaces in $(1+d)$-dimensional Minkowski spacetime. Unfortunately, the infinite number of points in our slices now requires infinite tensor products to be taken: to deal with this, we use as our field category the dagger compact category *Hilb of non-standard hyperfinite-dimensional Hilbert spaces, where such infinite products can be handled safely.

We set the scattering map to be the following non-standard unitary

$$U = 1 \oplus \sigma_X \exp\left(-im\varepsilon\sigma_X\right) \oplus 1, \tag{50}$$

where σ_X is the X Pauli matrix: this is the same unitary used in the Dirac QCA, but with the real parameter ε turned into an infinitesimal. Each application of U only inches infinitesimally further from the identity, but in the non-standard setting we are allowed to consider the cumulative effect across infinite sequences of infinitesimally close slices. The first order approximations to the Dirac equation derived in [17] turn into legitimate infinitesimal differentials, connecting the state on each slice to the state on the (infinitesimally close) next slice: once the standard part is taken, the lattice $\varepsilon\Omega$ ends up covering the entirety of $(1+d)$-dimensional Minkowski spacetime, the differentials get integrated and Ψ turns into a continuous-time field evolution following the Dirac equation.

Remark 47: *The power to express limiting constructions algebraically, without exiting the original framework, is one of the most attractive aspects of non-standard analysis. The dagger compact category *Hilb (and other categories derived from it) can be used to make categorical sense of constructions from quantum field theory [24, 32], including other cellular automata with field-theoretic continuous limits. The formulation of such limits within our framework is an point of great interest, but is left to future work.*

7 CONCLUSION AND FUTURE WORK

In this work, we have defined a functorial, theory-independent notion of causal field theory founded solely on the order-theoretic structure of causality. We have seen how the causality requirement for such field theories is automatically satisfied as a consequence of symmetry-breaking in the ordering on space-like slices. In an effort to connect to Algebraic Quantum Field Theory (AQFT), we have constructed complex spaces of states over regions of spacetime and discussed how the associated information redundancy can be reduced in selected cases. We have introduced symmetries in our framework and shown that Quantum Cellular Automata (QCA) can be modeled within it, both in their traditional discrete formulation and in their continuous limit.

Despite our efforts, we feel we have barely scratched the surface on the potential of this material. In the future, we envisage three lines of

research stemming from this work. Firstly, we believe that the connection with AQFT can be strengthened and honed to the point that the framework will be a tool for the construction of new models. This includes a thorough understanding of the structure of spaces of states for categories of slices more general than those induced by foliations. Secondly, we wish to further explore and fully characterise the possibilities associated with working in the continuous limit of QCAs, with an eye to applications in perturbative quantum field theory. Finally, we plan to extend the framework in a number of directions, including indefinite causal order—already achieved for QCAs by [30], at least in partial form—enrichment and the possibility of working with restricted classes of causal paths (in temporal analogy to categories of slices).

AUTHOR CONTRIBUTIONS

Research carried out by SG and by MS under the supervision of BC. Writing carried out by SG and MS, with editing by BC. Revision and approval of work carried out by SG, MS, and BC.

FUNDING

This work is supported by a grant form the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect on the views of the John Templeton Foundation. MS is supported by an EPSRC grant, ref. EP/P510270/1.

REFERENCES

1. Malament DB. The Class of Continuous Timelike Curves Determines the Topology of Spacetime. *J Math Phys* (1977) 18:1399–404. doi:10.1063/1.523436
2. Kronheimer EH, and Penrose R. On the Structure of Causal Spaces. *Math Proc Camb Phil Soc* (1967) 63:481–501. doi:10.1017/S030500410004144X
3. Hawking SW, King AR, and McCarthy PJ. A New Topology for Curved Space-Time Which Incorporates the Causal, Differential, and Conformal Structures. *J Math Phys* (1976) 17:174–81. doi:10.1063/1.522874
4. Martin K, and Panangaden P. Domain Theory and General Relativity. In: *New Structures for Physics*. Berlin, Heidelberg: Springer (2010). p. 687–703. doi:10.1007/978-3-642-12821-9_11
5. Martin K, and Panangaden P. Spacetime Geometry from Causal Structure and a Measurement. In: Mathematical Foundations of Information Flow: Clifford Lectures Information Flow in Physics, Geometry, and Logic and Computation; March 12-15, 2008; New Orleans, Louisiana. . Tulane University (2012). p. 213, 2012 . 71.
6. Bombelli L, Lee J, Meyer D, and Sorkin RD. Space-time as a Causal Set. *Phys Rev Lett* (1987) 59:521–4. doi:10.1103/PhysRevLett.59.521
7. [Dataset] Bombelli L. Causal Sets: Overview and Status. In: Presentation at QGA3 Conference (2006). Available at: https://ipfs.io/ipfs/QmfFjRpUFxrNTnV7mkYiCRXLiokJQaPCeum51zaYLyt8Wq.
8. Pinzani N, Gogioso S, and Coecke B. Categorical Semantics for Time Travel. In: 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2019). (IEEE Computer Society) (2019). doi:10.1109/LICS.2019.8785664
9. Haag R, and Kastler D. An Algebraic Approach to Quantum Field Theory. *J Math Phys* (1964) 5:848–61. doi:10.1063/1.1704187
10. Halvorson H, and Müger M. Algebraic Quantum Field Theory. In: *Handbook of the Philosophy of Physics*. (North Holland) (2006).
11. Heunen C, Landsman NP, and Spitters B. A Topos for Algebraic Quantum Theory. *Commun Math Phys* (2009) 291:63–110. doi:10.1007/s00220-009-0865-6
12. Doering A, and Isham C. What Is a Thing?. In: *Topos Theory in the Foundations of Physics*. Berlin, Heidelberg: Springer (2008). p. 753–940. doi:10.1007/978-3-642-12821-9_13
13. Lurie J. On the Classification of Topological Field Theories. *Curr Dev Math* (2008) 2008:129–280. doi:10.4310/cdm.2008.v2008.n1.a3
14. Atiyah M. Topological Quantum Field Theories. *Publications Mathématiques de l'Institut des Hautes Scientifiques* (1988) 68:175–86. doi:10.1007/bf02698547
15. Witten E. Topological Quantum Field Theory. *Commun.Math Phys* (1988) 117:353–86. doi:10.1007/bf01223371
16. D'Ariano GM, and Perinotti P. Quantum Cellular Automata and Free Quantum Field Theory. *Front Phys* (2016) 12:120301. doi:10.1007/s11467-016-0616-z
17. Arrighi P. An Overview of Quantum Cellular Automata. *Nat Comput* (2019) 18:885–99. doi:10.1007/s11047-019-09762-6
18. von Neumann J. *Theory of Self-Reproducing Automata*. Champaign, IL, USA: University of Illinois Press (1966).
19. Coecke B. Terminality Implies No-Signalling ...and Much More Than That. *New Gener Comput* (2016) 34:69–85. doi:10.1007/s00354-016-0201-6
20. Coecke B, and Lal R. Causal Categories: Relativistically Interacting Processes. *Found Phys* (2013) 43:458–501. doi:10.1007/s10701-012-9646-8
21. Kissinger A, and Uijlen S. A Categorical Semantics for Causal Structure. In: *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE (2017). p. 1–12. doi:10.1109/LICS.2017.8005095
22. Gogioso S. A Process-Theoretic Church of the Larger hilbert Space. EPTCS (QPL (2019). arxiv Preprint: arxiv.org/abs/1905.13117.
23. Gogioso S, and Scandolo CM. Categorical Probabilistic Theories. *Electron Proc Theor Comput Sci* (2017) 266:367–85. doi:10.4204/EPTCS.266.23
24. Gogioso S, and Genovese F. Quantum Field Theory in Categorical Quantum Mechanics. *Electron Proc Theor Comput Sci* (2018) 287:163–77. doi:10.4204/EPTCS.287.9
25. Coecke B, and Kissinger A. *Picturing Quantum Processes*. Cambridge, United Kingdom: Cambridge University Press (2017). doi:10.1017/9781316219317
26. Kock J. *Frobenius Algebras and 2D Topological Quantum Field Theories*. Cambridge, United Kingdom: Cambridge University Press (2003). doi:10.1017/CBO9780511615443
27. von Neumann J. On Infinite Direct Products. *Compositio Mathematica* (1939) 6: 1–77.
28. Feynman RP. Simulating Physics with Computers. *Int J Theor Phys* (1982) 21: 467–88. doi:10.1007/BF02650179
29. Arrighi P, and Dowek G. Causal Graph Dynamics. *Inf Comput* (2013) 223: 78–93. doi:10.1016/j.ic.2012.10.019
30. Arrighi P, and Martiel S. Quantum Causal Graph Dynamics. *Phys Rev D* (2017) 96:024026. doi:10.1103/PhysRevD.96.024026
31. Arrighi P, Nesme V, and Forets M. The Dirac Equation as a Quantum Walk: Higher Dimensions, Observational Convergence. *J Phys A: Math Theor* (2014) 47:465302. doi:10.1088/1751-8113/47/46/465302
32. Gogioso S, and Genovese F. Towards Quantum Field Theory in Categorical Quantum Mechanics. *Electron Proc Theor Comput Sci* (2017) 266:349–66. doi:10.4204/EPTCS.266.22

Conflict of Interest: 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 and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership