COMPUTATIONAL PROBABILITY AND MATHEMATICAL MODELING

EDITED BY: José Roberto Cantú-González, F-Javier Almaguer, Javier Morales-Castillo and Pavel Solin

frontiers Research Topics

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

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

COMPUTATIONAL PROBABILITY AND MATHEMATICAL MODELING

Topic Editors:

José Roberto Cantú-González, Universidad Autónoma de Coahuila, Mexico
F-Javier Almaguer, Universidad Autónoma de Nuevo León, Mexico
Javier Morales-Castillo, Universidad Autónoma de Nuevo León, Mexico
Pavel Solin, University of Nevada, Reno, United States of America
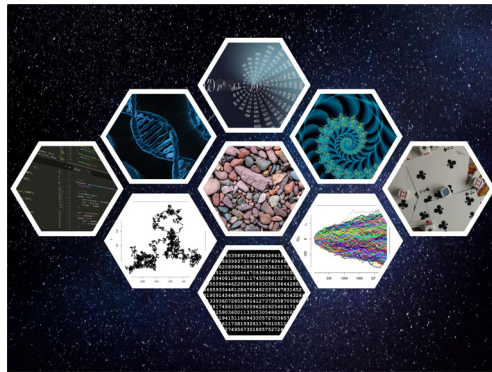
Image: Computational Probability and Mathematical Modeling.
By: José Roberto Cantú-González and F-Javier Almaguer.

In the present time, two of the most important approaches to tackle complex systems are probability and stochastic processes theory. Still from an analytic perspective, modeling and solving a problem using a stochastic approach is not a trivial issue, hence, a combination of the logic of probabilistic reasoning with computational science is needed to obtain qualitatively good solutions in a reasonable time.

This eBook presents an interesting view of applications associated to fields of probability, statistics, and mathematic modeling, all of them supported by a computational context though the approach of stochasticity and simulation used in most of them. This collection contains three chapters, which bring applications in fields of biology, finance and physics, each chapter contains work(s) with specific applications. An editorial is also contained with a summarized version of each work, and each of them are widely explained in a specific section, which include a state of art to support the nature of the individual research, a methodology to solve the defined problem and the results and conclusions.

We hope the present eBook can represent a potential source of knowledge for the academic community of implicated disciplines, and an inspirational starting point of starting for scientists in the amazing world of applied mathematics and the search to solve complex problems

Table of Contents

Editorial: Computational Probability and Mathematical Modeling - A Stochastic Approach in Applied Sciences

José Roberto Cantú-González[1]* and F-Javier Almaguer[2]*

[1] *Escuela de Sistemas PMRV, Campus Acuña, Universidad Autónoma de Coahuila, Saltillo, Mexico,* [2] *Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Mexico*

Editorial on the Research Topic

Computational Probability and Mathematical Modeling - A Stochastic Approach in Applied Sciences

INTRODUCTION

Probability and the stochastic processes theory, joint to the mathematical modeling and the respective computational support are clearly important work tools to tackle complex systems, which has been the fundament to develop the Research Topic of "Computational Probability and Mathematical Modeling – A Stochastic Approach in Applied Sciences".

In a brief analysis of the elements contained inside of the Research Topic, we find three important scopes: computational probability, mathematical modeling, and stochastics processes:

Computational probability, it is defined as the development of data structures and algorithms to automate the derivation of existing and new results in probability and statistics [1], Mathematical modeling is the translation of a specific problem from the natural sciences (experimental physics, chemistry, biology, geosciences) or the social sciences, or from technology, into a well-defined mathematical problem [2] and a stochastic process is a probability model that describes the evolution of a system evolving randomly in time [3]. In addition, advances in computer science have been very determinant for growing of science and technology, and all fields of mathematics included stochastic processes have not been bypassed by the digital revolution, because numerical calculation and computer simulation play a decisive role in present-day [4].

Probability and mathematical modeling are usually very adaptable for different proposes, and in definitive, they are the starting point to solve relevant problems [5].

Specifically probability and the stochastic processes theory can be applied in any field where random plays a preponderant role and /or when an analytic treatment is too complicated to be solved. In the other context, mathematical modeling as a descendent from applied mathematics rejects the abstractionism to support the science for understanding of the nature laws; which explains that mathematical models can attend disciplines as physics, biology, finances, economics, any engineering and a variety of fields to solve real problems though its included skills. In this perspective, both probability / stochastic processes theory as the mathematical modeling converge on the same objective: attention of real problems in complex systems.

RESEARCH TOPIC AND THE ARTICLES CONTAINED

This Research Topic of "Computational Probability and Mathematical Modeling – A Stochastic Approach in Applied Sciences" presents a little but substantial view of studies related to interesting scenarios from the applied science in fields of genetics, biology, physics, finances, and investment options. In addition, as a core and distinguished common value in these studies, we have that even though all of them come from different disciplines and topics, in a direct or indirect perspective, they possess the essence of probability and mathematical modeling, and they have an objective directed to solve a particular issue. The works contained in this Research Topic combine all the implicit elements of the general theme, and in the most of them it is mentioned the use of simulation as a primary resource in the handling of studied factors.

Individual works in alphabetical order are:

- A Stochastic Phylogenetic Algorithm for Mitochondrial DNA Analysis
- On the Aerodynamic Forces on a Baseball, With Applications
- Output-Feedback Control of Virus Spreading in Complex Networks With Quarantine
- Probabilistic Assessment of Investment Options in Honey Value Chains in Lamu County, Kenya
- Synchronization, Oscillator Death, and Frequency Modulation in a Class of Biologically Inspired Coupled Oscillators
- The Amnesiac Lookback Option: Selectively Monitored Lookback Options and Cryptocurrencies

A Stochastic Phylogenetic Algorithm for Mitochondrial DNA Analysis (Corona-Ruiz et al.):

This research is an exploratory analysis of the mitochondrial DNA and proposes new indices as functions of Shannon entropy, the Chargaff ratio, and fractal dimensions using rescaled-range analysis and DFA; the work suggests to utilize the triplet of indices to construct phylogenetic trees using clustering algorithms. Additionally, it is an initiative to identify the tendencies and correlations in the mutations that produce new species throughout evolutionary history.

On the Aerodynamic Forces on a Baseball, With Applications (Santos et al.):

This article combines experimental results, phenomenological and dimensional analysis; it is classifies as a review, and summarize from recent literature some methods regarding the reproduction and reconstruction of baseball trajectories from aerodynamic forces, and finally discusses their potential applications.

Output-Feedback Control of Virus Spreading in Complex Networks with Quarantine (Alarcn-Ramos et al.):

The research proposes a simple output-feedback control, which stabilizes the extinction state in a virus spreading process over a complex network with quarantine. Additionally, it is provided numerical simulation results to illustrate the functioning of the proposed control scheme for a scale-free network of $N = (10)^6$ nodes.

Probabilistic Assessment of Investment Options in Honey Value Chains in Lamu County (Wafula et al.):

This work, classified as method brings an analysis to demonstrate a comprehensive approach to decision-making in a project where decision outcomes are uncertain.

The structure of the decision model includes several strategies, of which the Stochastic Impact Evaluation approach (SIE) and Monte Carlo simulation clearly stands out. SIE approach is used as fundamental base for the decision analysis; particularly SIE prioritizes the critical uncertainties considering the risk factors that may compromise the success and the performance of the project. By the other hand, through the use of Monte Carlo simulation, it was possible to obtain the results of project performance over 10 years, to get it possible the model was ran a total of 10,000 times. This propose is substantially useful for the process of decision-making and the probabilistic assessment of investment options, the case of honey value chains in Lamu County, Kenya is analyzed as study case.

Synchronization, Oscillator Death, and Frequency Modulation in a Class of Biologically Inspired Coupled Oscillators (Franci et al.):

This research is focused on the basic mathematical principles that underlie the emergence of synchronous biological rhythms, in particular, the circadian clock. The study analyzes the role that the coupling strength, coupling type, and noise play in the synchronization of a system of coupled, non-linear oscillators. Authors report the following two phenomena found, whose are described as new from a non-linear collective phenomena perspective: In diffusively coupled cells, resting node dynamics imply global asymptotic stability; oscillating node dynamics imply global-synchronization for small coupling, and multistability between oscillator death and global synchronization for large coupling. In stochastic, linearly coupled populations, it is described the dynamical mechanisms through which coupling modulates the frequency of the synchronous oscillation. Finally, the article and its presented results, emphasizes the importance of simple mathematical models in understanding situations where synchronization of multiple oscillating populations appears. Additionally, Authors believe that results obtained may help to shed light on physiological and pathological phenomena involving synchronization of oscillators in important tissues as Parkinson's disease [6, 7], and epilepsy [8, 9].

The Amnesiac Lookback Option: Selectively Monitored Lookback Options and Cryptocurrencies (Chang and Li):

This research suggests the use of its properties to reduce risk exposure in cryptocurrency markets through blockchain enforced smart contracts and correct for informational inefficiencies surrounding prices and volatility. In addition, this work generalizes partial, discretely monitored lookback options that dilute premiums by selecting a subset of specified periods to determine payoff, which authors call amnesiac lookback options. As part of the utilized method, authors price Amnesiac lookbacks with Monte Carlo simulations of Gaussian random walks under equidistant and random periods. The paper concludes that the instrument provides an ideal space for investors to balance their risk, and as a prime candidate to hedge extreme volatility.

CONCLUDING REMARKS

This collection of articles included in this volume has provided an interesting view of applications associated to fields of probability, statistics, and mathematic modeling, all of them supported by a computational context; additionally it has been determinant the approach of stochasticity and simulation used in most of them as a key element in the utilized methodology. Each article contained, describes a particular object of study, bring a state of art to support the nature of the research, presents a methodology to solve the defined problem and finally provides results and conclusions as the evidence of the efficiency of the entire propose.

We hope the present collection of articles can be attractive for readers from the academic community of mathematics and the applied sciences in general; also, it is our desire that these papers can provide a starting point for future researches in solving of practical and complex problems.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

REFERENCES

1. Drew JH, Evans DL, Glen AG, Leemis LM. *Computational Probability, Algorithms and Applications in the Mathematical Sciences*. Cham: Springer International Publishing (2017).

2. Eck C, Garcke H, Knabner P. *Mathematical Modeling*. Cham: Springer International Publishing (2017).

3. Kulkarni VG. *Introduction to Modeling and Analysis of Stochastic Systems*. New York, NY: Springer Science+Business Media, LLC (2011).

4. Schuster P. *Stochasticity in Processes, Fundamentals and Applications to Chemistry and Biology*. Cham: Springer International Publishing AG (2016).

5. Heinz S. *Mathematical Modeling*. Berlin; Heidelberg: Springer-Verlag (2011).

6. Hammond C, Bergman H, Brown P. Pathological synchronization in Parkinson's disease: networks, models and treatments. *Trends Neurosci.* (2007) **30**:357–64. doi: 10.1016/j.tins.2007.05.004

7. Schnitzler A, Gross J. Normal and pathological oscillatory communication in the brain. *Nat Revi Neurosci.* (2005) **6**:285–96. doi: 10.1038/nrn1650

8. Engel JJ, Bragin A, Staba R, Mody I. High-frequency oscillations: what is normal and what is not? *Epilepsy.* (2009) **50**:598–604. doi: 10.1111/j.1528-1167.2008.01917.x

9. Velazquez J, Carlen P. Gap junctions, synchrony and seizures. *Trends Neurosci.* (2000) **23**:68–74. doi: 10.1016/S0166-2236(99)

Output-Feedback Control of Virus Spreading in Complex Networks With Quarantine

Luis A. Alarcón-Ramos[1], Roberto Bernal Jaquez[2*] and Alexander Schaum[3]

[1] Posgrado en Ciencias Naturales e Ingeniería, Universidad Autónoma Metropolitana-Cuajimalpa, Mexico City, Mexico,
[2] Departamento de Matemáticas Aplicadas y Sistemas, Universidad Autónoma Metropolitana-Cuajimalpa, Mexico City, Mexico, [3] Chair of Automatic Control, Kiel-University, Kiel, Germany

In this paper the problem of designing an output-feedback control for the stabilization of the extinction steady-state in a virus spreading process over a complex network with quarantine is considered. Sufficient conditions are established for the choice of those nodes for which sensor information is necessary and those which should be controlled using notions from constructive control theory. A simple output-feedback control is proposed which exponentially stabilizes the extinction state. Numerical simulation results are provided to illustrate the functioning of the proposed control scheme for a scale-free network of $N = 10^6$ nodes.

Keywords: complex networks, virus spreading, feedback control, quarantine, sensor location

1. INTRODUCTION

Studies on the propagation and control of viruses and infectious diseases in human and animal populations [1–8] have gained a great importance in the last years. Understanding and controlling spreading processes is a problem of interdisciplinary nature [9–12]. The use of mathematical and physical inspired models that give account of the dynamics of propagation of infections have gained acceptance since the London cholera epidemics in September 1853 when John Snow divided a map of London into sectors (in a way, nowadays, equiparable to a Voronoi diagram) to calculate who was most likely to use each water pump in the city and in this way, discovered that the 40 Broad Street pump was the main focus of the cholera infection.

Some decades later, McKendrick [13] and Kermack and McKendrick [14] proposed the SIR model that divides the population in three different compartments or groups of individuals: Susceptible, Infective, and Recovered (or Removed) that had the disease and become immune or died. This pioneering work gave birth to the SIS (Susceptible-Infected-Susceptible) model mainly because many diseases do not confer any immunity. Extensions such as the SIQ (Susceptible-Infected-Quarantine) model (see e.g., [15–17]) appeared in order to account for the dynamics of the spreading of infections when quarantine has been implemented as a means to control the massive spreading of the infectious disease. Accordingly, a new class of quarantined individuals is included. These individuals are removed, with some probability from the class of infected individuals.

Although these models were proposed for modeling spreading in populations with no structure, nowadays the SIS and SIQ models have been extended to model the spreading process in a population mapped on a complex network.

In recent years, Markov-chain based models for a Susceptible-Infected-Susceptible (SIS) dynamics over complex networks have been used [4–8, 18] to describe spreading processes in networks. Using these models, it is possible to determine the macroscopic properties of the system as well as the description of the dynamics of individual nodes. One interesting result is that the calculated infection threshold depends on the value of the spectral radius of the adjacency matrix, thus relating the network structure with spreading behavior.

In the present work we will study the spreading of diseases using a Markov-chain based model for a Susceptible-Infected-Quarentine (SIQ) dynamics over a complex network that has been used in a simplified version for the determination of stability tresholds in recent studies by the authors [15, 17]. We will show that, even when the built-in control strategy of quarantine is not able to ensure virus extinction, it is always possible to solve the problem of epidemic spreading extinction. In order to accomplish this task, we determine sufficient conditions for stabilizing the extinction state in a complex network of arbitrary topology. One remarkable point is that, these conditions give a clue to identify the nodes that do not need any control to reach the extinction state and distinguish them from the nodes that need to be controlled. Inspired on the ideas of control theory, the set of nodes that do not need to be controlled (in order to reach the extinction state) will be associated with the zero dynamics of the system [19, 20] . At the same time, the set of nodes to be monitored and controlled will be identified and a decentralized feedback control will be applied in order to stabilize the extinction state. Accordingly, it is proven that the extinction state is an exponentially stable fixed point for the zero dynamics.

We have performed numerical simulations using a scale-free complex network with 1 million nodes constructed as proposed by Barabasi [21] and found a complete agreement of the numerical results with our theoretical findings.

The paper is organized as follows: section Methods presents the problem statement for the SIQ model mapped on a complex network of arbitrary topology and the control problem we want to solve together with some definitions. In section Results, sufficient conditions for the stabilization of the extinction state are derived. Using these conditions we establish a selection criterion that allows to identify the set of nodes that need to be controlled in order to reach the extinction state. Afterwards, we design a simple stabilizing output-feedback control and present our simulation results. In section Discussion and Conclusions we summarize our main results and present our main conclusions.

2. METHODS

As pointed out in the introduction, the SIQ model belongs to a class of models that are used to capture possible human interactions employed to impede the spreading of an infection process and it is essentially an extension of the SIS model in which the model and control strategies are co-developed to yield a kind of closed-loop control model [22]. In order to reveal the dynamics and essential mechanisms of this model, we proceed to

(i) formulate the SIQ model in a complex network (ii) introduce the specific control problem.

2.1. The SIQ Model in a Complex Network

Consider a network of N nodes described by an undirected graph $G(V, E)$ of any topology. Let $V = \{v_1, v_2, \ldots, v_N\}$ being the set of nodes and $E = \{e_{i,j}\}$ the set of connecting edges. The adjacency matrix associated to $G(V, E)$ is given by $A = \{a_{ij}\}$, where $a_{ij} = 1$ if $e_{i,j} \in E$ and zero otherwise. The set of neighbor nodes of a node $v_i \in V$ is defined as

$$V_i = \{v_j \in V \mid a_{ij} = a_{ji} = 1\} \subset V, \qquad (1)$$

and the number of neighbors or degree of a node v_i is given by $\mathcal{N}_i = |V_i|$.

As formulated in Bernal Jaquez et al. [15], the underlying process for every node is depicted in **Figure 1**, as a discrete time Markov process. A node v_i can be in state I (Infected) with probability $p_i(t)$, in state Q (Quarantine) with probability $q_i(t)$, or in state S (Susceptible) with probability $s_i(t) = 1 - p_i(t) - q_i(t)$. At each time step, the probability functions $p_i(t), s_i(t)$ are updated due to the fact that every node can transit from state S to state I with probability $1 - \zeta_i$ or from state I to state Q with probability τ_i and from state Q to state S with probability μ_i because nodes are interacting.

According to the transition diagram shown in **Figure 1**, we have the following dynamical system:

$$\begin{aligned}
p_i(t+1) &= (1-\tau_i)p_i(t) + (1-\zeta_i(P(t)))s_i(t), \\
q_i(t+1) &= (1-\mu_i)q_i(t) + \tau_i p_i(t), \qquad (2) \\
s_i(t+1) &= \zeta_i(P(t))s_i(t) + \mu_i q_i(t). \\
p_i(t) &+ q_i(t) + s_i(t) = 1, \quad v_i \in V,
\end{aligned}$$

where, for every node $v_i \in V$, τ_i is the internment probability associated to quarantine, μ_i is the recovery probability and
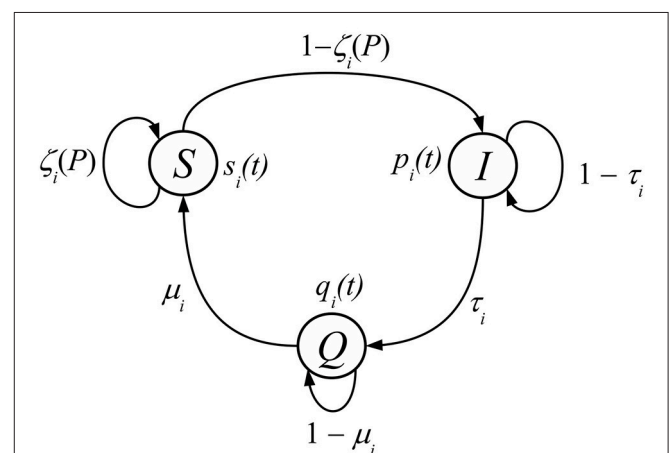
FIGURE 1 | State transition diagram for each node $v_i \in V$. The states Infected, Quarantine and Susceptible are represented by I, Q, and S, respectively.

$\zeta_i(P(t))$ is the probability of a node v_i of not being infected by any neighbor at time t, which is given by

$$\zeta_i(P(t)) = \prod_{j=1}^{N}(1 - \beta_i r_{ij} a_{ij} p_j(t)). \qquad (3)$$

In the above equation, we have defined the vector $P(t) = [p_1(t), p_2(t), ..., p_N(t)]^T \in [0, 1]^N$ and β_i is the probability of infection during a single contact, r_{ij} is the probability that the node v_i performs at least one contact intent with its neighbor $v_j \in V_i$. The probability r_{ij} is also known as the connection probability and depends on the number of contact intents (or interaction rate).

Note that in system (2)

$$0 \leq \mu_i, \tau_i, r_{ij}, \beta_i, \zeta_i(P(t)), p_i(t), q_i(t), s_i(t) \leq 1$$

From the model (2) we point out that

- Unlike the model proposed in Bernal Jaquez et al. [15] the system (2) considers non-homogeneous properties given by μ_i, τ_i, β_i and r_{ij}.
- Quarantine is a control mechanism which is introduced in order to prevent virus spreading, however it could probably not be sufficient for achieving the extinction state, depending on the stability conditions (to be determined below). We can consider, instead, the objective of this work to design a controller to stabilize the extinction state, e.g., by adapting inherent propagation parameters like the propability of infection β_i of some nodes to be specified.

We consider that each node has a manipulable variable $u_i(t)$, which is amenable for control. For the system 2, this manipulable variable can be β_i or r_{ij}, i.e., we consider that, for each node v_i, it is possible to improve its health or avoid to perform several contact attempts with its neighbor nodes, the control will adapt one of these parameters, that will be selected according to our mathematical analysis.
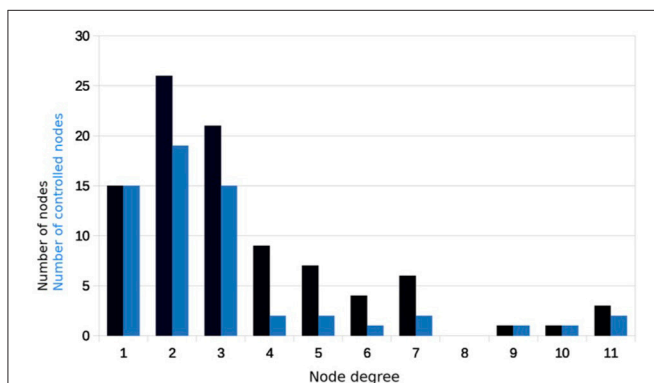
FIGURE 2 | Number of nodes with a given degree (black) in a non-homogenous scale-free network with $N = 100$ nodes and normally distributed parameters and number of nodes of the given degree that need to be controlled (blue).

2.2. Control Problem

Quarantine is an heuristic control mechanism or strategy that intents to reduce a virus propagation in a population. As stated in the last subsection, these kind of models probably are not sufficient to ensure the system to reach the extinction state. This is due to the fact that no adaptation of the network parameters is implemented in dependence of the actual system state. Such a mechanism is proposed in the sequel including decisions on (i) the subset of nodes $V_M \subset V$ which should be monitored, i.e., whose actual state must be known at each time instant, (ii) the subset of nodes $V_C \subset V$ which have to be controlled, i.e., whose interaction parameters ($u_i = \beta_i$ or $u_i = r_{ij}$) should be subject to *on-line* adaptation, and (iii) the specific control law $u_i(t) = \varphi_i(P(t))$ which should be used for this adaptation.

The chosen approach follows the constructive (i.e., passivity-based) control idea, and consist in two steps:

a) Assigning the necessary outputs so that the associated zero dynamics is asymptotically stable.
b) Designing controllers $u_i(t) = \varphi_i(Y_i(t))$ so that for some $0 \leq \gamma < 1$ it holds that

$$p_i(t) \leq p_{io}\gamma^t, \quad p_i(t = 0) = p_{io}.$$

3. RESULTS

In this section the main results are presented. In particular sufficient conditions for the stabilization of the extinction state are derived including (i) a selection criterion for the nodes to be monitored in terms of the connectivity parameters, the infection probability and the graph topology and (ii) the design of a simple stabilizing output-feedback control scheme. Simulation results illustrate the functioning of the proposed control scheme for a scale-free network with $N = 10^6$ nodes.

3.1. Selection of Monitored and Controlled Nodes

Taking into account that $p_i(t) + q_i(t) + s_i(t) = 1$ for each $v_i \in V$ and to be consistent with the control idea, we consider that $\zeta_i(P) = \zeta_i(P, U)$, where the vector $U(t) = [u_{c1}(t), u_{c2}(t), \ldots, u_{cK}(t)]^T$, $v_{ci} \in V_C$ represent the manipulable set of parameters associated with β_i or r_{ij}. So, we can rewrite system (2) as follows

$$p_i(t+1) = (1 - \tau_i)p_i(t) + (1 - \zeta_i(P(t), U(t)))(1 - p_i(t) - q_i(t)),$$
$$q_i(t+1) = (1 - \mu_i)q_i(t) + \tau_i p_i(t), \qquad (4)$$
$$\zeta_i(P, U) = \prod_{j=1}^{N}(1 - \beta_i r_{ij} a_{ij} p_j(t)).$$

The fixed points associated with the dynamics (4) for some constant U^* (i.e., β_i and r_{ij} are set to some constant value) can be determined by substituting the relations $p_i(t + 1) =$

$p_i(t) = p_i^*$ and $q_i(t+1) = q_i(t) = q_i^*$. After some algebra it follows that

$$p_i^* = \frac{1 - \zeta_i(P^*, U^*)}{\tau_i + \left(1 + \frac{\tau_i}{\mu_i}\right)\left(1 - \zeta_i(P^*, U^*)\right)},$$

$$q_i^* = \frac{\tau_i}{\mu_i} \frac{1 - \zeta_i(P^*, U^*)}{\tau_i + \left(1 + \frac{\tau_i}{\mu_i}\right)\left(1 - \zeta_i(P^*, U^*)\right)}, \qquad (5)$$

$$s_i^* = \frac{\tau_i}{\tau_i + \left(1 + \frac{\tau_i}{\mu_i}\right)\left(1 - \zeta_i(P^*, U^*)\right)}.$$

Note that the extinction state $(p_i^* = 0, q_i^* = 0, s_i^* = 1)$ in (2), for $v_i \in V$ is a fixed point when $\zeta_i(0, U^*) = 1$. However, up to this point, it is not clear if the extinction state or any other fixed point given by (5) are stable. The extinction state means that no viruses are propagated over the network and, as we will see in the following Lemma, knowledge of the conditions under which this state is reached, will give us a clue of how viruses are propagated.

The approach followed subsequently exploits ideas from Wang et al. [4], and Bernal Jaquez et al. [15] by establishing a linear bounding dynamics for (4) that has the origin $(P, Q) = (0, 0)$ as exponentially stable fixed point.

Lemma 1. *Consider the dynamics (4) on a complex network with graph $G(V, E)$ and adjacency matrix $\mathbf{A}$. The extinction state $(P, Q) = (0, 0)$ is globally exponentially stable if the constant vector U^* (i.e., for some constant values of β_i and r_{ij}) is such that*

$$\sigma(\mathbf{H}) < 1, \quad U^* = [u_1^*, \ldots, u_N^*], \qquad (6)$$

where $\sigma(\cdot)$ is the spectral radius of the matrix $\mathbf{H}$ defined as

$$\mathbf{H} = \begin{bmatrix} \mathbf{I} - \mathbf{T} + \mathbf{BR} & \mathbf{0} \\ \mathbf{T} & \mathbf{I} - \mathbf{M} \end{bmatrix}, \qquad (7)$$

where

$$\mathbf{T} = diag(\tau_i), \quad \mathbf{B} = diag(\beta_i), \quad \mathbf{R} = [r_{ij}a_{ij}], \quad \mathbf{M} = diag(\mu_i). \qquad (8)$$

and $\mathbf{I}$ being the identity matrix.

Proof: As it is proved in Wang et al. [4], and Bernal Jaquez et al. [15], $\zeta_i(P, U^*)$ can be bounded as follows

$$1 - \zeta_i(P(t), U^*) \le \sum_{j=1}^{N} \beta_i r_{ij} a_{ij} p_j(t). \qquad (9)$$

Substituting this bound into the first equation in (4) and after some algebra one obtains

$$p_i(t+1) \le (1 - \tau_i)p_i(t) + (1 - p_i(t) - q_i(t))\sum_{j=1}^{N} \beta_i r_{ij} a_{ij} p_j(t),$$

$$\le (1 - \tau_i)p_i(t) + \sum_{j=1}^{N} \beta_i r_{ij} a_{ij} p_j(t).$$

This can be written in matrix form as

$$\begin{bmatrix} P(t+1) \\ Q(t+1) \end{bmatrix} \le \begin{bmatrix} \mathbf{I} - \mathbf{T} + \mathbf{BR} & \mathbf{0} \\ \mathbf{T} & \mathbf{I} - \mathbf{M} \end{bmatrix} \begin{bmatrix} P(t) \\ Q(t) \end{bmatrix},$$

where $P(t) = [p_1(t), \ldots, p_N(t)]^T$ and $Q(t) = [q_1(t), \ldots, q_N(t)]^T$, and the inequality being interpreted as element-wise. Therefore, the solutions of $p_i(t)$ are bounded by the linear dynamics

$$x_i(t+1) = (1 - \tau_i)x_i(t) + \sum_{j=1}^{N} \beta_i r_{ij} a_{ij} x_j(t),$$

$$w_i(t+1) = \tau_i x_i(t) + (1 - \mu_i)w_i(t), \quad v_i \in V \qquad (10)$$

i,e., forall $t \ge 0$ it holds that

$$0 \le p_i(t) \le x_i(t), \quad q_i(t) = w_i(t) \qquad (11)$$

if $p(0) = x(0), q(0) = w(0)$. The linear dynamics can be expressed in matrix form as

$$\begin{bmatrix} X(t+1) \\ W(t+1) \end{bmatrix} = \begin{bmatrix} \mathbf{I} - \mathbf{T} + \mathbf{BR} & \mathbf{0} \\ \mathbf{T} & \mathbf{I} - \mathbf{M} \end{bmatrix} \begin{bmatrix} X(t) \\ W(t) \end{bmatrix} =: \mathbf{H} \begin{bmatrix} X(t) \\ W(t) \end{bmatrix}, \qquad (12)$$

where $X(t) = [x_1(t), \ldots, x_N(t)]^T$ and $W(t) = [w_1(t), \ldots, w_N(t)]^T$. It holds that $(X^T, W^T)^T = (0^T, 0^T)^T$ is exponentially stable if and only if the eigenvalues of the associated matrix $\mathbf{H}$ are contained in the open unit circle $\mathbb{C}_1 = \{\lambda \in \mathbb{C} \,|\, |\lambda| < 1\}$. Taking into account (11) it follows that a necessary and sufficient condition for global exponential stability of $(P^T, Q^T)^T = 0$ is given by (6). $\square$

The expontial stability condition (6) is very general and does not provide any idea on how to select nodes to be controlled or monitored in order to reach the extinction state. However, using this result as a point of departure, we can get insight into the condition that every node $v_i \in V$ has to fulfill in order to ensure that $P(t)$ converges to $P^* = 0$ as shown in the following Lemma.

Lemma 2. *For a constant U^* (i.e., for some constant value for β_i and r_{ij}), the state vector $(P^T(t), Q^T(t))^T$ globally exponentially converges to the extinction state $(P^T, Q^T)^T = 0$ if for every node $v_i \in V$ it holds that*

$$\beta_i < \frac{\tau_i}{\sum_{j=1}^{N} r_{ij} a_{ij}}. \qquad (13)$$

Proof: In virtue of Lemma 1 it is sufficient to show that if the condition (13) holds, the matrix $\mathbf{H}$ defined in (6) has only eigenvalues within the open unit circle $\mathbb{C}_1 \subset \mathbb{C}$, or equivalently, that its spectral radius $\sigma(\mathbf{H}) < 1$. Note that due to the block-diagonal structure of the matrix $\mathbf{H}$ its eigenvalues are given by the eigenvalues of the two matrices on its diagonal, i.e., in terms of the matrix spectra

$$\mathcal{S}(\mathbf{H}) = \mathcal{S}(\mathbf{I} - \mathbf{T} + \mathbf{BR}) \cap \mathcal{S}(\mathbf{I} - \mathbf{M}),$$

where $\mathcal{S}(\mathbf{A})$ denotes the spectrum of a matrix $\mathbf{A}$, i.e., the union of all its eigenvalues. The eigenvalues of the matrix $\mathbf{I} - \mathbf{M}$ are contained in $\mathbb{C}_1$, given that $\mathbf{M}$ is diagonal with entries $0 < \mu_i < 1$. Thus, it remains to show that $\mathcal{S}(\mathbf{I} - \mathbf{T} + \mathbf{BR}) \subset \mathbb{C}_1$. This can be analyzed by applying Gerschgorin's Theorem [23], which provides an upper-bound estimate for the spectral radius of a given matrix. In the following, let λ represent an arbitrary eigenvalue of the matrix $\mathbf{I} - \mathbf{T} + \mathbf{BR}$. The application of this theorem to the matrix $\mathbf{I} - \mathbf{T} + \mathbf{BR}$ provides the following inequality

$$|\lambda| \leq |1 - \tau_i + \beta_i \sum r_{ij}a_{ij}| = 1 - \tau_i + \beta_i \sum r_{ij}a_{ij} \quad (14a)$$

$i = 1, \ldots, N$. Thus, $|\lambda| < 1$ is satisfied if

$$\sum_{j=1}^{N} \beta_i r_{ij} a_{ij} < \tau_i, \quad \forall v_i \in V$$

or equivalently if (13) holds true. This complete the proof. $\square$

Note that, as stated in the proof the convergence to zero of X (and thus P) is independent of the behavior of W (or accordingly Q). This resides in the fact that the dynamics represent a cascade structure. Accordingly, for the purpose of stabilizing the extinction state it is sufficient to ensure that less nodes get infected than pass into quarantine. This intuitively clear condition is exactly what is formally stated in Lemma 2.

The condition (13) gives a criterion on how to choose the nodes to be monitored. If condition (13) does not hold for some set of nodes, then it is appropriate to consider this collection as the set of nodes to be monitored V_M. Additionally, according to (13), the set of controlling nodes can be set as $V_C = V_M$. That is, we can consider that all controlled nodes are monitored. In this way, we can select β_i (or r_{ij}) as the parameter amenable for control purposes for those nodes that do not satisfy (13).

Further note that in case that the network parameters are homogenous, i.e., all nodes have the same parameter values $\mu, \beta, \tau, r_{ij} = r$, the selection criterion for a node to be controlled is directly related to its degree $\mathcal{N}_i = \sum_{i=1}^{N} a_{ij}$ (i.e., here simply the number of neighboring nodes) and can be expressed as

$$v_i \in V_C \,\forall\, i: \mathcal{N}_i > \frac{\tau}{r\beta}.$$

Nevertheless, when the parameter distribution is non-homogenous, it is possible that nodes with high degree do not have to be controlled and nodes with small degree have to be, given the particular constellation between τ_i, β_i and r_{ij} according to condition (13). This is illustrated in **Figure 2** for a scale-free network with $N = 100$ nodes and normally distributed parameters.

Note that in the notion of constructive control theory (see e.g., [19, 24]), the dynamics of the nodes that are not controlled establishes the zero dynamics (i.e., the dynamics resulting from the restriction that $p_i = 0$ for all i such that $v_i \in V_c$) and by construction this dynamics is asymptotically stable. Therefore, the zero dynamics correspond to a spreading process over a reduced network, from which the monitored (and thus controlled) nodes have been withdrawn, given that for $p_i(t) \equiv 0$ the node v_i does no longer interact with its neighbors.

3.2. Feedback Control Design

The question addressed in this section is how to design the feedback control for the nodes $v_j \in V_C$ so that $\lim_{t \to \infty} |p_j(t) - p_j^*| = 0$. Up to this point, we have not considered the dependency of $\zeta(P, U)$ on U, given that U was considered as a set of constant parameters. This dependency will permit to explicitly determine a control law $u_i(t) = \beta_i(t)$ that steers the nodes $v_j \in V_C$ to their desired values $p_j^* = 0$. With this idea, the following theorem is established

Theorem 1. *Consider the dynamics given by (4) and consider the nodes that do not satisfy (13) as the set V_C of nodes to be controlled. Let $V_M = V_C$, that is, all controlled nodes are monitored. If the controls $u_i(t)$ satisfy*

$$0 \leq u_i(t) < \bar{\beta}_i = \frac{\tau_i}{\sum_{j=1}^{N} r_{ij}a_{ij}}, \quad (15)$$

then $(P^T, Q^T)^T = 0$ is exponentially stable.

Proof: The Theorem can be easily proven as follows. Let $\bar{\beta}_i$ be given as in (15). It follows that

$$0 \leq u_i(t) < \bar{\beta}_i,$$
$$\Leftrightarrow \quad 0 \leq u_i(t)r_{ij}a_{ij} < \bar{\beta}_i r_{ij}a_{ij},$$
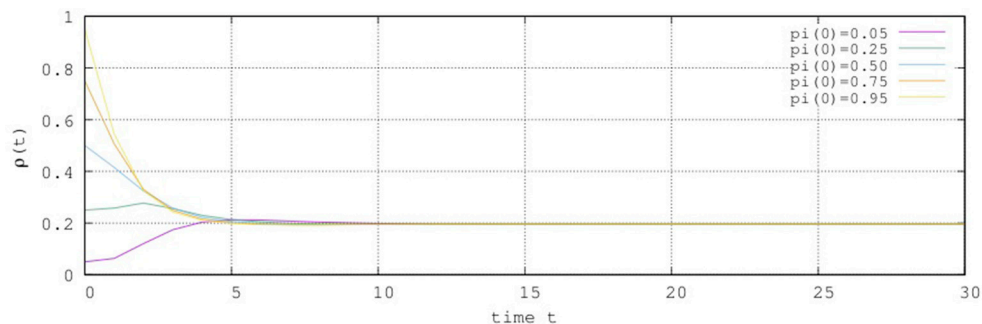
FIGURE 3 | Behavior of $\rho(t)$ for several initial conditions without control.

$$\Leftrightarrow \quad 0 \le \sum_{j=1}^{N} u_i(t) r_{ij} a_{ij} < \sum_{j=1}^{N} \bar{\beta}_i r_{ij} a_{ij}.$$

Given that for $\beta_i \le \bar{\beta}_i$ inequality (13) is satisfied, one obtains

$$0 \le \sum_{j=1}^{N} u_i(t) r_{ij} a_{ij} < \sum_{j=1}^{N} \bar{\beta}_i r_{ij} a_{ij} < \tau_i,$$

or equivalently

$$\sum_{j=1}^{N} u_i(t) r_{ij} a_{ij} < \tau_i$$

implying that inequality (13) is satisfied. In consequence, the exponential stability of $(P^T, Q^T)^T = 0$ follows. □

Note that following a similar reasoning, a control law can be established for the case that $u_i(t) = r_{ij}(t)$. Nevertheless, this approach will not be explicitly elaborated at this place.

From Theorem 1, we have that in order to stabilize the extinction state for the dynamics (4) it is sufficient to design feedback controls $u_i(t)$ for all nodes $v_i \in V_C$, which take values below the upper-bound $\bar{\beta}_i$ defined in (15). For example, a simple linear feedback control given by

$$u_i(t) = \beta_i(t) = \gamma \bar{\beta}_i (1 - p_i(t)), \quad \gamma \in (0, 1) \tag{16}$$

does satisfy this condition. Note that the advantage of using a time varying control law $\beta_i(t)$ consists in actively adapting the infection probability on the actual needs, i.e., the actual network state. In comparison with imposing a constant value for β_i this possibly enables to optimize the control effort over time, because according to (16) $\beta_i \rightarrow \gamma \bar{\beta}_i$ with $p_i \rightarrow 0$. The performance using this simple output-feedback control is illustrated in the subsequent section.

3.3. Simulations
In order to verify our results, we perform several simulations of the dynamical system (4), for different initial conditions, with the following considerations

- We used a scale-free network of $N = 10^6$ described by $G(V, E)$, that incorporates preferential attachment according to Barabási and Albert [21]. We started with a small number $m_o = 9$ of vertices, linked randomly, and at every step we add a new node or vertex with $m = 3$ edges until we reach $N = 10^6$. We emphasize that our results are independent of the network's topology.
- To facilitate our simulations, we consider that $r_{ij} = r_{ji}$.
- The constant values for the recovery probability μ_i, the probability of infection β_i, the internment probability τ_i, and the contact probability r_i were distributed uniformly over the nodes with values in the interval $[0.2, 0.7]$.

FIGURE 4 | Behavior of the zero dynamics.

FIGURE 5 | Closed-loop network behavior with the linear feedback control given by (18).

- To show and corroborate our results, we calculated the average probability $\rho(t)$ given by

$$\rho(t) = \frac{1}{N} \sum_{i=1}^{N} p_i(t). \tag{17}$$

- The sets V_C and V_M are chosen according to Lemma 2 as the sets of nodes that do not satisfy (13). For the case considered here 424569 nodes out of $N = 10^6$ (i.e., about 42.5%) have to monitored and controlled.

Simulation studies have been performed considering the following scenarios:

I. Absence of control: All parameters of the system (4) have constant values which not necessarily satisfy condition (13).
II. Zero dynamics: The output was constrained to $Y(t) = 0$ by setting $\beta_j = 0$ and $p_j(0) = 0$ for all $v_j \in V_M = V_C$.
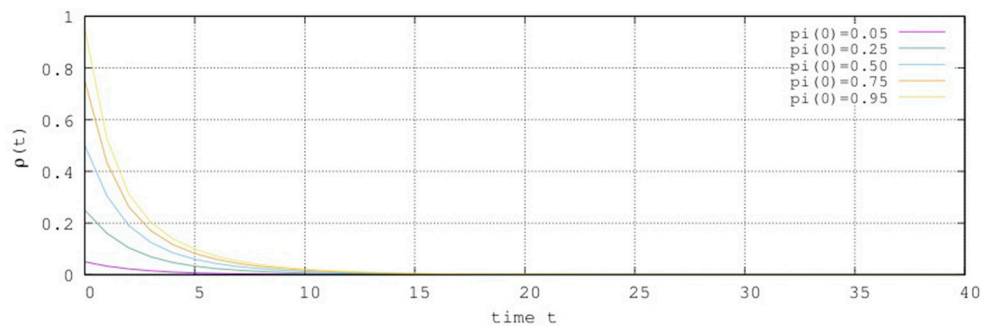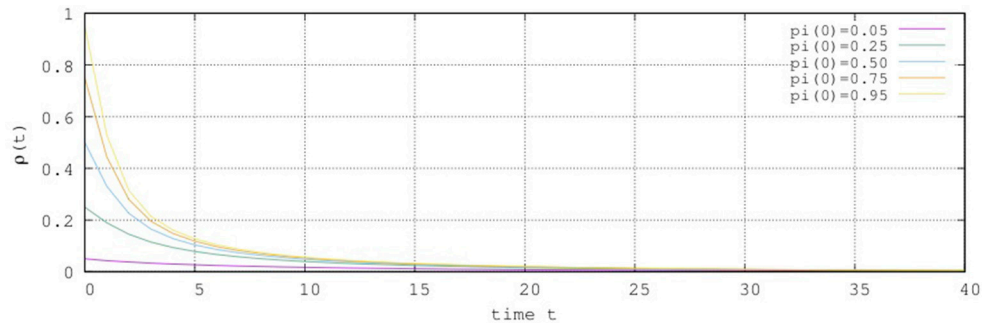III. Linear feedback control: The linear feedback control given by (16) has been implemented for all $v_i \in V_M = V_C$.

I. Absence of Control

In the absence of control the network reaches an endemic attractor state where about a 20% of the nodes are probably infected after about 20 time units as can be seen in **Figure 3**.

II. Zero Dynamics

To corroborate that the zero dynamics are exponentially stable, we set $Y(t) = 0$ for those nodes that do not satisfy (13). This is achieved by setting $\beta_i = 0$ for $v_i \in V_M$.

Figure 4 shows the behavior of the nodes associated with the zero dynamics. Note that the extinction state is reached after 18 time steps.

III. Linear Feedback Control

To show the performance of the proposed simple output-feedback control scheme (16), it has been implemented for all monitored and controlled nodes with the gain $\gamma = 0.9$ and the upper bound $\bar{\beta}_i$ calculated according to (15). This yields the output-feedback controller.

$$\beta_i(t) = \frac{\gamma \tau_i (1 - p_i(t))}{\sum_{j=1}^{N} r_j a_{ij}}, \quad \text{for} \quad \gamma = 0.9 \quad \text{and} \quad \forall v_i \in V_C = V_M. \tag{18}$$

Note that the control (18), depends on the state of the node i given by $p_i(t)$, and the properties of its neighbors given by $\bar{\beta}_i$. **Figure 5** shows the result of the simulation with the applied control (18). As predicted in Theorem 1 the extinction state is a close-loop attractor, and is reached in about 40 time steps.

4. DISCUSSIONS AND CONCLUSIONS

The problem of deciding which nodes in a complex network with quarantine should be controlled and how to control them in order to achieve that in the closed-loop system the extinction state becomes a (global) attractor has been studied. Sufficient conditions for virus extinction have been derived using a constructive control approach by suitably identifying the zero dynamics according to a threshold condition for the stability of the extinction state. The associated node selection criterion does depend on the transmission probability between the nodes and their neighbors, the degree of each node and its probability to pass into quarantine. It has been shown that in spite of the strongly nonlinear dynamics of the spreading process the extinction state can be efficiently stabilized using a simple linear bounded output-feedback control if the nodes to be controlled are selected according to the proposed scheme. The performance and behavior of the spreading process without and with control has been illustrated for a scale-free network with $N = 10^6$ nodes.

AUTHOR CONTRIBUTIONS

All authors designed and did the research, in particular the stability analysis. LA-R and AS conceived and designed the control and the numerical experiments. LA-R performed the numerical experiments. All authors analyzed the data. LA-R and RB wrote the paper. All authors reviewed the manuscript. All authors have read and approved the final manuscript.

FUNDING

REFERENCES

1. Wang W, Tang M, Zhang HF, Gao H, Do Y, Liu ZH. Epidemic spreading on complex networks with general degree and weight distributions. *Phys Rev E* (2014) **90**:042803. doi: 10.1103/PhysRevE.90.042803
2. Leventhal GE, Hill AL, Nowak MA, Bonhoeffer S. Evolution and emergence of infectious diseases in theoretical and real-world networks. *Nat Commun.* (2015) **6**:6101. doi: 10.1038/ncomms7101
3. Tomovski I, Trpevski I, Kocarev L. Topology independent SIS process: An engineering viewpoint. *Commun Nonlinear Sci Numer Simul.* (2014) **19**:627–37. doi: 10.1016/j.cnsns.2013.06.033

4. Wang Y, Chakrabarti D, Wang C, Faloutsos C. Epidemic spreading in real networks: an eigenvalue viewpoint. In: *22nd International Symposium on Reliable Distributed Systems, 2003.* Florence. (2003). p. 23–5.
5. Chakrabarti D, Wang Y, Wang C, Leskovec J, Faloutsos C. Epidemic thresholds in real networks. *ACM Trans Inf Syst Secur.* (2008) **10**:1–26. doi: 10.1145/1284680.1284681
6. Gómez S, Arenas A, Borge-Holthoefer J, Meloni S, Moreno Y. Discrete-time Markov chain approach to contact-based disease spreading in complex networks. *Europhys Lett.* (2010) **89**:38009. doi: 10.1209/0295-5075/89/38009

7. Gómez S, Arenas A, Borge-Holthoefer J, Meloni S, Moreno Y. Probabilistic framework for epidemic spreading in complex networks. *Int J Complex Syst Sci.* (2011) **1**:47–54. Available online at: http://www.ij-css.org/volume-01_01/ijcss01_01-047.pdf

8. Meloni S, Arenas A, Gomez S, Borge-Holthoefer J, Moreno Y. Modeling Epidemic Spreading in Complex Networks: Concurrency and Traffic. In: Thai MT, Pardalos PM, editors. *Handbook of Optimization in Complex Networks.* New York, NY: Springer (2012). p. 435–62.

9. Moreno Y, Nekovee M, Pacheco AF. Dynamics of rumor spreading in complex networks. *Phys Rev E* (2004) **69**:066130. doi: 10.1103/PhysRevE.69.066130

10. Borge-Holthoefer J, Moreno Y. Absence of influential spreaders in rumor dynamics. *Phys Rev E* (2012) **85**:026116. doi: 10.1103/PhysRevE.85.026116

11. Singh A, Singh YN. Rumor Spreading and Inoculation of Nodes in Complex Networks. In: *Proceedings of the 21st International Conference on World Wide Web. WWW '12 Companion.* New York, NY: ACM (2012). p. 675–8.

12. Xie M, Jia Z, Chen Y, Deng Q. Simulating the spreading of two competing public opinion information on complex network. *Appl Math.* (2012) **3**:1074–8. doi: 10.4236/am.2012.39158

13. Mc'Kendrick AG. Applications of Mathematics to Medical Problems. *Proc Edinburgh Math Soc.* (1925) **44**:98–130.

14. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics. II. —The problem of endemicity. *Proc R Soc Lond A Math Phys Eng Sci.* (1932) **138**:55–83.

15. Bernal Jaquez R, Schaum A, Alarcon L, Rodriguez C. Stability analysis for virus spreading in complex networks with quarantine. *Publ Mat Uruguay* (2013) 14:221–33. Available online at: http://pmu.uy/pmu14/pmu14.pdf

16. Chimmalee B, Sawangtong W, Wiwatanapataphee B. The effects of community interactions and quarantine on a complex network. *Cogent Math.* (2016) **3**:1249141. doi: 10.1080/23311835.2016.1249141

17. Alarcon-Ramos LA, Schaum A, Rodriguez Lucatero C, Bernal Jaquez R. Stability analysis for virus spreading in complex networks with quarantine

18. Schaum A, Bernal Jaquez R. Estimating the state probability distribution for epidemic spreading in complex networks. *Appl Math Comput.* (2016) 29:197–206. doi: 10.1016/j.amc.2016.06.037

19. Monaco S, Normand-Cyrot D. Zero dynamics of sampled nonlinear systems. *Syst Control Lett.* (1988) **11**:229–34.

20. Isidori A. *Nonlinear Control Systems.* London: Springer (2000).

21. Barabási AL, Albert R. Emergence of Scaling in Random Networks. *Science* (1999) **286**:509–12.

22. Nowzari C, Preciado VM, Pappas GJ. Analysis and control of epidemics: a survey of spreading processes on complex networks. *IEEE Control Syst.* (2016) 36:26–46. doi: 10.1109/MCS.2015.2495000

23. Cullen CG. *Matrices and Linear Transformations.* New York, NY: Dover Publications (1990).

24. Byrnes CI, Lin W. Losslessness, feedback equivalence, and the global stabilization of discrete-time nonlinear systems. *IEEE Trans Automatic Control* (1994) 39:83–98.

Synchronization, Oscillator Death, and Frequency Modulation in a Class of Biologically Inspired Coupled Oscillators

Alessio Franci[1]*, Marco Arieli Herrera-Valdez[1], Miguel Lara-Aparicio[1] and Pablo Padilla-Longoria[2]*

[1] Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico,
[2] Departamento de Matemáticas y Mecánica, Instituto de Investigación en Matemáticas Aplicadas y Sistemas, Universidad Nacional Autónoma de Mexico, Mexico City, Mexico

The general purpose of this paper is to build up on our understanding of the basic mathematical principles that underlie the emergence of synchronous biological rhythms, in particular, the circadian clock. To do so, we study the role that the coupling strength, coupling type, and noise play in the synchronization of a system of coupled, non-linear oscillators. First, we study a deterministic model based on Van der Pol coupled oscillators, modeling a population of diffusively coupled cells, to find regions in the parameter space for which synchronous oscillations emerge and to provide conditions under which diffusive coupling kills the synchronous oscillation. Second, we study how noise and coupling interact and lead to synchronous oscillations in linearly coupled oscillators, modeling the interaction between various pacemaker populations, each having an endogenous circadian clock. To do so, we use the Fokker-Planck equation associated to the system. We show how coupling can tune the frequency of the emergent synchronous oscillation, which provides a general mechanism to make fast (ultradian) pacemakers slow (circadian) and synchronous via coupling. The basic mechanisms behind the generation of oscillations and the emergence of synchrony that we describe here can be used to guide further studies of coupled oscillations in biophysical non-linear models.

Keywords: synchronization, oscillator death, circadian rhythm, coupled non-linear oscillators, Fokker-Planck

1. INTRODUCTION

The study of circadian rhythms has been a subject of great interest for a long time. The majority of the first studies were mainly based on observations in plants [1–4]. The study of circadian rhythms from a mathematical perspective reached a milestone with the work of Kalmus and Wigglesworth, a biologist and a mathematician, respectively, who associated of circadian rhythms to the existence of a limit cycle, using a hydraulic system as analogy. Kalmus and Wigglesworth presented their work entitled "Shock excited system as models for biological rhythms" along with several mathematical models of circadian rhythms at a Symposium on Biological Rhythms carried out at Cold Spring Harbor in the United States of America in 1960 [5–7]. Lots of other important works on circadian rhythms were presented in this symposium, but the work

of Kalmus and Wigglesworth was key in establishing a better mathematical formalism for the study of circadian rhythms. Although many researchers followed the theoretical path proposed by Kalmus and Wigglesworth, the mathematical study of circadian rhythms was finally established by Arthur Winfree (biologist and mathematician), who introduced topology for the description of several aspects of circadian rhythms. An excellent summary of many of the earlier works can be found in Arthur Winfree's master book entitled "The Geometry of Biological Time" [8]. The number of studies about biological rhythms at large has increased greatly in the last two decades, in part due to new technological advances. Particularly related to circadian rhythms, there is now evidence that there is rhythmic patterns of activity at the molecular [9, 10], cellular [11, 12], tissue [13, 14], and systems levels [15–17], and that circadian regulation is involved in jointly regulating activity in all those different levels of biological organization [18–20], and also, taking into account interactions with the environment [21] and perturbations induced by behavior [22, 23].

Mathematical modeling and experimental characterizations of different properties of circadian rhythms have been combined to produce explanations and hypotheses about the rhythmicity in biological phenomena [24–28]. Of particular interest, the ontogeny of circadian rhythm in the crayfish has been studied by Lara-Aparicio et al. [29] combining theoretical and experimental perspectives, building phenomenological mathematical models that capture a series of experimental results involving the synchronization of electro-retinogram activity in crayfish [30–32]. These models couple two van der Pol oscillators [33] represented by state vectors, and include parameters representing the frequency of the oscillators, the radii of the limit cycles, and the first coordinate of the center of the limit cycle. The system is setup such that one oscillator is driving the behavior of the other oscillator, but not vice-versa. One of the main findings with this model is that the driving oscillator induces an Andronov-Hopf bifurcation in the driven oscillator and regulates its frequency.

The model by Aparicio et al. simulates, explains, and has suggested new biological experiments, it is simple enough to allow analytical approaches, and it has provided useful insights about questions referring to the ontogeny of the circadian rhythm in crayfish from the childhood to adult stages. For instance, a hypothesis about the existence of a hormone, which was experimentally detected, was generated from the model. The model also allowed Lara-Aparicio et al. to study synchronization of circadian rhythms with external signals like day and night light cycle [34]. By studying basic principles underlying the generation of oscillations in coupled non-linear systems, these researchers were able to conjecture that circadian rhythms can result from coupling systems of cells, each one oscillating with an ultradian oscillation [35–37]. Synchronization among cells emerges naturally as a motivating theme that has been studied through systems of non-linear Equations [38] representing n oscillators with the classical van der Pol non-linear damping for the terms responsible for the oscillatory dynamics.

In the present paper, we extend the work in Lara-Aparicio et al. [29] and Barriga-Montoya et al. [38] by analyzing two qualitative mathematical models, each capturing a different level of organization in the ontogeny of circadian rhythms. Inspired by gap junction coupling between neurons, or similarly, by chemical coupling in self oscillating networks, we study the bifurcation structure in a deterministic model assuming that the coupling between the oscillators is diffusive. The resulting dynamics resemble neuronal activity at the cellular level. Then, using graph theoretical methods and center manifold theory, we show that synchronous oscillations appear via a Hopf bifurcation in a population of pacemaker oscillators. In this case, the bifurcation parameter is thought of as a representative of the developmental stage of the neurons. We further explore the phenomenon of oscillator death: although the single neurons are endogenously oscillating for sufficiently large values of the bifurcation parameter, the population oscillation dies for sufficiently large coupling, which suggests that the weak coupling hypothesis must be satisfied for robust synchronous oscillations to occur. In the case of all-to-all coupling, we provide necessary conditions for oscillator death to occur and leave the derivation of sufficient conditions to a future report.

Frequency modulation is also an important phenomenon that can be studied with these models. In consideration of the results from the first analysis, we shift our focus to the frequency modulations that emerge as a result of the interconnection of various circadian pacemaker populations. In doing so, we estimate the synchronization frequency as a function of the intrinsic frequencies of the oscillators, their coupling strength, and the topology of the network. To do so, we construct a second model that can be thought of as a stochastic, larger-scale extension of the first model we present. In this case, each population is assumed to be an endogenous oscillator and the coupling is assumed to be linear. Using linear stochastic analysis and under the assumption that the population oscillations are synchronized, we derive a scalar Fokker-Plank equation. The model captures an important feature of circadian rhythm ontogeny: the emergence of low frequency (circadian) oscillations from coupled high frequency (ultradian) oscillators [35–38]. Future work will aim at deriving conditions on the intrinsic frequencies, the coupling strength, and the network topology, that ensure synchronization of the population oscillations.

It is worth noticing that, although the motivation for the present paper is centered around circadian rhythms, the model captures more general phenomena. Our findings include that in diffusively coupled cells resting node dynamics imply global asymptotic stability, oscillating node dynamics imply global synchronization for small coupling, and multistability between oscillator death and global synchronization for large coupling. In stochastic, linearly coupled populations, we describe the dynamical mechanisms through which coupling modulates the frequency of the synchronous oscillation. To the author's knowledge, both phenomena are new from a non-linear collective phenomena perspective. Among other reasons, it is surprising that passive coupling like diffusive coupling could kill an oscillation and create multistability. Similarly, we are not aware of any work providing mechanistic explanations on how coupling can tune a global oscillation frequency.

The paper is organized as follows. In section 2, we present and analyze the model of diffusively coupled oscillators. Two theorems are proved about global asymptotic stability for resting cells and global synchronization for oscillating cells and weak coupling. We then derive sufficient conditions for diffusive coupling-induced oscillator death. In section 3 we present and analyze the model of linearly coupled oscillators. In particular, we derive an explicit formula for the emergent synchronization frequency as a function of the coupling topology and oscillator natural frequencies. Finally we discuss the presented results in section 4.

2. GLOBAL SYNCHRONIZATION AND OSCILLATOR DEATH IN DIFFUSIVELY COUPLED OSCILLATORS

We regard a network of N coupled oscillators as a directed graph $\mathcal{G}$ with N vertices, with a network topology codified by a matrix $A = [a_{ij}]_{i,j=1}^{N}$, where $a_{ij} \geq 0$ represent connection weights. If oscillator i receives signals from oscillator j, then the graphical representation of $\mathcal{G}$ has an arrow from j to i, and $a_{ij} > 0$. If $a_{ij} > 0$, the signal received by oscillator i from oscillator j is $a_{ij}(x_j - x_i)$. Assume that the dynamics for each oscillator satisfies the following coupled oscillator dynamics

$$\dot{x}_i = y_i + \mu \sum_{j=1}^{N} a_{ij}(x_j - x_i) \tag{1a}$$

$$\dot{y}_i = \left(\lambda - x_i^2 - \frac{y_i^2}{\omega^2}\right) y_i - \omega_i^2 x_i \tag{1b}$$

where μ is the global coupling strength, which uniformly scales the coupling weights a_{ij}.

In the absence of coupling the oscillator dynamics reduces to the modified van der Pol equation

$$\ddot{x}_i = \left(\lambda - x_i^2 - \frac{y_i^2}{\omega_i^2}\right) \dot{x}_i - \omega_i^2 x.$$

These equations define a simple dynamical system that can transition between global asymptotic stability and almost global convergence to a hyperbolic limit cycle through variations of the control parameter $\lambda \in \mathbb{R}$, providing a simple model for various biological systems that exhibit the same transition, in particular neurons and molecular oscillators. For $\lambda < 0$ the non-linear dissipation coefficient $-\left(\lambda - x^2 - \frac{y^2}{\omega_i^2}\right)$ is always positive, which leads to damped oscillations. For $\lambda > 0$ the dissipation coefficient becomes negative close to the origin, which leads to sustained oscillations through a Hopf bifurcation. A generic trajectory belonging to the family of periodic orbits born at the Hopf bifurcation has the form

$$\sqrt{\lambda}(\cos(\omega_i t + \theta_0), \sin(\omega_i t + \theta_0)), \tag{2}$$

where the θ_0 is the initial phase.

In the presence of coupling, Equations (1) represent a generic network of diffusively coupled non-linear oscillators. As

mentioned earlier, the diffusive form of the coupling can be thought of as gap junction coupling in a neuronal population, or diffusive coupling between non-linear molecular oscillators. In both interpretations, the graph topology is necessarily undirected, that is, $a_{ij} = a_{ji}$. However, the mathematical results presented in this section hold under more general assumptions and we present them in the general form.

2.1. Global Synchronization

We start by recalling some basic graph-theoretical definitions and facts. The graph $\mathcal{G}$ is said to be *strongly connected* if for each pair of nodes in $\mathcal{G}$, there exists a directed path between them. $\mathcal{G}$ is *balanced* if $\sum_{j=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ji}$ for all i.

The Laplacian matrix $L = [L_{ij} : i, j = 1, ..., n]$ for the graph $\mathcal{G}$ is such that $L_{ij} = -a_{ij}$ if $i \neq j$, and $L_{ii} = \sum_{j=1}^{N} a_{ij}$. Note that the vector of ones is always a right null eigenvector of L and zero is always an eigenvalue of L ($L1_N = 0$). It can be shown that a graph is strongly connected if and only if zero is a simple eigenvalue of the Laplacian matrix [39]. Obviously, symmetric graphs (i.e., satisfying $a_{ij} = a_{ji}$) are balanced, but the converse is not true. Consider, for instance, a directed ring.

The global behavior of the system (1) for $\lambda < 0$ and $\mu \geq 0$, for a network with connectivity represented by a generic balanced graph is characterized by the next theorem (**Figure 1**).

Theorem 2.1. *Assume that the graph $\mathcal{G}$ is balanced and that $\omega_i = \omega_j = \omega$ for all $i, j = 1, \ldots, N$. If $\lambda < 0$, then the origin is globally asymptotically stable and locally exponentially stable for all $\mu \geq 0$.*

Proof. We consider the quadratic Lyapunov function

$$V(x, y) = \sum_{i=1}^{N} (x_i^2 + y_i^2/\omega^2).$$

The derivative of V along the trajectories of the system (1) gives

$$\dot{V} = \sum_{i=1}^{N} \left(2x_i \dot{x}_i + 2y_i \dot{y}_i/\omega^2\right) \tag{3}$$

which after substitution of the derivatives $\dot{x}_i$ and $\dot{y}_i$, can be bounded from above as follows

$$= \sum_{i=1}^{N} \left(2x_i y_i - 2y_i \frac{\omega^2 x_i}{\omega^2} + 2\frac{y_i}{\omega^2}\left(\lambda - x_i^2 - \frac{y_i^2}{\omega^2}\right) y_i \right.$$
$$\left. -2\mu x_i \sum_{j=1}^{N} a_{ij}(x_i - x_j)\right)$$

$$\leq 2\frac{\lambda}{\omega^2} \sum_{i=1}^{N} y_i^2 - 2\mu \sum_{i,j=1}^{N} x_i a_{ij}(x_i - x_j)$$

$$= 2\frac{\lambda}{\omega^2} \sum_{i=1}^{N} y_i^2 - 2\mu \sum_{i,j=1}^{N} a_{ij}(x_i^2 - x_j x_i)$$

$$= 2\frac{\lambda}{\omega^2} \sum_{i=1}^{N} y_i^2 - 2\mu \sum_{i,j=1}^{N} a_{ij}(x_i^2/2 - x_j x_i + x_i^2/2)$$
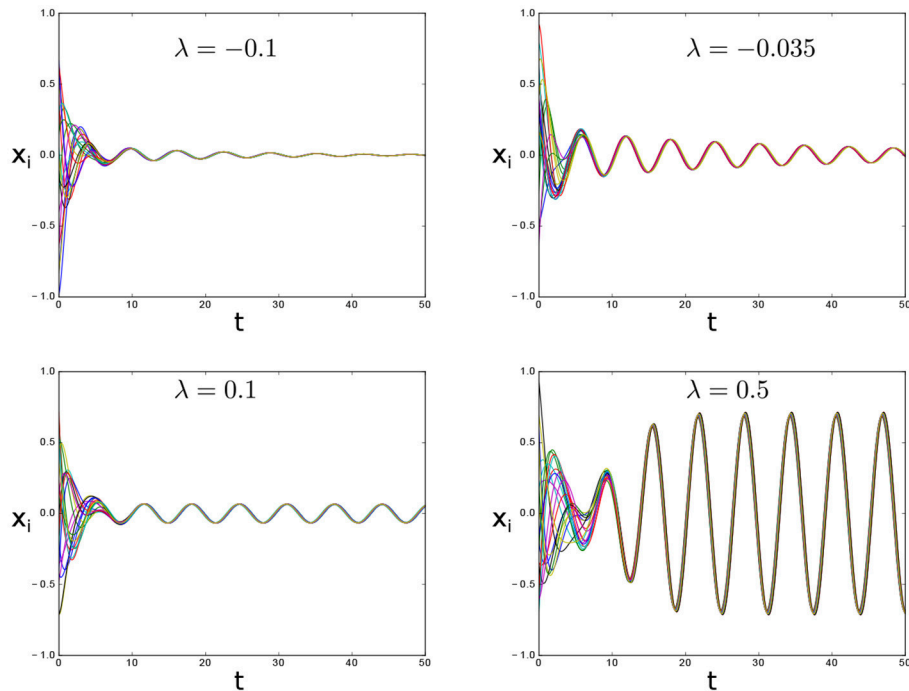
FIGURE 1 | Transition between global asymptotic stability and synchronous oscillation via Hopf bifurcation in system (1) for $\mu = 0.05$, $N = 20$, natural frequencies uniformly distributed in the interval $[0.9, 1.1]$ and varying λ. The interconnection topology is all-to-all.

To show that the last term is negative definite, we use the balanced interconnection hypothesis. Let $d_j = \sum_{i=1}^{N} a_{ij} = \sum_{i=1}^{N} a_{ji}$. Then,

$$\sum_{i,j=1}^{N} a_{ij}x_j = \sum_{j=1}^{N} d_j x_j = \sum_{i=1}^{N} d_i x_i = \sum_{i,j=1}^{N} a_{ij}x_i.$$

As a consequence,

$$\dot{V} \leq \frac{\lambda}{\omega^2}\sum_{i=1}^{N} y_i^2 - \mu \sum_{i,j=1}^{N} a_{ij}(x_i - x_j)^2 \leq \frac{\lambda}{\omega^2}\sum_{i=1}^{N} y_i^2.$$

Then the global part of statement follows by LaSalle invariance principle [40]. The local part follows by observing that for $\lambda < 0$ the linearization of model (1) at the origin is non-singular and therefore asymptotic stability of the origin implies that all eigenvalues have negative real part. □

Remark. Because exponentially stability implies robustness to small perturbations, Theorem 2.1 remains true for small heterogeneity in the natural frequencies.

Note that, for $\lambda < 0$, the system in Equation (1) exhibits exponentially damped oscillations toward the origin (**Figure 1**, top panels).

Next we show that, at $\lambda = 0$ and identical natural frequencies model (1) undergoes a supercritical Hopf bifurcation inside the consensus space

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^{2N} : x_i = x_j, y_i = y_j, \forall i, j = 1, \dots, N\},$$

provided the graph is strongly connected. The linearization of the system (1) is given by

$$J = \begin{bmatrix} -\mu L & I_N \\ -\omega^2 I_N & \lambda I_N \end{bmatrix}, \tag{4}$$

where I_N is the N-dimensional identity matrix and L is the network Laplacian defined in section 2.1. Let 1_N be the N-dimensional vector of ones. Given a (complex) vector $\nu = (v, w)$ in the consensus space $\mathcal{C}$, that is, $v = a1_N$ and $w = b1_N$ for some $a, b \in \mathbb{C}$, the eigenvalue problem for the Jacobian matrix Equation (4), restricted to the consensus space, takes the form

$$\begin{aligned} J\nu &= J\begin{bmatrix} a1_N \\ b1_N \end{bmatrix} \\ &= \begin{bmatrix} -\mu L a1_N + b1_N \\ -\omega^2 a + \lambda b1_N \end{bmatrix} \\ &= \begin{bmatrix} b1_N \\ (-\omega^2 a + \lambda b)1_N \end{bmatrix} \\ &= \xi \begin{bmatrix} a1_N \\ b1_N \end{bmatrix} \end{aligned}$$

where ξ is a (complex) eigenvalue. In the third equality we used the fact that 1_N is a right null eigenvector of L. The last equality shows that $\nu \in \mathcal{C}$ is an eigenvector of J with eigenvalue ξ if and only if its components a and b satisfy

$$\begin{bmatrix} 0 & 1 \\ -\omega^2 & \lambda \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \xi \begin{bmatrix} a \\ b \end{bmatrix}. \tag{5}$$

We can now easily solve Equation (5) to obtain the eigenvalues/eigenvectors pairs

$$\xi_\pm(\lambda) = \frac{\lambda}{2} \pm \frac{1}{2}\sqrt{\lambda^2 - 4\omega^2}, \quad \begin{bmatrix} a_\pm \\ b_\pm \end{bmatrix} = \begin{bmatrix} \frac{1}{\omega^2}(\lambda - \xi_\pm) \\ 1 \end{bmatrix} \quad (6)$$

For $\lambda = 0$ we got two purely imaginary eigenvalues which correspond to a supercritical Hopf bifurcation of model (1) inside the consensus space, as summarized in the following theorem and illustrated in (**Figure 1**, bottom panels).

Theorem 2.2. *For almost all balanced, strongly connected interconnection topologies the following holds. For all $\mu > 0$, the system (1) undergoes a supercritical Hopf bifurcation at $\lambda = 0$ with center manifold given by the consensus space $\mathcal{C}$. Moreover, the family of periodic solutions born at the Hopf bifurcation are exponentially asymptotically stable and correspond to synchronous oscillations of the oscillator network.*

Proof. By Theorem 2.1 the origin is locally exponentially stable for $\lambda < 0$. We further observe that, if the interconnection topology is strongly connected, then zero is a simple eigenvalue of L and therefore no either eigenvalue of J satisfies the same eigenvalue problem defined by Equation (5). It then follows by the center manifold theorem [41] and Equation (6), that the system (1) possesses a two-dimensional center manifold $\mathcal{W}^c$ that is tangent to the consensus space $\mathcal{C}$, for $\lambda = 0$. Moreover, this center manifold is exponentially attractive. By the Hopf bifurcation theorem [42], Equation (6) also implies that the system Equation (1) undergoes a supercritical Hopf bifurcation inside $\mathcal{W}^c$ when λ crosses zero from negative to positive. By direct substitution inside the model equations, we see that along a generic member of the family of periodic orbits born at the Hopf bifurcation, oscillators are synchronously oscillating with each oscillator orbit given by (2). □

Remark 2.3. *Because Hopf bifurcation is codimension-zero (in the sense of [43]), it is persistent under small perturbations, which ensures that Theorem 2.2 remains true for small heterogeneity in the natural frequencies.*

2.2. Oscillator Death and Multi-Stability for Stronger Coupling

We now explore the phenomenon of "oscillator death," induced by strong coupling in model (1). We restrict our attention to the all-to-all coupling case, i.e., $a_{ij} = 1$ for all $i \neq j$. For $\lambda > 0$ and μ sufficiently small the synchronous oscillations born at Hopf bifurcation (Theorem 2.2) attract all trajectories. However, the system is *multistable*, as can be noted from the fact that increasing μ leads to the appearance of a family of steady states that attract some of the trajectories, but the synchronous periodic orbits remain locally exponentially stable (**Figure 2**). Indeed, depending on the initial conditions, only some trajectories converge to the synchronous oscillations. In the following we will provide geometric insights, without formal proof, about the mechanisms underlying oscillator death and multi-stability in model (1).

We start by observing that the oscillator death state is characterized by the presence of two dead oscillator clusters.

Inside each cluster, oscillators converge to the same steady state. To analyze the appearance of oscillator death steady-states, we can simplify the model by assuming that $(x_i, y_i) = (x_1, y_1)$ for all $i = 1, \ldots, N_1$ and $(x_i, y_i) = (x_2, y_2)$ for all $i = 1, \ldots, N_2$, where $N_1, N_2 < N$, $N_1 + N_2 = N$, are the cluster sizes. The pairs (x_1, y_1) and (x_2, y_2) define the cluster states.

The cluster state dynamics can be easily derived and read

$$\dot{x}_1 = y_1 + \mu N_2(x_2 - x_1), \quad (7a)$$

$$\dot{y}_1 = -\omega^2 x_1 + \left(\lambda - x_1^2 - \frac{y_1^2}{\omega^2}\right) y_1, \quad (7b)$$

$$\dot{x}_2 = y_2 + \mu N_1(x_1 - x_2), \quad (7c)$$

$$\dot{y}_2 = -\omega^2 x_2 + \left(\lambda - x_2^2 - \frac{y_2^2}{\omega^2}\right) y_2. \quad (7d)$$

Each cluster state dynamics has the form

$$\dot{x} = y + \mu N_j(x_j - x), \quad (8a)$$

$$\dot{y} = -\omega^2 x + \left(\lambda - x^2 - \frac{y^2}{\omega^2}\right) y, \quad (8b)$$

where N_j is the other cluster size and x_j the other cluster state. A sufficient condition for the appearance of multiple steady states is that there must exist values of x_j for which the model(8) has multiple steady-states. This condition can easily be verified by analyzing the dependence of the intersection of the nullclines of model (8) as a function of x_j and the parameters μ and λ (**Figure 3**).

If the coupling strength is too small the origin is the only steady state (**Figure 3**). This steady state is unstable and all trajectories are attracted toward the synchronous periodic orbit. However, new steady states appear for larger values of μ. The critical value of μ for which the new steady-states appear can be found by computing the slope of the nullclines at the origin. The slope of the x-nullcline is evidently μN_j. The slope of the x-nullcline can be computed by implicit differentiation and is given by $\frac{\omega^2}{\lambda}$. Multiple steady-states appear if $\mu > \frac{\omega^2}{N_j\lambda}$ (**Figure 2**).

3. SYNCHRONIZATION AND FREQUENCY MODULATION IN LINEARLY COUPLED OSCILLATORS

In this section we present an alternative approach to study synchronization under the influence of noise using the Fokker-Planck equation (FPE). The modeling in this section can be thought of in the context of interacting populations of oscillators. Related work in the context of populations of synchronized neurons can be found in the work by Jiao et al. [44]. Introducing noise in the equation is natural from the biological perspective. However, even in the absence of noise, the introduction of random perturbations allows the extraction of information about the deterministic system. This is done by letting the perturbation amplitudes go to zero. To investigate the dependence of the synchronization frequency on the frequencies of the coupled oscillators, and the different coupling parameters, we assume

FIGURE 2 | Emergence of oscillator death in model (1) for $\lambda = 0.5$, $N = 20$, natural frequencies uniformly distributed in the interval $[0.9, 1.1]$ and varying μ. The interconnection topology is all-to-all. Note that for the same value of $\mu = 0.25$ both oscillatory and oscillator death states are possible.

FIGURE 3 | Nullclines of the cluster state dynamics for incresing values of μ for $x_j = 0$, $\omega = 1.0$, $N = 20$ and $N_1 = N_2 = N/2$, $\lambda = 0.5$.

synchronization, which reduces the FPE equation to an equation in two variables.

We divide this section in two parts. First, we consider a simple deterministic system in which the effect of coupling can be understood. In the second part, we randomly perturb a more general version of the previous model to show that the FP equation provides an approximation for the synchronization frequency, and obtain some insights on the effect of noise.

Let us then consider a system similar to the one already studied

$$\ddot{x}_i = -\omega^2 x_i + \nu \left[1 - \left(x_i^2 + \frac{\dot{x}_i^2}{\omega^2} \right) \right] \dot{x}_i$$

$$+ \mu \sum_{j=1}^{N} a_{ij} x_j, \quad i = 1, \ldots, N, \tag{9}$$

Notice that $x(t) = sin(\omega t)$ is still a solution for the uncoupled system ($\mu = 0$), independently of the value of ν. This system has the advantage of allowing direct calculations around the limit cycle, which can be written explicitly.

If we linearize the equation and take $\nu << 1$, we can neglect the contribution of the disipative term. The resulting linear system is

$$\ddot{x}_i = -(\omega^2 - \mu A_i)x_i,$$

where $A_i = \sum_{j=1}^{N} a_{ij}x_j$, for $i = 1, \ldots, N$. If all the $a_{ij} = 1$, corresponding to a fully connected network of oscillators, the previous system can be reduced to the equation

$$\ddot{x} = -\left(\omega^2 - \mu(N-1)\right)x,$$

assuming that a synchronized regime is established. This assumption may not always be biologically realistic, but it allows us to obtain the common synchronization frequency in a simple way. Later on we consider the general case and recover this formula as a particular case. This provides an estimate for the synchronization frequency of

$$\Omega_{sync} = \sqrt{\omega^2 - \mu(N-1)}.$$

Moreover, this reduction suggests that synchronized oscillatory behavior takes place for sufficiently small μ. That is, when

$$\omega^2 - \mu(N-1) > 0,$$

Otherwise, exponentially large growth can be expected. Notice that unless the a_{ij} are equal, the previous reasoning is not consistent and no conclusion can be drawn. We claim that introducing random perturbations and using the FP equation allows us to circumvent this difficulty and analyze the general case. This is the content of what follows. First of all, we write the system in the form

$$\dot{x}_i = y_i$$

$$\dot{y}_i = f_i(x_i, y_i) + \mu \sum_{j=1}^{N} a_{ij}x_j, \quad i = 1, \ldots, N.$$

Notice that in the linearized regime, anologous to the reasoning for small ν in the previous example, we might naturally assume that

$$f_i(x_i, y_i) \approx -\omega_i^2.$$

Perturbing the equation with Brownian noise we have

$$dx_i = y_i \, dt + \sqrt{2\varepsilon} \, dW_{i1}$$

$$dy_i = \left[-\omega_i^2 x_i + \mu \sum_{j=1}^{N} a_{ij}x_j \right] dt + \sqrt{2\varepsilon} \, dW_{i2}, \quad i = 1, \ldots, N,$$

where the W_{ij} are uncorrelated Brownian motions for $i, j \in \{1, \ldots, N\}$. The probability density, $u(x_1, \ldots, x_n, y_1, \ldots, y_n, t)$ of the

system being in the state $x_1, \ldots, x_n, y_1, \ldots, y_n$ at time t satisfies the FP Equation

$$\frac{\partial u}{\partial t} = \varepsilon \Delta u + \nabla(F(x, y)u), \tag{10}$$

where F is the vector field determined by the right hand side of the stochastic system. Looking for stationary solutions, i.e., $u_t = 0$ and explicitly substituting F in terms of x and y, the equation becomes

$$\varepsilon \Delta u + \sum_i \left(y_i u_{x_i} + (-\omega_i^2 x_i + \mu \sum_{j \neq i} a_{ij}x_j)u_{y_i} \right) = 0. \tag{11}$$

If we use the synchronization condition $x_1 = \ldots = x_n$ and $y_1 = \ldots = y_n$, we obtain the equation

$$\varepsilon \Delta + n y u_x + (-\sum_i \omega_i^2 + \mu \sum_{i,j} a_{ij})x u_y = 0.$$

If we let

$$b = \sum_i \omega_i^2 - \mu \sum_{i,j} a_{ij},$$

we can write the Equation (11) as

$$\varepsilon \Delta + n y u_x - b x u_y = 0.$$

Assuming ε is small, it is reasonable to expect that the probability u will concentrate around the characteristic curves of the first order equation

$$n y u_x - b x u_y = 0,$$

that are solutions to the system

$$\dot{x} = ny,$$
$$\dot{y}_i = bx.$$

By taking the scalar product with the vector $(x/n, y/b)$ we obtain the relation

$$\dot{x}\frac{x}{n} + \dot{y}\frac{y}{b} = 0,$$

or equivalently,

$$\frac{x^2}{n} + \frac{y^2}{b} = constant,$$

which defines the characteristic curves as ellipses. In turn, interpreting these as curves in the phase portrait, the resulting solutions would correspond to periodic trajectories with frequency

$$\Omega_{sync}^2 = -\frac{b}{n} = \frac{(\sum_i \omega_i^2 - \mu \sum_{i,j} a_{ij})}{n}, \tag{12}$$

which provides an estimate for the synchronization frequency in terms of the original frequencies and the coupling

FIGURE 4 | Synchronization and modulation of the synchronization frequency via coupling strength in model (9), for $N = 10$, natural frequencies uniformly distributed in the interval [0.95, 1.05], all-to-all coupling and varying μ.
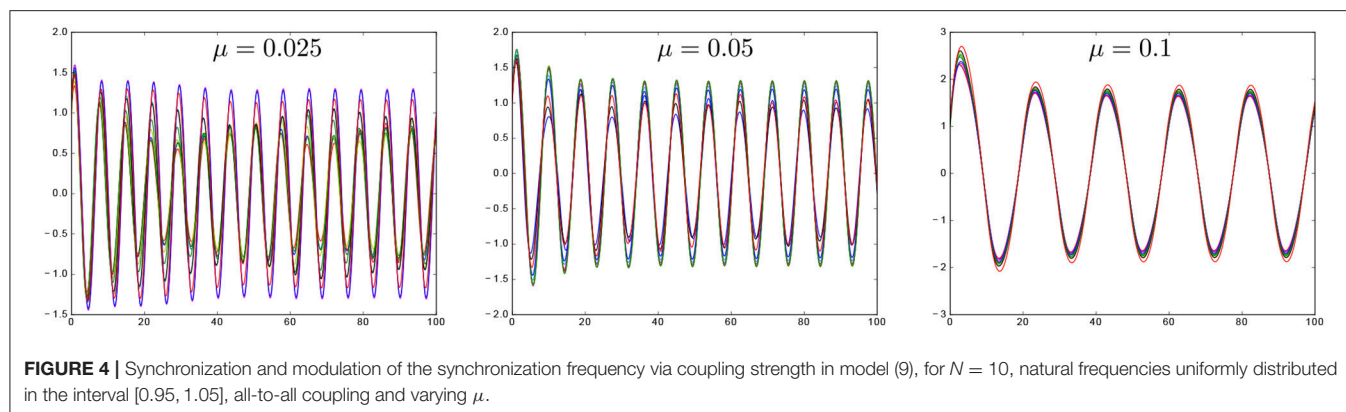
parameters. Importantly, it shows that the synchronization frequency decreases with the coupling strength. In particular, formula Equation (12) can be used to study how coupled ultradian oscillations can give rise to circadian oscillations (**Figure 4**).

4. DISCUSSION AND SUMMARY

We have described, through basic geometrical analysis, the relationship between the dissipation coefficient, an intrinsic property of the oscillators we study, and the coupling strength μ in a network of diffusively coupled non-linear oscillators. Our analysis predicts the emergence of sustained oscillations for increasing values of the parameter λ in the system, only for a limited range of coupling strengths. It is reasonable to conjecture from this result, that there is a functional limit in the coupling strength for oscillating tissues in nature above which the tissue oscillations dies. To the best of our knowledge, it is the first time that diffusive coupling has been shown to be able to induce such oscillator death.

We have also derived an estimation for the synchronization frequency of a linearly coupled network of non-linear oscillators in terms of the oscillator natural frequencies and the coupling parameters [Equation (12)]. The presented results are indeed local, that is, the synchronous oscillation is only locally asymptotically stable. The oscillators are not synchronized at the beginning of the simulations, but their spread in state space is very small to ensure convergence to the synchronous oscillation. For a larger spread, a more complex behavior is observed.

We believe that these results constitute predictions that, although possibly difficult to test experimentally, would be worth verifying in light of the existing evidence about the joint frequency modulation of activity between different tissues during the day [23, 45].

The results we have presented thus far emphasize the importance of simple mathematical models in understanding situations where synchronization of multiple oscillating populations appears. The results presented here may help to shed light on both physiological and pathological phenomena involving synchronization of oscillators in different tissues (Parkinson's disease [46, 47], epilepsy [48, 49]). The other way around, it is also of potential importance to unravel mechanisms underlying the disappearance of coordinated oscillatory regimes. In a future publication, we plan to formally justify our estimations, and further, integrate the analysis of oscillations in the cellular and network levels of biological organization, to build up on our understanding of coupling oscillators at the tissue level. Two important extensions of the current models that we are studying are, the full characterization in higher codimension of the bifurcation structures of the system (1), and also, replacements of the van der Pol dynamics with biophysical models of excitable cells [50]. This last extension may prove useful to explain possible compensatory mechanisms that take place during the beginning of a pathology [51].

AUTHOR CONTRIBUTIONS

All authors wrote the paper. ML-A proposed the original set of equations, AF and PP-L proposed revised sets of equations and extended models. All authors performed the analysis. AF and MH-V performed simulations.

REFERENCES

1. Bünsow RC. The circadian rhythm of photoperiodic responsiveness in Kalanchoe. In: Stillman B, Stewart D, Grodzicker T, editors. *Cold Spring Harbor Symposia on Quantitative Biology*. Vol. 25. Cold Spring Harbor Laboratory Press (1960). p. 257–60.

2. Pol B. Biological rhythms considered as relaxation oscillations. *J Intern Med*. (1940) **103**:76–88.

3. Sholl D. Regularities in growth curves, including rhythms and allometry. *Dyn Growth Process*. (1954) 224–241.

4. Pavlidis T, Zimmerman W, Osborn J. A mathematical model for the temperature effects on circadian rhythms. *J Theor Biol*. (1968) **18**:210–21.

5. Klotter K. Theoretical analysis of some biological models. In: Stillman B, Stewart D, Grodzicker T, editors. *Cold Spring Harbor Symposia on Quantitative Biology*. Vol. 25. Cold Spring Harbor Laboratory Press (1960). p. 189–96.

6. Klotter K. General properties of oscillating systems. In: Stillman B, Stewart D, Grodzicker T, editors. *Cold Spring Harbor Symposia on Quantitative Biology*. Vol. 25. Cold Spring Harbor Laboratory Press (1960). p. 185–87.

7. Sweeney BM. The photosynthetic rhythm in single cells of Gonyaulax polyedra. In: Stillman B, Stewart D, Grodzicker T, editors. *Cold Spring Harbor Symposia on Quantitative Biology*. Vol. 25. Cold Spring Harbor Laboratory Press (1960). p. 145–48.

8. Winfree AT. *The Geometry of Biological Time*. Vol. 12. New York, NY: Springer Science & Business Media (2001).

9. Hirota T, Lee JW, Lewis WG, Zhang EE, Breton G, Liu X, et al. High-throughput chemical screen identifies a novel potent modulator of cellular circadian rhythms and reveals CKIα as a clock regulatory kinase. *PLoS Biol*. (2010) **8**:e1000559. doi: 10.1371/journal.pbio.1000559

10. Partch CL, Green CB, Takahashi JS. Molecular architecture of the mammalian circadian clock. *Trends Cell Biol*. (2014) **24**:90–9. doi: 10.1016/j.tcb.2013.07.002.

11. Welsh DK, Yoo SH, Liu AC, Takahashi JS, Kay SA. Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression. *Curr Biol*. (2004) **14**:2289–95. doi: 10.1016/j.cub.2004.11.057

12. Hastings MH, Maywood ES, O'Neill JS. Cellular circadian pacemaking and the role of cytosolic rhythms. *Curr Biol*. (2008) **18**:R805–15. doi: 10.1016/j.cub.2008.07.021

13. Nader N, Chrousos GP, Kino T. Interactions of the circadian CLOCK system and the HPA axis. *Trends Endocrinol Metab*. (2010) **21**:277–86. doi: 10.1016/j.tem.2009.12.011

14. Abruzzi KC, Rodriguez J, Menet JS, Desrochers J, Zadina A, Luo W, et al. Drosophila CLOCK target gene characterization: implications for circadian tissue-specific gene expression. *Genes Dev*. (2011) **25**:2374–86. doi: 10.1101/gad.174110.111

15. Manoogian EN, Panda S. Circadian rhythms, time-restricted feeding, and healthy aging. *Ageing Res Rev*. (2017) **39**:59–67. doi: 10.1016/j.arr.2016.12.006.

16. Tahara Y, Aoyama S, Shibata S. The mammalian circadian clock and its entrainment by stress and exercise. *J Physiol Sci*. (2017) **67**:1–10. doi: 10.1007/s12576-016-0450-7

17. Rusak B, Zucker I. Neural regulation of circadian rhythms. *Physiol Rev*. (1979) **59**:449–526.

18. Guo H, Brewer JM, Lehman MN, Bittman EL. Suprachiasmatic regulation of circadian rhythms of gene expression in hamster peripheral organs: effects of transplanting the pacemaker. *J Neurosci*. (2006) **26**:6406–12. doi: 10.1523/JNEUROSCI.4676-05.2006

19. Tahara Y, Shibata S. Circadian rhythms of liver physiology and disease: experimental and clinical evidence. *Nat Rev Gastroenterol Hepatol*. (2016) **13**:217. doi: 10.1038/nrgastro.2016.8

20. Oster H, Damerow S, Kiessling S, Jakubcakova V, Abraham D, Tian J, et al. The circadian rhythm of glucocorticoids is regulated by a gating mechanism residing in the adrenal cortical clock. *Cell Metab*. (2006) **4**:163–73. doi: 10.1016/j.cmet.2006.07.002

21. Ishida A, Mutoh T, Ueyama T, Bando H, Masubuchi S, Nakahara D, et al. Light activates the adrenal gland: timing of gene expression and glucocorticoid release. *Cell Metab*. (2005) **2**:297–307. doi: 10.1016/j.cmet.2005.09.009

22. Kiessling S, Eichele G, Oster H. Adrenal glucocorticoids have a key role in circadian resynchronization in a mouse model of jet lag. *J clin Invest*. (2010) **120**:2600–9. doi: 10.1172/JCI41192

23. Damiola F, Le Minh N, Preitner N, Kornmann B, Fleury-Olela F, Schibler U. Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes Dev*. (2000) **14**:2950–61. doi: 10.1101/gad.183500

24. Kronauer RE, Czeisler CA, Pilato SF, Moore-Ede MC, Weitzman ED. Mathematical model of the human circadian system with two interacting oscillators. *Am J Physiol Regul Integr Comp Physiol*. (1982) **242**:R3–17.

25. olde Scheper T, Klinkenberg D, Pennartz C, Van Pelt J. A mathematical model for the intracellular circadian rhythm generator. *J Neurosci*. (1999) **19**:40–47.

26. Glass L. Synchronization and rhythmic processes in physiology. *Nature* (2001) **410**:277. doi: 10.1038/35065745

27. Ueda HR, Hagiwara M, Kitano H. Robust oscillations within the interlocked feedback model of Drosophila circadian rhythm. *J Theor Biol*. (2001) **210**:401–6. doi: 10.1006/jtbi.2000.2226

28. Jiang J, Liu Q, Niu L. Theoretical investigation on models of circadian rhythms based on dimerization and proteolysis of PER and TIM. *Math Biosci Eng*. (20170 **14**:1247–59. doi: 10.3934/mbe.2017064

29. Lara-Aparicio M, de Medrano SL, Fuentes-Pardo B, Moreno-Sáenz E. A qualitative mathematical model of the ontogeny of a circadian rhythm in crayfish. *Bull Math Biol*. (1993) **55**:97–110.

30. Fanjul-Moles ML, Moreno-Sáenz E, Villalobos-Hiriart N, Fuentes-Pardo B. ERG circadian rhythm in the course of ontogeny in crayfish. *Comp Biochem Physiol A Physiol*. (1987) **88**:213–9.

31. Fanjul-Moles ML, Miranda-Anaya M, Fuentes-Pardo B. Effect of monochromatic light upon the erg crcadian rhythm during ontogeny in crayfish (Procambarus clarkii). *Comp Biochem Physiol A Physiol*. (1992) **102**:99–106.

32. Fuentes-Pardo B, Fanjul-Moles ML, Moreno-Sáenz E. Synchronization by light of the ERG circadian rhythm during ontogeny in the crayfish. *Biol Rhythm Res*. (1992) **23**:81–91.

33. Van der Pol B. On "relaxation-oscillations." *Lond Edinb Dublin Philos Mag J Sci*. (1926) **2**:978–92.

34. Fuentes-Pardo B, Lara-Aparicio M, de Medrano SL. Perturbation of a circadian rhythm by single and periodic signals and its mathematical simulation. *Bull Math Biol*. (1995) **57**:175–89.

35. Dowse HB, Hall JC, Ringo JM. Circadian and ultradian rhythms in *period* mutants of *Drosophila melanogaster*. *Behav Genet*. (1987) **17**:19–35.

36. Goldbeter A. From Ultradian Biochemical Oscillations to Circadian Rhythms. In: Vanden Driessche T, Guisset JL, Petiau-de Vries GM, editors. *Membranes and Circadian Rythms*. Springer (1996). p. 67–93.

37. Lloyd D, Rossi EL. *Ultradian Rhythms in Life Processes: An Inquiry Into Fundamental Principles of Chronobiology and Psychobiology*. London: Springer Science & Business Media (2012).

38. Barriga-Montoya C, Padilla-Longoria P, Lara-Aparicio M, Fuentes-Pardo B. Ultradian rhythms underlying the dynamics of the circadian pacemaker. In: Vonend O, editior. *Aspects of Pacemakers-Functions and Interactions in Cardiac and Non-Cardiac Indications*. InTech (2011).

39. Agaev R, Chebotarev P. On the spectra of nonsymmetric Laplacian matrices. *Linear Algebra Appl*. (2005) **399**:157–68. doi: 10.1016/j.laa.2004.09.003

40. La Salle JP. An invariance principle in the theory of stability, differential equations and dynamical systems. In: *Proceedings of the International Symposium, Puerto Rico* (1967). p. 277–86.

41. Guckenheimer J, Holmes P. Local bifurcations. In: Guckenheimer J, Holmes P, editors. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York, NY: Springer (1983), p. 117–165. doi: 10.1007/978-1-4612-1140-2_3

42. Hassard BD, Kazarinoff ND, Wan YH. *Theory and Applications of Hopf Bifurcation*. Vol. 41. Cambridge: CUP Archive (1981).

43. Golubitsky M, Stewart I, Schaeffer DG. *Singularities and Groups in Bifurcation Theory*. Vol. 1. Verlag: Springer (1985).

44. Jiao X, Wang R. Synchronous firing patterns of neuronal population with excitatory and inhibitory connections. *Int J Non Linear Mech*. (2010) **45**:647–51. doi: 10.1016/j.ijnonlinmec.2008.11.020

45. Le Minh N, Damiola F, Tronche F, Schütz G, Schibler U. Glucocorticoid hormones inhibit food-induced phase-shifting of peripheral circadian oscillators. *EMBO J*. (2001) **20**:7128–36. doi: 10.1093/emboj/20.24.7128

46. Hammond C, Bergman H, Brown P. Pathological synchronization in Parkinson's disease: networks, models and treatments. *Trends Neurosci.* (2007) **30**:357–64. doi: 10.1016/j.tins.2007.05.004

47. Schnitzler A, Gross J. Normal and pathological oscillatory communication in the brain. *Nat Revi Neurosci.* (2005) **6**:285–96. doi: 10.1038/nrn1650

48. Engel Jr J, Bragin A, Staba R, Mody I. High-frequency oscillations: what is normal and what is not? *Epilepsia* (2009) **50**:598–604. doi: 10.1111/j.1528-1167.2008.01917.x

49. Velazquez JLP, Carlen PL. Gap junctions, synchrony and seizures. *Trends Neurosci.* (2000) **23**:68–74. doi: 10.1016/S0166-2236(99)01497-6

50. Herrera-Valdez MA, McKiernan EC, Berger SD, Ryglewski S, Duch C, Crook S. Relating ion channel expression, bifurcation structure, and diverse firing patterns in a model of an identified motor neuron. *J comput Neurosci.* (2013) **34**:211–29. doi: 10.1007/s10827-012-0416-6

51. McKiernan EC. A genetic manipulation of motor neuron excitability does not alter locomotor output in *Drosophila larvae*. *PeerJ PrePrints* (2015) **3**:e469v3. doi: 10.7287/peerj.preprints.469v3

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

A Stochastic Phylogenetic Algorithm for Mitochondrial DNA Analysis

M. Corona-Ruiz[1], Francisco Hernandez-Cabrera[1]*, José Roberto Cantú-González[2], O. González-Amezcua[1] and Francisco Javier Almaguer[1]*

[1] *Facultad de Ciencias Físico-Matemáticas, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Mexico,*
[2] *Escuela de Sistemas PMRV, Unidad Acuña, Universidad Autónoma de Coahuila, Saltillo, Mexico*

This paper presents an exploratory analysis of the mitochondrial DNA (mtDNA) of 32 species in the subphylum Vertebrata, divided in 7 taxonomic classes. Multiple stochastic parameters, such as the Hurst and detrended fluctuation analysis (DFA) exponents, Shannon entropy, and Chargaff ratio are computed for each DNA sequence. The biological interpretation of these parameters leads to defining a triplet of novel indices. These new functions incorporate the long-range correlations, the probability of occurrence of nucleic bases, and the ratio of pyrimidines-to-purines. Results suggest that relevant regions in mtDNA can be located using the proposed indices. Furthermore, early results from clustering algorithms indicate that the indices introduced might be useful in phylogenetic studies.

Keywords: DNA, random-walk, Hurst exponent, detrended fluctuation analysis, Shannon entropy, coefficient of disequilibrium

1. INTRODUCTION

Previous mathematical studies on DNA sequences have seen a variety of approaches and frequently involve a numerical representation of the nucleotide chains. For instance, distance matrices have been constructed using different metrics (Randi et al., 2003; Liao and Wang, 2004; Zhang and Tan, 2007; Kandiah and Shepelyansky, 2013). These matrices, in combination with clustering methods, are used to evaluate phylogenetic relationships among species (Yu and Huang, 2013).

Other studies involve the representation of DNA sequences as random-walks, known as *DNA-walks* (Peng et al., 1994). The main objectives of these studies focus on the long-range correlations among nucleotides; i.e., "how the frequency of each nucleotide of a pairing nucleotide couple changes locally" (Namazi and Kiminezhadmalaie, 2015). These DNA-walk studies find differences in the long-range correlation between coding and non-coding DNA sequences (Peng et al., 1994).

Recently, DNA-walk analysis has been used in combination with the fractal dimension and Hurst exponent to identify mosaic structures in DNA that allow distinguishing between healthy and cancerous cells (Namazi and Kiminezhadmalaie, 2015).

Additionally, alternative statistical tools frequently used in DNA sequence analysis include Shannon entropy, which is a measure of the amount of "information" stored within a system (López-Ruiz et al., 1995). In a biological sense, Shannon entropy evaluates the probability of independent occurrences of each nucleic base in a DNA sequence. In recent studies, fluctuations in local Shannon entropy in DNA sequences have been analyzed to identify regions of repeating patterns of one or more nucleotides, known as *tandem repeats* (Thanos et al., 2018). The capability of Shannon entropy to highlight important segments in DNA sequences has led to the supported notion that entropy studies might be used for biological classifications of species (Melnik and Usatenko, 2014).

Similarly, the concept of complexity has played a central role in various DNA sequence analyses. For instance, López-Mancini-Calbet (LMC) complexity, employed in this paper, has led to the development of an effective gene-predicting technique (López-Ruiz et al., 1995; Monge and Crespo, 2015). In a recent study, the symbolic complexity of DNA sequences is used to identify segments resulting from random duplication, as well as changes in the speed of accumulation of point mutations (Salgado-Garcia and Ugalde, 2016).

Our objective is to examine the parameters previously mentioned to determine a small number of coefficients with biological relevance that may be used to determine rates of change in nucleotide bases, establish comparisons between regions, and better understand the relation among species in a phylogenetic sense.

This paper is structured as follows: section 2 introduces the concepts and methodology; section 3 presents the results obtained and the variables introduced; and section 4 is devoted to a discussion of the results, comments on the methodology in general, and final remarks. Tables and figures are incorporated in sections 2 and 3, respectively. The **Supplementary Material** includes a table with the identification codes for the data.

2. METHODOLOGY

GenBank® is the *National Institutes of Health*'s genetic sequence database made possible by the collaboration of several organizations. All datasets used within this work were obtained through GenBank because of its availability of access, encouragement of use, and the advantage that the information stays up-to-date.

A total of 32 complete mtDNA sequences of different species in the subphylum Vertebrata were selected. The lengths vary from $16,207$ to $18,254$ base pairs (bp). The choice of this type of DNA presents multiple advantages: it is relatively small in size (in contrast, human chromosomal DNA contains hundreds of millions bp); the sequences contain conserved regions, can be compared in blocks among different species, and contain a small percentage of non-coding regions; and the interpretation of the mutations in mtDNA as an estimator of evolutionary change (Barton and Jones, 1983). For these reasons, the exploratory nature of this study does not require additional information on the species themselves. Thus, the selection criteria focused on 32 different members from 7 groups intuitively related in taxonomic classes. The 32 NCBI codes from the data files have been attached in the **Table S1**.

A pre-processing of the data files consists of a realignment of the sequences to set the control region of the heavy chain (H-chain) in the direction of transcription as the new ending point. This realignment is done once. The *displacement loop*, or D-loop, is within the control region and the most varying region in mtDNA, with substantial differences observed even among individuals of the same species (Yamamoto, 2001). See **Figure S1** (Supplementary Material). Additionally, the header information was removed, which contains the identification key and the name of the organism. The downloaded files (in *.fasta* format)

were processed using the programming language R version 3.4.4 (2018-03-15). The packages used are *stringr* and *fractal*.

2.1. DNA-Walk

DNA consists of sequences of nitrogenous bases: adenine (A), guanine (G), thymine (T), and cytosine (C). The length and distribution of the bases fluctuate from species to species. Several mappings have been introduced based on properties intrinsic to DNA. Moreover, adenine and guanine have a two-ring structure and belong to the *purine* group, while cytosine and thymine have a one-ring structure and belong to the *pyrimidine* group. Furthermore, adenine bonds with thymine through a double hydrogen bond, which is called a *weak bond*, while guanine and cytosine bond through a triple hydrogen bond, which is called a *strong bond*. **Figure 1** illustrates these descriptions. In summary, we have:

- *Purine (R)*: {A, G} / *Pyrimidine (Y)*: {C, T}
- *Strong Hydrogen bond (S)*: {G, C} / *Weak Hydrogen bond (W)*: {A, T}
- *Keto (K)*: {G, T} / *Amino (M)*: {A, C}

Considering the properties described previously, it is possible to read a DNA sequence and assign either a $+1$ or -1 depending on whether the respective nucleotide is a purine or pyrimidine (*RY* rule). This can be interpreted as random steps x_i of a one-dimensional walk. Then, the final position after n steps is given by

$$X_n = x_0 + \sum_{i=1}^{n} x_i \tag{1}$$

where $x_0 = 0$ by definition.

Let $S = \{s_1 s_2 \ldots s_M\}$ be a nucleotide sequence of length M, where $s_k \in \{A, C, G, T\}$ for $k \in \{1, 2, \ldots, M\}$. Hence, a one-dimensional DNA-walk can be defined through the following rules:

- *RY rule:*

$$x_k = \begin{cases} 1 & \text{if } s_k \in R = \{A, G\} \\ -1 & \text{if } s_k \in Y = \{C, T\} \end{cases} \tag{2}$$

- *SW rule:*

$$x_k = \begin{cases} 1 & \text{if } s_k \in S = \{C, G\} \\ -1 & \text{if } s_k \in W = \{A, T\} \end{cases} \tag{3}$$

- *KM rule:*

$$x_k = \begin{cases} 1 & \text{if } s_k \in M = \{A, C\} \\ -1 & \text{if } s_k \in K = \{G, T\} \end{cases} \tag{4}$$

where s_k is the $k-$th nucleotide and x_k is the value of the $k-$th assigned step in a DNA sequence. The path of the DNA-walk after n steps is then defined as the partial sums $X_n = x_0 + \sum_{k=1}^{n} x_k$, where $n \in \{1, 2, \ldots, M\}$ and $x_0 = 0$.

In the context of DNA-walks, Equation (2) evaluates the tendency of changes between purines and pyrimidines. Transversions (substitutions of purines for pyrimidines, or vice versa) are less likely to happen and have been used to evaluate

FIGURE 1 | Chemical structure of DNA. Adenine, guanine, cytosine, and thymine are shown in colors green, blue, red, and purple, respectively. Notice the double-ring structure of the purines (A, G) and the single-ring structure of the pyrimidines (C, T). Similarly, the type of bond is readily observable: double- and triple-Hydrogen bonds for A, T and G, C, respectively. This illustration, by Madeleine Price Ball, has a Creative Commons Zero (CC0, i.e., "No Rights Reserved") license and has been published in previous articles (Wikimedia Commons Contributors, 2018).

molecular evolution (Stoltzfus and Norris, 2016). Thus, using this rule within corresponding blocks of nucleotides in different species, it is possible to observe changes in the DNA-walk that could be interpreted as an evolutionary variation. Similarly, Equation (4) is associated with the rate of recombination between transversions and transitions (purine-purine or pyrimidine-pyrimidine substitutions).

Moreover, Equation (3) refers to the difference in abundance of the GC bond with respect to the AT bond. A higher GC content suggests a significantly higher temperature for DNA denaturing (melting temperature T_m). Previous studies have shown that GC content is associated to an age-related natural selection and environmental factors (Min and Hickey, 2008). Finally, it is assumed that each DNA-walk is an ergodic stochastic process. Specifically, the conceived notion adopted is that each DNA sequence may be used to represent the ensemble of DNA sequences of individuals within the same species.

In summary, the three assignment rules provide insight into the evolutionary aspects of the organisms considered.

2.2. Hurst Exponent and DFA Exponent

Additional information of the long-range correlations of DNA-walks can be obtained via stochastic methods such as *rescaled-range analysis* and *detrended fluctuation analysis*. With these methods, it is possible to obtain the Hurst exponent,

which represents a quantitative measure of the fractal nature of DNA sequences.

The Hurst exponent, here denoted by α, satisfies $0 < \alpha < 1$. In comparisons of mtDNA sequences, each Hurst exponent can be interpreted as a measure of the tendency of changes between nucleotides according to the rules mentioned in the previous section. The calculations used to obtain the Hurst exponent have been reported in previous studies (Peng et al., 1994; Buldyrev et al., 1995).

The Hurst exponent is directly related to the fractal dimension α' by the relation:

$$\alpha' = 2 - \alpha. \tag{5}$$

The fractal dimension evaluates changes in detail of the pattern of a DNA-walk with respect to the scale used for measurement.

An alternative method to calculate the Hurst exponent of a DNA-walk is DFA. In contrast to the rescaled-range analysis, DFA analyzes the random fluctuations of the DNA-walk without trend in the data (Peng et al., 1994; Buldyrev et al., 1995). The DFA exponent is computed using the following algorithm:

- Given a numerical sequence $X = \{X_1, X_2, \ldots, X_M\}$, calculate the cumulative sum

$$y_k = \sum_{i=1}^{k} (X_i - \overline{X}) \tag{6}$$

where $k = 1, 2, \ldots, M$ and $\overline{X}$ is the mean value of X.
- Divide y_k into M/L subintervals of length L. For each window, calculate the polynomial linear fit (the local trend) $y_{k,L}$ via least-squares minimization.
- Calculate the fluctuation, which is an average of the squares of the detrended sequence given by

$$F^2(L) = \frac{1}{M} \sum_{k=1}^{M} \left| y_k - y_{k,L} \right|^2. \tag{7}$$

- The slope β of the linear regression analysis in the scale $\log F(L) / \log L$ is an estimator of the Hurst exponent.

This method tests for self-similarity at different window sizes L. No correlation (or short-range correlations) gives stochastic properties such as those of a random-walk, so $\beta = 0.5$; in contrast, long-range correlations give a value of $\beta \neq 0.5$. Specifically, correlation yields $\beta > 0.5$, while anti-correlation gives $\beta < 0.5$.

This paper adopts a minimum block size of 4 nucleotides, while the maximum is $B = \frac{M}{2}$, corresponding to half the length of the sequence in question. Should M be odd, B is rounded down.

2.3. Chargaff Ratio

In a remarkable discovery, Erwin Chargaff determined that there is a balance held in DNA by the nucleobases (Chargaff, 1950), known as *Chargaff's Rule*. These state: (1) that globally (i.e., considering both strands of DNA) adenine is equal to thymine in quantity, and (2) that guanine is equal to cytosine in quantity. This result was the basis for the Watson-Crick model, which

determined that adenine binds with thymine and that guanine binds with cytosine (Watson and Crick, 1953).

On this basis, and in the context of this work, the *Chargaff ratio* is defined as the ratio of pyrimidines to purines:

$$\xi = \frac{N_C + N_T}{N_A + N_G} \qquad (8)$$

where N_C, N_T, N_A, N_G represent the amount of cytosine, thymine, adenine, and guanine, respectively, within one strand of DNA. Note that this value is always positive. If $0 \leq \xi < 1$, there are more purines than pyrimidines (i.e., $N_C + N_T < N_A + N_G$); similarly, $\xi > 1$ reflects an excess of pyrimidines over purines. A Chargaff ratio with value 1 results from an equal number of either type of nucleotide bases.

2.4. Shannon Entropy

In his seminal paper, Claude Shannon introduced the concept of *information entropy*. It measures the "amount" of information or uncertainty of a system (Shannon and Weaver, 1998). Let $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ be a set of events where each ω_i has probability of occurrence $p_i \in [0, 1]$, for $i = 1, 2, \ldots, N$. Thus, the Shannon entropy of the system is defined as

$$\mathcal{H} = -K \sum_{i=1}^{N} p_i \log_2(p_i), \qquad (9)$$

where K is a positive constant chosen appropriately according to the units desired for measurement (thus, for this work, $K = 1$). For the case when $p_i = 0$, $p_i \log_2(p_i) = 0$ in the limit definition. Also, note that the logarithm is in base 2; this is because information in a computer is encoded in *binary digits*, or *bits*, which are the basic units of measurement of information.

For $N = 2$, events ω_1 and ω_2 have probability p and $1 - p$, respectively, see **Figure S2** (Supplementary Material). Thus, it can be seen that a maximum is attained at $p = 1 - p = \frac{1}{2}$. This result can be extended to the general case with N events. The proof requires *Jensen's inequality* for a concave function (in this case, the *logarithmic* function), and is given below. Using some algebra to rewrite Equation (9) with $K = 1$ yields

$$\mathcal{H} = \log_2 \left(\prod_{i=1}^{N} \left(\frac{1}{p_i} \right)^{p_i} \right)$$

By the *weighted arithmetic-mean and geometric-mean* inequality, this implies that

$$2^{\mathcal{H}} = \prod_{i=1}^{N} \left(\frac{1}{p_i} \right)^{p_i} \leq \sum_{i=1}^{N} p_i \left(\frac{1}{p_i} \right) = N$$

where equality (the maximum) is satisfied when $p_1 = p_2 = \cdots = p_N$. That is, when

$$\mathcal{H} = \log_2(N). \qquad (10)$$

To evaluate Shannon entropy in the context of DNA sequence analysis, it seems rather reasonable to define the set of possible

events as $\Omega = \{A, G, C, T\}$. However, it is expected that the probability of occurrence of each nucleotide in a DNA sequence will likely be different for different species; thus, these associated probabilities will be calculated empirically for each DNA sequence in a straightforward fashion. That is, by counting the amount of each nucleotide within the sequence and taking the corresponding proportion by dividing by the total amount of nucleotides M. Thus, the probabilities will be given by

$$p_A = \frac{N_A}{M}, \qquad p_C = \frac{N_C}{M}, \qquad p_G = \frac{N_G}{M}, \qquad p_T = \frac{N_T}{M}, \qquad (11)$$

where N_A, N_C, N_G, N_T are the amount of *adenine*, *cytosine*, *guanine*, and *thymine*, respectively.

In the context of DNA sequence analysis, maximum entropy is attained whenever the nucleic bases within a DNA sequence are found with equiprobability. It may thus be interpreted that such a sequence is the result of a random combination of these events. Any departure from the maximum value of the Shannon entropy due to an underlying structure might contribute to determining any tendencies present in a sequence, see **Figure S3** (Supplementary Material).

In a more general sense, the entropy fluctuations could be analyzed by means of the Local Shannon entropy. By studying the local fluctuations of entropy at a given scale, and across scales, an "entropic microscope" could highlight areas with a high degree of variation or, equally interesting, low degree of variation, as seen in previous studies (Melnik and Usatenko, 2014; Thanos et al., 2018).

2.5. Coefficient of Disequilibrium

Additional information of DNA sequences can be derived from the deviations from equiprobability of occurrence of each
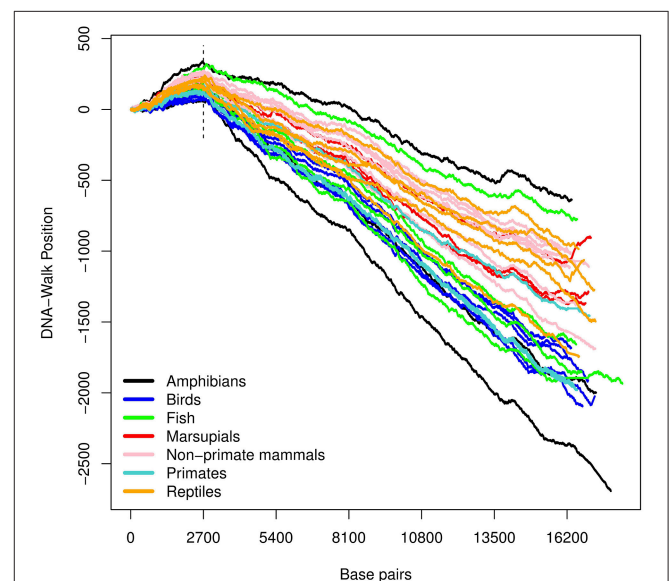
FIGURE 2 | DNA-walk illustration for various species using the *purine-pyrimidine* rule. Observe the vicinity of nucleotide 2, 700 and the change in tendency from a purine-rich region (positive slope) to a predominance of pyrimidines for the remaining DNA-walk (negative slope).

nucleotide. This measure is known as *disequilibrium* (López-Ruiz et al., 1995). The events in the set Ω have probability p_i for $i = 1, 2, 3, 4$. The coefficient of disequilibrium, $\mathcal{D}$, is defined as:

$$\mathcal{D} = \sum_{i=1}^{N=4} \left(p_i - \frac{1}{4} \right)^2. \tag{12}$$
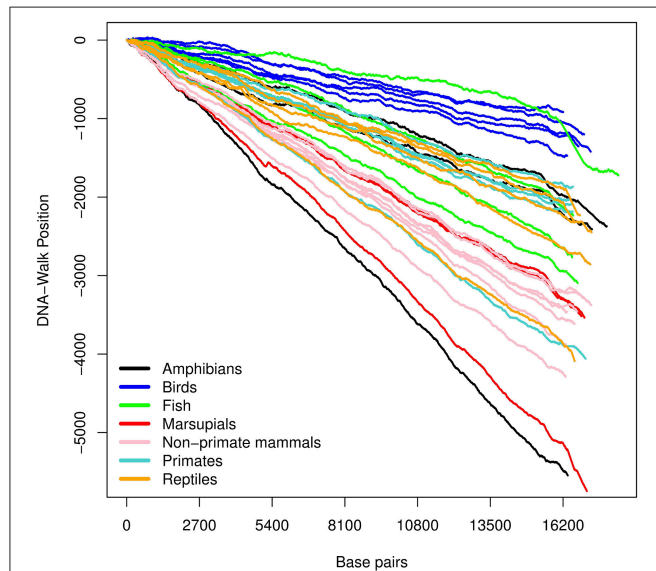
FIGURE 3 | DNA-walk illustration for various species using the *strong- and weak-bond* rule. Observe the immediate (and consistent) tendency. This indicates that mtDNA is rich in adenine and thymine, whose type of bond is weaker than that of cytosine and guanine.
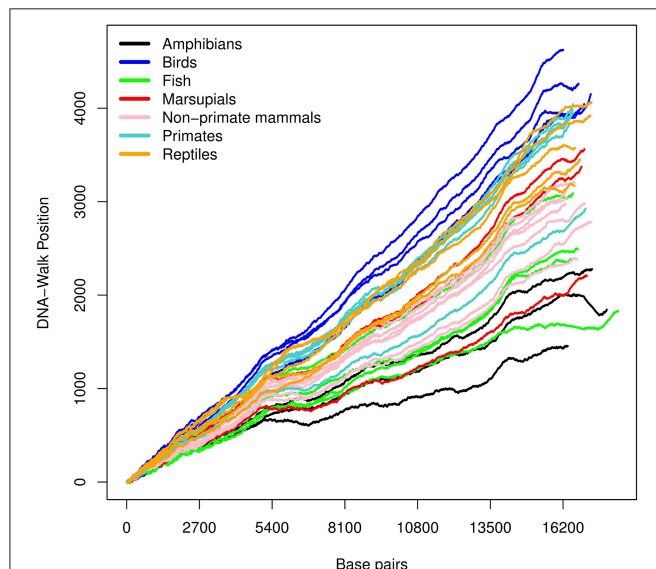
FIGURE 4 | DNA-walk illustration for various species using the *keto and amino* rule. The figure shows a higher amount of adenine and cytosine.

This sum of squared distances can be seen as a type of variance. Note that $\mathcal{D} = 0$ in the case of equilibrium. Any deviation from this would result in $\mathcal{D} > 0$. The maximum disequilibrium value, $\mathcal{D}_{\max} = \frac{3}{4}$ can be obtained using multivariate calculus.

The coefficient of disequilibrium may represent a measure of relatedness between a DNA sequence and one resulting from a random process if each (independent) event has a probability p_i of occurrence. That is, larger deviations from an equiprobable space yield higher coefficients of disequilibrium. It can be observed that this behavior counters that of the Shannon entropy in an intuitive manner.

2.6. Coefficient of Complexity
The coefficient of complexity $\mathcal{C}$ is then given by the product of the Shannon entropy (9) and the coefficient of disequilibrium (12), as in (13). It can be seen from (12) that $\mathcal{D}$ resembles the definition of

TABLE 1 | Results of the Chargaff ratio and Shannon entropy for all groups.

Scientific name (common name)	ξ	$\mathcal{H}$
Ambystoma tigrinum tigrinum (Eastern tiger salamander)	1.081	1.9059
Bufo gargarizans (Chusan Island toad)	1.2617	1.9598
Rana plancyi (Eastern golden frog)	1.3562	1.9591
Ara ararauna (Blue-and-yellow macaw)	1.2537	1.9421
Archilochus colubris (Ruby-throated hummingbird)	1.2296	1.9409
Columba livia (Rock pigeon)	1.2664	1.9381
Gallus gallus (Red junglefowl)	1.2851	1.9316
Ninox strenua (Powerful owl)	1.2421	1.926
Carcharodon carcharias (Great white shark)	1.249	1.9444
Cyprinus carpio (Common carp)	1.0981	1.9577
Dicentrarchus labrax (European seabass)	1.2372	1.9765
Poecilia reticulata (Guppy)	1.2228	1.9529
Didelphis virginiana (Virginia Opossum)	1.1117	1.8969
Macropus giganteus (Eastern gray kangaroo)	1.1762	1.9275
Vombatus ursinus (Common wombat)	1.164	1.9254
Bos taurus (Cattle)	1.1332	1.9339
Canis lupus familiaris (Dog)	1.1848	1.9441
Capra aegagrus (Wild goat)	1.1441	1.9292
Felis catus (Domestic cat)	1.1398	1.9429
Mus musculus musculus (House mouse)	1.1316	1.9154
Oryctolagus cuniculus (Common rabbit)	1.2169	1.9403
Rattus rattus (House rat)	1.1465	1.9219
Gorilla gorilla gorilla (Western lowland gorilla)	1.2706	1.9322
Homo sapiens (Human)	1.2716	1.9305
Lemur catta (Ring-tailed lemur)	1.1869	1.9246
Pan paniscus (Bonobo)	1.2711	1.9272
Pan troglodytes (Common chimpanzee)	1.2717	1.9293
Alligator mississippiensis (American alligator)	1.2338	1.9383
Chelydra serpentina (Common snapping turtle)	1.1259	1.9205
Crocodylus niloticus (Nile crocodile)	1.1347	1.9504
Crotalus horridus (Timber rattlesnake)	1.1898	1.9337
Naja naja (Indian cobra)	1.1597	1.9324

TABLE 2 | Results of the Hurst exponent for all groups and each of the three random-walk rules.

Scientific name (common name)	α_{RY}	α_{SW}	α_{KM}
Ambystoma tigrinum tigrinum (Eastern tiger salamander)	0.91798	0.91328	0.90701
Bufo gargarizans (Chusan Island toad)	0.91688	0.91187	0.91191
Rana plancyi (Eastern golden frog)	0.91695	0.91259	0.91228
Ara ararauna (Blue-and-yellow macaw)	0.91657	0.91337	0.9133
Archilochus colubris (Ruby-throated hummingbird)	0.91621	0.91298	0.91332
Columba livia (Rock pigeon)	0.91696	0.91336	0.91383
Gallus gallus (Red junglefowl)	0.91564	0.91109	0.91368
Ninox strenua (Powerful owl)	0.91569	0.91662	0.91341
Carcharodon carcharias (Great white shark)	0.91506	0.91385	0.91056
Cyprinus carpio (Common carp)	0.91759	0.91463	0.91045
Dicentrarchus labrax (European seabass)	0.91881	0.90116	0.91412
Poecilia reticulata (Guppy)	0.91631	0.91447	0.90864
Didelphis virginiana (Virginia Opossum)	0.91844	0.91408	0.90997
Macropus giganteus (Eastern gray kangaroo)	0.91811	0.91388	0.91113
Vombatus ursinus (Common wombat)	0.9179	0.91391	0.91207
Bos taurus (Cattle)	0.91704	0.9137	0.91125
Canis lupus familiaris (Dog)	0.91666	0.91426	0.91009
Capra aegagrus (Wild goat)	0.91783	0.9136	0.91174
Felis catus (Domestic cat)	0.91755	0.91438	0.91172
Mus musculus musculus (House mouse)	0.91641	0.91368	0.91138
Oryctolagus cuniculus (Common rabbit)	0.91665	0.91411	0.91117
Rattus rattus (House rat)	0.91655	0.91301	0.9119
Gorilla gorilla gorilla (Western lowland gorilla)	0.91509	0.91436	0.91224
Homo sapiens (Human)	0.91549	0.91484	0.91255
Lemur catta (Ring-tailed lemur)	0.91821	0.91424	0.91033
Pan paniscus (Bonobo)	0.91545	0.91465	0.91235
Pan troglodytes (Common chimpanzee)	0.91548	0.9146	0.91225
Alligator mississippiensis (American alligator)	0.91704	0.91213	0.91343
Chelydra serpentina (Common snapping turtle)	0.91732	0.9142	0.91211
Crocodylus niloticus (Nile crocodile)	0.91653	0.91448	0.91326
Crotalus horridus (Timber rattlesnake)	0.91366	0.91336	0.91345
Naja naja (Indian cobra)	0.91379	0.91192	0.913

TABLE 3 | Results of the DFA exponent for all groups and each of the three random-walk rules.

Scientific name (common name)	β_{RY}	β_{SW}	β_{KM}
Ambystoma tigrinum tigrinum (Eastern tiger salamander)	0.67836	0.90728	0.71664
Bufo gargarizans (Chusan Island toad)	0.75691	0.76766	0.76934
Rana plancyi (Eastern golden frog)	0.78803	0.74711	0.74653
Ara ararauna (Blue-and-yellow macaw)	0.74963	0.65734	0.86363
Archilochus colubris (Ruby-throated hummingbird)	0.74416	0.6971	0.86625
Columba livia (Rock pigeon)	0.76494	0.67966	0.86371
Gallus gallus (Red junglefowl)	0.7581	0.66958	0.87402
Ninox strenua (Powerful owl)	0.75282	0.6407	0.88804
Carcharodon carcharias (Great white shark)	0.74703	0.80776	0.79693
Cyprinus carpio (Common carp)	0.67192	0.75648	0.8331
Dicentrarchus labrax (European seabass)	0.75312	0.73671	0.72254
Poecilia reticulata (Guppy)	0.73809	0.78308	0.78935
Didelphis virginiana (Virginia Opossum)	0.70386	0.90263	0.7744
Macropus giganteus (Eastern gray kangaroo)	0.73178	0.8363	0.84188
Vombatus ursinus (Common wombat)	0.72691	0.83327	0.85255
Bos taurus (Cattle)	0.69678	0.84215	0.82662
Canis lupus familiaris (Dog)	0.71743	0.84081	0.79415
Capra aegagrus (Wild goat)	0.69553	0.84634	0.83395
Felis catus (Domestic cat)	0.70012	0.82755	0.82021
Mus musculus musculus (House mouse)	0.68457	0.87555	0.82526
Oryctolagus cuniculus (Common rabbit)	0.7394	0.82727	0.80349
Rattus rattus (House rat)	0.70334	0.85943	0.82893
Gorilla gorilla gorilla (Western lowland gorilla)	0.76455	0.7491	0.85718
Homo sapiens (Human)	0.76264	0.73476	0.8657
Lemur catta (Ring-tailed lemur)	0.72169	0.86066	0.81856
Pan paniscus (Bonobo)	0.76222	0.75973	0.86114
Pan troglodytes (Common chimpanzee)	0.76283	0.75342	0.86122
Alligator mississippiensis (American alligator)	0.74351	0.76308	0.84625
Chelydra serpentina (Common snapping turtle)	0.68238	0.85194	0.83671
Crocodylus niloticus (Nile crocodile)	0.69992	0.75112	0.83504
Crotalus horridus (Timber rattlesnake)	0.70735	0.75833	0.86203
Naja naja (Indian cobra)	0.69597	0.79567	0.85368

$$\mathcal{C} = \mathcal{HD} = \left(-\sum_{i=1}^{N} p_i \log_2(p_i) \right) \left(\sum_{i=1}^{N} \left(p_i - \frac{1}{N} \right)^2 \right). \quad (13)$$

variance; thus, the coefficient of complexity can be interpreted as a measure of dispersion within the information stored in a system (López-Ruiz et al., 1995).

The coefficient of complexity may thus be regarded as the Shannon entropy weighted by the coefficient of disequilibrium, which can be interpreted as the tendency of a random sequence.

TABLE 4 | New variables.

Scientific name (common name)	v_1	v_2	v_3
Ambystoma tigrinum tigrinum (Eastern tiger salamander)	−4.31380	−1.10950	−2.39820
Bufo gargarizans (Chusan Island toad)	−4.23240	−1.35480	−2.26260
Rana plancyi (Eastern golden frog)	−4.09750	−1.29530	−2.25390
Ara ararauna (Blue-and-yellow macaw)	−4.23570	−1.83360	−2.19300
Archilochus colubris (Ruby-throated hummingbird)	−4.27030	−1.84760	−2.21910
Columba livia (Rock pigeon)	−4.21960	−1.84640	−2.20990
Gallus gallus (Red junglefowl)	−4.17150	−1.91490	−2.17750
Ninox strenua (Powerful owl)	−4.21900	−2.02220	−2.21770
Carcharodon carcharias (Great white shark)	−4.18720	−1.46250	−2.31750
Cyprinus carpio (Common carp)	−4.47010	−1.58480	−2.28400
Dicentrarchus labrax (European seabass)	−4.35900	−1.18640	−2.12810
Poecilia reticulata (Guppy)	−4.20920	−1.42200	−2.30390
Didelphis virginiana (Virginia Opossum)	−4.29970	−1.33300	−2.40300
Macropus giganteus (Eastern gray kangaroo)	−4.28390	−1.68110	−2.34200
Vombatus ursinus (Common wombat)	−4.31840	−1.74230	−2.33970
Bos taurus (Cattle)	−4.37060	−1.56840	−2.34500
Canis lupus familiaris (Dog)	−4.28210	−1.42680	−2.35010
Capra aegagrus (Wild goat)	−4.35320	−1.60750	−2.34750
Felis catus (Domestic cat)	−4.39050	−1.53750	−2.34000
Mus musculus musculus (House mouse)	−4.33330	−1.55190	−2.37410
Oryctolagus cuniculus (Common rabbit)	−4.24440	−1.48650	−2.33690
Rattus rattus (House rat)	−4.33380	−1.58590	−2.35240
Gorilla gorilla gorilla (Western lowland gorilla)	−4.16280	−1.80220	−2.27510
Homo sapiens (Human)	−4.16480	−1.85860	−2.26910
Lemur catta (Ring-tailed lemur)	−4.24340	−1.54580	−2.36720
Pan paniscus (Bonobo)	−4.15450	−1.82670	−2.28690
Pan troglodytes (Common chimpanzee)	−4.15590	−1.82770	−2.28130
Alligator mississippiensis (American alligator)	−4.26190	−1.71650	−2.26170
Chelydra serpentina (Common snapping turtle)	−4.37120	−1.61300	−2.35910
Crocodylus niloticus (Nile crocodile)	−4.44740	−1.61700	−2.27800
Crotalus horridus (Timber rattlesnake)	−4.31660	−1.78940	−2.27140
Naja naja (Indian cobra)	−4.35360	−1.72560	−2.28610

3. RESULTS

The three DNA-walks for the 7 groups are depicted in **Figures 2–4**. Results for the Chargaff ratio ξ and Shannon entropy $\mathcal{H}$ are shown in **Table 1**, while **Tables 2**, **3** contain the Hurst and DFA exponents for each type of random-walk and for each sequence.

In **Figure 2**, there is an initial upward trend that is present irrespective of the species. The *RY* rule (Equation 2) implies that a (local) inclination toward the positive direction of the vertical axis corresponds to a (local) majority of purines (adenine or guanine). Similarly, the downward trend in **Figure 3** reflects a consistent predominance of the weakly-pairing bases, adenine or thymine (considering rule *SW*). Thus, adenine dominates within the range $0- \sim 3,000$ bp.

The Hurst exponents for the rules *RY*, *SW*, and *KM* (Equations 2–4, respectively) fall in the range of $0.900 - 0.912$ and imply a long-term positive autocorrelation. To put it into perspective, a Hurst exponent value of 0.9 indicates that, on average, the tendency of changes between nucleotides varies slightly as the sub-sequence size is changed. Moreover, the proximity of the Hurst exponent toward unity suggests that either purines or pyrimidines are predominant; it cannot distinguish, however, which one prevails. Similarly, the DFA exponents fall within $0.64 - 0.91$ which implies the existence of strong long-range correlations in the sequences even after detrending. Interestingly, neither the Hurst nor DFA exponent values are near zero in any of the species considered. A possible explanation is that the tendency of changes between nucleotides does not vary randomly; i.e., mtDNA has an informational structure.

For all the DNA sequences, the Chargaff ratio is positive with $\xi > 1$, implying a larger amount of pyrimidines than purines. This implication is visually reflected in the overall downward tendency of the curves in **Figure 2**.

The disequilibrium coefficient takes values $\mathcal{D} \in (0.01 - 0.03)$. From Equation (12), values near 0 imply that the probabilities $p_i \approx \frac{1}{4}$ for any of the four nucleic bases. In other words, the disequilibrium values obtained suggest that the four nucleotide bases appear with almost the same proportion within each of the 32 mtDNA sequences. This is further supported by the Shannon entropy values. In this case, Equation (10) and $N = 4$ yield a (theoretical) maximum entropy value $\mathcal{H} = \log_2(4) = 2$. Hence, the empirical entropy values $\mathcal{H} \in (1.89 - 1.97)$ suggest near-equiprobability among the nucleic bases.

A graph of $\mathcal{D}$ vs. the Shannon entropy $\mathcal{H}$ suggests a linear relation. On this account, the disequilibrium coefficient is omitted for the remainder of the study. In addition, the complexity coefficient is omitted due to its direct proportionality to $\mathcal{D}$. See **Figure S4** (Supplementary Material).

This work proposes three new evolutionary indices as functions of Shannon entropy, the Chargaff ratio, and the fractal dimensions derived from the Hurst and DFA exponents:

$$v_1 = \mathcal{H} * \log\left[\alpha'_{RY} * \xi * \log\left(\alpha'_{KM}\right)\right] \tag{14}$$

$$v_2 = \log\left[\beta'_{RY} * \log\left(\beta'_{KM}\right)\right] \tag{15}$$

$$v_3 = \log\left[\beta'_{SW} * \log\left(\alpha'_{SW}\right)\right]. \tag{16}$$

These indices reflect the long-range correlations found in DNA-walks and the information given by Shannon entropy and the Chargaff ratio.

The fractal dimensions α' and β' are derived from the Hurst and DFA exponents, respectively, using Equation (5).
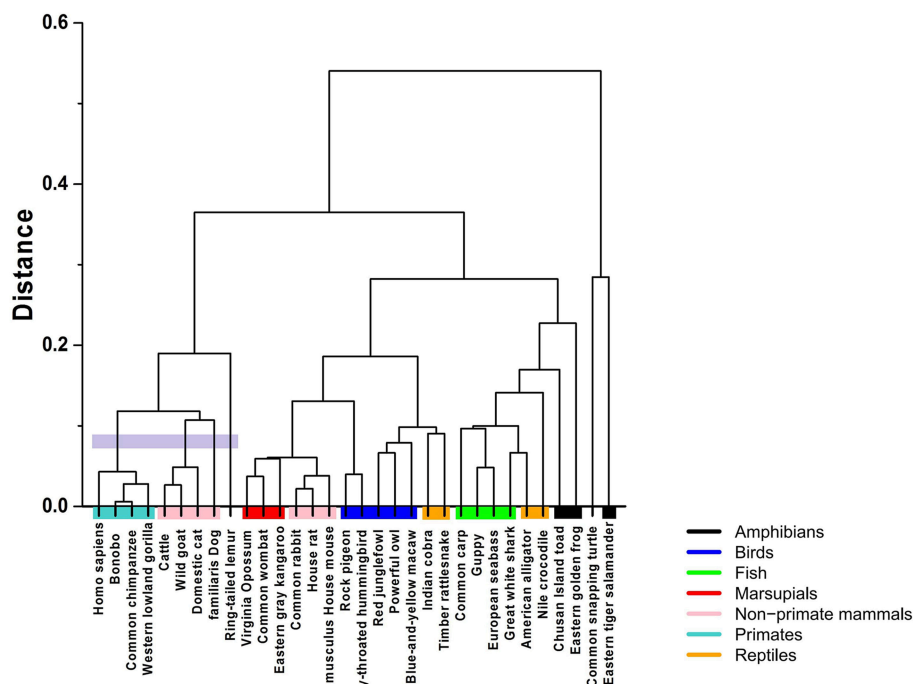
FIGURE 5 | Hierarchical clustering of the 32 species using the Hurst exponent metric with and without tendency, weighted by the Chargaff ratio and Shannon entropy.

The natural logarithm can be seen as a transformation that maximizes the differences between the coefficients. Equations (14), (15), and (16) are defined from an evolutionary perspective, while Equation (16) provides information on the energy content of sequences.

In Equation (14), the logarithm of the fractal dimension derived from the Hurst exponent using the *KM* rule provides information regarding the transversions and transitions of the entire DNA sequence. On the other hand, the Chargaff ratio is used as a weighting factor for the fractal dimension derived using the *RY* rule. The logarithm of the product of these quantities provides an evolutionary measure related to the long-range correlations. The last term in the equation (the Shannon entropy) evaluates the probability of independent nucleotide changes for a given DNA sequence.

Equation (15) uses the fractal dimensions of the DFA exponents, which are computed using the detrended DNA-walks. Therefore, it is not accurate to include the Chargaff ratio or Shannon entropy as normalization parameters. Finally, Equation (16) represents a measure of the natural selection factors in relation to the environment. Results for v_1, v_2, v_3 are shown in **Table 4**.

Clustering algorithms may benefit from the proposal. Preliminary results, shown in **Figure 5**, suggest a possible application in studies centering on the evolutionary relations among species. The proposed indices are used in the *group-average agglomerative clustering* algorithm with Euclidean metric and the sum of distances as the clustroid. Furthermore, an additional grouping was constructed using a traditional program, ClustalW, which is frequently applied to the study of phylogenetic trees, as seen in **Figure 6**.

The implementation of the algorithm using the R programming language is not computationally demanding, with running times of about 15–20 min. In comparison, ClustalW requires about 2 and a half hours for the construction of the phylogenetic tree of 32 mtDNA sequences.

The comparative analysis between the two methods shows consistency among the group of primates and other mammals sharing a common ancestry of similar lineage to the lemur. On the other hand, the marsupials and rodents (including the common rabbit) are more closely grouped with the stochastic algorithm and present a common ancestor, just as calculated by the traditional method. Other groups that share proximity with both methods are the reptiles and the birds, as well as the fish group and some amphibians.

The most pronounced differences are found in certain taxa. The proposed method relates the rabbit more closely to rodents, with characteristics similar to marsupials. Meanwhile, the traditional method positions the rabbit closer to primates. Another interesting point is that the proposed stochastic method shows that small reptiles and birds are more closely related, while the traditional method relates the birds closer to large reptiles.

4. CONCLUSIONS

As has been suggested by other studies, Shannon entropy and Hurst and DFA exponents provide insight into the properties of DNA sequences (Peng et al., 1994; Oiwa and Glazier, 2004; Melnik and Usatenko, 2014; Monge and Crespo, 2015; Namazi and Kiminezhadmalaie, 2015; Salgado-Garcia and Ugalde, 2016; Thanos et al., 2018). This exploratory analysis combines various
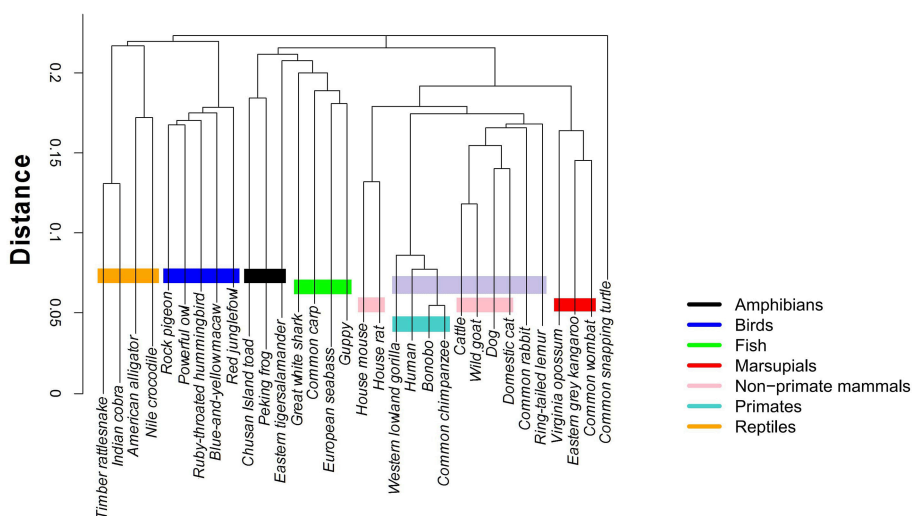
FIGURE 6 | Hierarchical clustering of the 32 species using ClustalW https://www.ebi.ac.uk/Tools/msa/clustalo/.

measures utilized in the literature to establish a biologically meaningful measure of distinction among species.

Our proposal defines new indices as functions of Shannon entropy, the Chargaff ratio, and fractal dimensions using rescaled-range analysis and DFA. These indices can be employed to construct phylogenetic trees using clustering algorithms.

Long-range correlations attributed to DNA-walks can be identified during our study. These can represent data with persistence in its evolutionary memory; i.e., that mtDNA sequences contain highly conserved regions among similar species.

The comparison between the traditional and the proposed clustering method shows clear agreements; however, there are differences that must be analyzed under an evolutionary perspective. For example, we notice that the mtDNA sequences of the common rabbit and the common snapping turtle show different properties in both methods. According to the established phylogeny, the placement of the rabbit is closer to the rodents. Interestingly, results of the stochastic hierarchical clustering suggest a potential application for phylogenetic studies.

Evolutionary processes are associated to an adaptive selection of the species throughout millions of years. However, the fluctuations of the changes in nucleotide bases could be random in order to find new sequence combinations. The proposed method attempts to measure the stochastic fluctuations to yield indices that allow the observation of tendencies and correlations in the mutations that produce new species throughout evolutionary history.

REFERENCES

Barton, N., and Jones, J. (1983). Mitochondrial DNA: new clues about evolution. *Nature* 306, 317–318. doi: 10.1038/306317a0

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

Buldyrev, S. V., Goldberger, A. L., Havlin, S., Mantegna, R. N., Matsa, M. E., Peng, C.-K., et al. (1995). Long-range correlation properties of coding and noncoding dna sequences: genbank analysis. *Phys. Rev. E* 51, 5084–5091. doi: 10.1103/PhysRevE.51.5084

Chargaff, E. (1950). Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* 6, 201–209. doi: 10.1007/BF02173653

Kandiah, V., and Shepelyansky, D. L. (2013). Google matrix analysis of DNA sequences. *PLoS ONE* 8:e61519. doi: 10.1371/journal.pone.0061519

Liao, B., and Wang, T.-M. (2004). 3-d graphical representation of DNA sequences and their numerical characterization. *J. Mol. Struct.* 681, 209–212. doi: 10.1016/j.theochem.2004.05.020

López-Ruiz, R., Mancini, H., and Calbet, X. (1995). A statistical measure of complexity. *Phys. Lett. A* 209, 321–326. doi: 10.1016/0375-9601(95)00867-5

Melnik, S., and Usatenko, O. (2014). Entropy and long-range correlations in DNA sequences. *Comput. Biol. Chem.* 53, 26–31. doi: 10.1016/j.compbiolchem.2014.08.006

Min, X. J., and Hickey, D. A. (2008). An evolutionary footprint of age-related natural selection in mitochondrial DNA. *J. Mol. Evol.* 67:412. doi: 10.1007/s00239-008-9163-8

Monge, R., and Crespo, J. (2015). Analysis of data complexity in human dna for gene-containing zone prediction. *Entropy* 17, 1673–1689. doi: 10.3390/e17041673

Namazi, H., and Kiminezhadmalaie, M. (2015). Diagnosis of lung cancer by fractal analysis of damaged dna. *Comput. Math. Methods Med.* 2015, 1–11. doi: 10.1155/2015/242695

Oiwa, N. N., and Glazier, J. A. (2004). Self-similar mitochondrial DNA. *Cell Biochem. Biophys.* 41, 41–62. doi: 10.1385/CBB:41:1:041

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Phys. Rev. E* 49, 1685–1689. doi: 10.1103/PhysRevE.49.1685

Randi, M., Vrako, M., Ler, N., and Plavi, D. (2003). Novel 2-d graphical representation of DNA sequences and their numerical characterization. *Chem. Phys. Lett.* 368, 1–6. doi: 10.1016/S0009-2614(02)01784-0

Salgado-Garcia, R., and Ugalde, E. (2016). Symbolic complexity for nucleotide sequences: a sign of the genome structure. *J. Phys. A Math. Theor.* 49:445601. doi: 10.1088/1751-8113/49/44/445601

Shannon, C. E., and Weaver, W. (1998). *The Mathematical Theory of Communication.* Urbana and Chicago, IL: University of Illinois Press.

Stoltzfus, A., and Norris, R. W. (2016). On the causes of evolutionary transition: transversion bias. *Mol. Biol. Evol.* 33, 595–602. doi: 10.1093/molbev/msv274

Thanos, D., Li, W., and Provata, A. (2018). Entropic fluctuations in dna sequences. *Physica A* 493, 444–457. doi: 10.1016/j.physa.2017.11.119

Watson, J. D., and Crick, F. H. (1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171, 737–738. doi: 10.1038/171737a0

Wikimedia Commons Contributors (2018). *File:DNA Chemical Structure.svg—Wikimedia Commons, the Free Media Repository.* Available online at: https://commons.wikimedia.org/w/index.php?title=File:DNA_chemical_structure.svg&oldid=328708739 (Accessed Feb 22, 2019).

Yamamoto, Y. (2001). "D-loop," in *Encyclopedia of Genetics*, eds S. Brenner and J. H. Miller (New York, NY: Academic Press), 539–540.

Yu, H.-J., and Huang, D.-S. (2013). Graphical representation for dna sequences via joint diagonalization of matrix pencil. *IEEE J. Biomed. Health Inform.* 17, 503–511. doi: 10.1109/TITB.2012.2227146

Zhang, Y., and Tan, M. (2007). Visualization of dna sequences based on 3dd-curves. *J. Math. Chem.* 44, 206–216. doi: 10.1007/s10910-007-9302-2

Probabilistic Assessment of Investment Options in Honey Value Chains in Lamu County, Kenya

Joshua Wafula[1]*, Yusuf Karimjee[1], Yvonne Tamba[1], Geoffrey Malava[1], Caroline Muchiri[1], Grace Koech[1], Jan De Leeuw[1,2], Josephat Nyongesa[1], Keith Shepherd[1] and Eike Luedeling[1,3]

[1] World Agroforestry Centre (ICRAF), Nairobi, Kenya, [2] ISRIC World Soil Information, Wageningen University, Wageningen, Netherlands, [3] Department of Horticultural Sciences, University of Bonn, Bonn, Germany

Designing and implementing biodiversity-based value chains can be a complex undertaking, especially in places where outcomes are uncertain and risks of project failure and cost overruns are high. We used the Stochastic Impact Evaluation (SIE) approach to guide the Intergovernmental Authority on Development (IGAD) on viable investment options in honey value chains, which the agency considered implementing as an economic incentive for communities along the Kenya-Somalia border to conserve biodiversity. The SIE approach allows for holistic analysis of project cost, benefit, and risk variables, including those with uncertain and missing information. It also identifies areas that pose critical uncertainties in the project. We started by conducting a baseline survey in Witu and Awer in Lamu County, Kenya. The aim of the survey was to establish the current farm income from beekeeping as a baseline, against which the prospective impacts of intervention options could be measured. We then developed an intervention decision model that was populated with all cost, benefit and risk variables relevant to beekeeping. After receiving training in making quantitative estimates, four subject-matter experts expressed their uncertainty about the proposed variables in the model by specifying probability distributions for them. We then used Monte Carlo simulation to project decision outcomes. We also identified variables that projected decision outcomes were most sensitive to, and we determined the value of information for each variable. The variable with the highest information value to the decision-maker in Witu was the honey price. In Awer, no additional information on any of the variables would change the recommendation to invest in honey value chains in the region. The analysis demonstrates a novel and comprehensive approach to decision-making for different stakeholders in a project where decision outcomes are uncertain.

Keywords: value chains, probabilistic projection, decision outcomes, uncertainty, value of information

INTRODUCTION

How to approve and prioritize among projects that aim at biodiversity conservation has been highlighted as one of the most critical decisions that conservation planners face [1]. This is not surprising, because conservation outcomes are often achieved through complex mechanisms, and the success of conservation actions is rarely guaranteed, with many uncertainties preventing precise

impact prediction. Success is even harder to predict, when conservation agencies aim to strengthen biodiversity indirectly, e.g., by supporting livelihoods and economies of local people as an incentive for them to conserve biodiversity outcomes [2]. Investing in biodiversity based value chains does not necessarily result in positive biodiversity outcomes. Negative impacts can arise, when value chain development results in depletion of the biodiversity that forms the resource base, on which the value chain depends (e.g. fisheries or non-timber forest products).

The production of honey is an example of a biodiversity based value chain that strengthens rather than erodes the conservation of biodiversity [3]. This is because honey producers have an interest to conserve the vegetation and plant species that produce the nectar and pollen that supports the value chain. The development of honey value chains typically revolves around a combination of introducing improved bee keeping and honey production techniques and improved access to markets for honey [4]. Yet, while attractive at first sight, such improved techniques are not always easily adopted [5]. An important reason for this is uncertainty among farmers about the financial outcomes of their investment in improved honey production techniques.

A detailed cost-benefit analysis on beekeeping projects can be considered to reduce the perceived uncertainty. However, there are rarely sufficient data on all relevant aspects of an investment decision to allow precise, purely data-driven projections to support decision-making [6]. Given such a lack of perfect knowledge, decision-makers need appropriate tools for handling uncertainties, and for identifying and prioritizing knowledge gaps, whose narrowing would reduce their chance of selecting a suboptimal decision option [7, 8]. Furthermore, decision-makers need improved capabilities to quantify risks surrounding proposed interventions, because failure to adequately account for risk can lead to high chances of project failure [9].

The Stochastic Impact Evaluation (SIE) approach allows for a structured decision analysis that incorporates all relevant variables, even those with uncertain and missing information [10]. It considers risk factors that may compromise project success or affect project performance. The approach incorporates Value of Information analysis that prioritizes critical uncertainties in a project, where further research has the greatest potential of enhancing clarity on the decisions. The present study uses the SIE approach to assess investment decisions in honey value chains for the Intergovernmental Authority on Development (IGAD) in its program on Biodiversity Management (BMP).

Study Background

IGAD-BMP partnered with the World Agroforestry Centre (ICRAF) to implement the program's biodiversity-based interventions along the Kenya-Somalia border. During the project inception phase, stakeholders were consulted, and they proposed participatory honey value chain development as one of the economic incentives for biodiversity conservation. Communities from Witu and Awer were selected to pilot the beekeeping project in Lamu County. One hundred farmers from both communities were selected and trained in beekeeping

techniques to boost their honey production knowledge and improve their access to formal markets.

The training was held in May 2015 in Malindi, Kenya. Malindi was selected for training due to its proximity to the Arabuko-Sokoke Forest, where surrounding communities intensively practice modern beekeeping. This provided an opportunity for the trainees to learn from established beekeepers. The training brought together different honey value chain actors, including representatives from the Ministry of Livestock and Fisheries in Kenya. During the workshop, stakeholders agreed that honey value chains had potential for being a viable investment option for Witu and Awer communities. A baseline survey (**Figure 1**) was commissioned to establish the current net income from beekeeping, the actual number of beekeepers among the trained farmers, beekeeping practices and risks associated with beekeeping.

In Awer, the average net income from beekeeping was 82 thousand Kenyan Shillings (Ksh) per year, approximately Ksh 225 per day (**Table 1**). Traditional beekeeping was widely practiced and characterized by high productivity and well-established markets. In Witu, the average net income from beekeeping per farmer was Ksh 39 thousand (approximately Ksh 107 per day). Very few farmers were currently practicing beekeeping. The region also had very low honey productivity and low farm gate prices, although a higher percentage of interviewed farmers were educated compared to Awer.

The baseline survey informed the project implementation approach for both communities. In Awer, it seemed sensible to support traditional beekeeping, since the region was characterized by high honey productivity and favorable farm gate prices. Establishment of a honey collection center and support of value addition activities, such as packaging, branding and processing of other bee products, were the most promising investment options. In Witu, the favored approach was to invest in modern beekeeping with emphasis on intensive farmer training and engagement of all honey value chain actors to boost productivity. Bridging existing gaps between beekeepers and formal markets was also a priority for the implementer to ensure favorable farm gate prices for the beekeepers.

While clear strategies for HVC development in the two regions thus emerged from the initial consultations and baseline survey, IGAD-BMP was still lacking certainty that the investment would raise farm incomes. To clarify the prospects for the intervention, the analysis aimed at (i) projecting impacts on beekeeping farms using a probabilistic impact evaluation approach, (ii) identifying and mitigating critical uncertainties in the project, and (iii) providing recommendations for project monitoring to reduce the chance of negative outcomes.

MATERIALS AND METHODS

We used ICRAF's Stochastic Impact Evaluation approach (**Figure 2**), which is based on the principles of Applied Information Economics (AIE) [9, 11], to project the impact of the decision on different stakeholders in the project. The method makes use of all available information, including expert
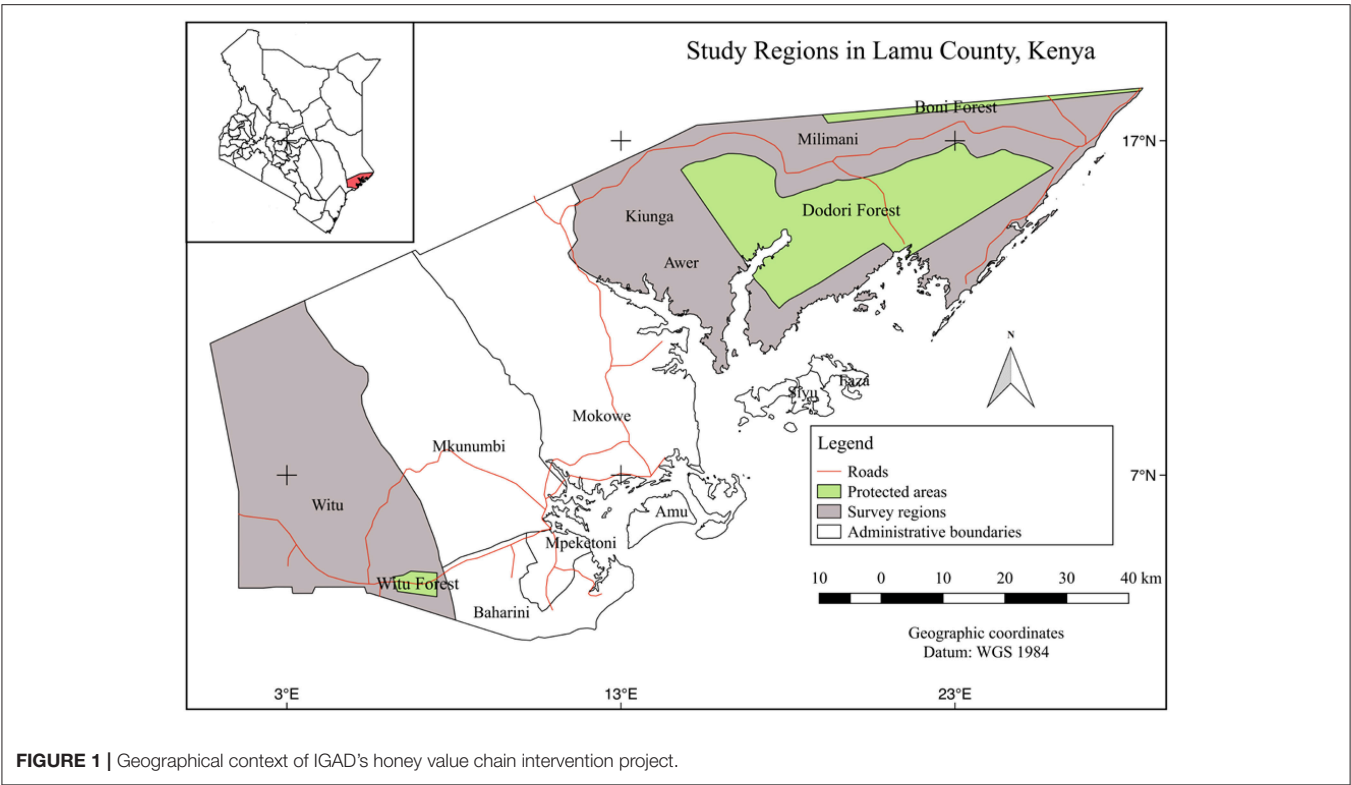
FIGURE 1 | Geographical context of IGAD's honey value chain intervention project.

TABLE 1 | Summary of the baseline survey results for net income from beekeeping in Witu and Awer.

Variable	Category	Awer		Witu	
		Number	(%)	Number	(%)
Gender	Men	37	(66)	37	(64)
	Women	19	(34)	21	(36)
Level of education	No education	32	(57)	13	(22)
	Primary	22	(39)	21	(36)
	Secondary	2	(4)	15	(26)
	Higher education	0	(0)	9	(16)
Bee keeping methods	None	2	(4)	43	(74)
	Traditional	29	(52)	8	(14)
	Modern	14	(25)	1	(2)
	Both	11	(19)	6	(10)
Hives per household		2 to 4		1 to 3	
Honey per harvest (liters)		5 to 12		2 to 6	
Income per beekeeper (thousand Ksh)		−19 to 299		−28 to 171	

knowledge, in making impact projections. The SIE approach can be applied even in the absence of perfect data. We quantified all risks, costs and benefits based on the current state of knowledge about them, considering the uncertainties about intervention outcomes that result from this. We also embraced the concept of Value of Information, which is useful for determining decision-specific knowledge gaps that decision-supporting research should address [12].

Modeling Process

The SIE process begins by thoroughly defining the decision to be made. According to Hubbard [9], a decision may merit structured decision analysis if it has two or more realistic alternatives, invokes some form of uncertainty or dilemma, has potentially negative consequences if it turns out that the wrong position was taken, and involves a decision-maker. For this study, HVC development was identified as a viable investment option during the consultative workshop, but stakeholders were still unsure that the decision to invest in HVCs would benefit them. This led to the question: *"Should IGAD-BMP invest in the proposed honey value chain intervention as an economic incentive for Witu and Awer communities to conserve biodiversity?"*

To clearly define this question, a decision analysis team was constituted with the aim of clarifying whether IGAD-BMP was interested in return on investment, who exactly the agency was targeting to raise incomes from beekeeping, and how the project was going to be implemented. The team that was convened to develop the HVC decision model consisted of 8 members: IGAD-BMP's HVC intervention project manager, the biodiversity management program coordinator, two decision analysis experts, one participatory modeling scientist and four research assistants. The initial conceptual model (**Table 2**), which aimed to reflect the overall structure of the decision at hand, was then sent to a principal bee health scientist from the International Centre of Insect Physiology and Ecology (icipe) for review. The updated model (**Figure 3**) was translated into a set of equations that represented the team's understanding of the decision's impact pathway [13].
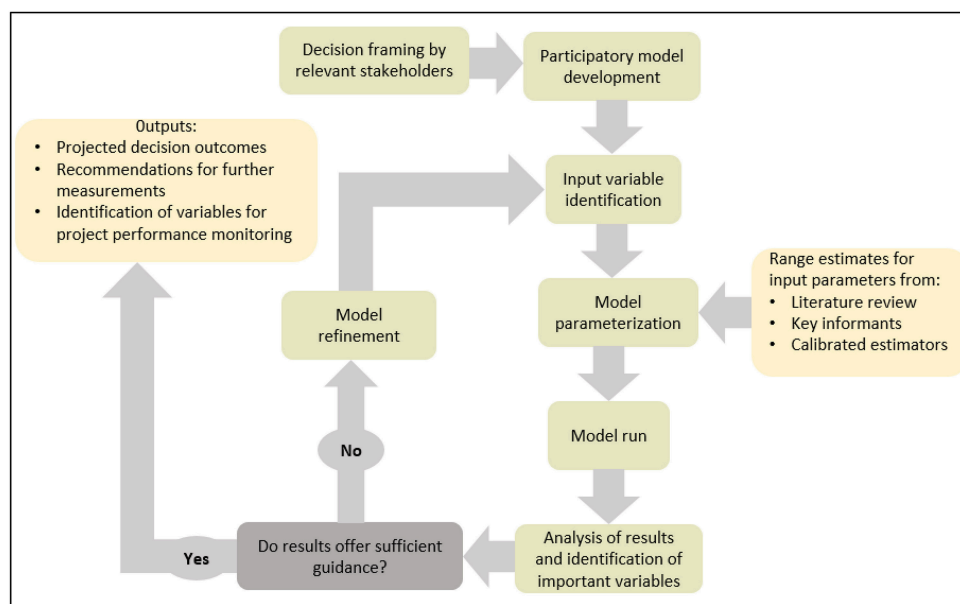
FIGURE 2 | Sequence of activities in the SIE decision modeling approach (modified from [8]).

Decision outcomes were projected using the decisionSupport package [14] for the R programming environment [15]. For this, the model equations were coded as a welfare function in R programming language (Data Sheets 3, 4). The model was parameterized using estimated probability distributions for variable values (Data Sheets 1, 2) (often specified as 90% confidence intervals), which were obtained from literature review and five established beekeepers from Mwingi in Kitui County, Kenya. The process required the team to undergo calibration training—a process that improves the capacity to make range estimates for which one is 90% confident that the actual value lies within the provided range [9]. This was done for all variables, including those for which no other information was available. Using the estimated variables as inputs, the decision model was run as a Monte Carlo simulation [16, 17]. This approach produced a distribution of possible decision outcomes by running the model a large number of times, each time fed with a different set of random draws from the defined distributions for the variables in the equations [18]). To identify variables that were most uncertain in the analysis, we used Partial Least Squares (PLS) regression [19], which was also implemented in the decisionSupport package [14].

The most uncertain variables in a project, which outcome projections are most sensitive to, are not necessarily most important to the decision-maker, because new information on them may not be able to change the recommendation emerging from the decision model. The value of additional information on a variable for decision-makers is determined by whether this information has the potential to change the sign of the expected value of the decision, which would change the preferred decision option [9].

To identify high-value variables, the Expected Value of Perfect Information (EVPI) was calculated by detecting correlations

TABLE 2 | List of variables used to populate the beekeeping model.

Benefits	Costs	Risks
Direct benefits	Initial investment costs	• Insecurity
• Revenue from sale of honey	• Cost of labor to establish apiary	• Land conflicts
	• Transport of equipment cost	• Fire outbreaks
• Revenue from sale of wax	• Cost of beehives	• Theft
	• Cost of setting up collection centers	
Indirect benefits	• Cost of farmer packages	
• Increased farm yields (from improved pollination)	• Other initial costs	
	Recurring costs	
	• Harvesting cost	
	• Wax processing cost	
	• Honey/wax transportation cost	
	• Honey processing cost	
	• Hive maintenance materials cost	
	• Opportunity cost	
	• Staff salaries	
	• Farmer training cost	
	• Other recurring costs	

between input and output variables and identifying those that were of importance to a decision-maker, i.e., the variables, whose measurement could help reduce the expected opportunity loss of the decision. Since there was no basis for deciding on the functional form of the relationship between test and outcome variable, a non-parametric test was used to detect these relationships. We used Spearman's rank correlation, using the usual statistical cutoff criterion of $\alpha < 0.05$ to exclude all variables that were not correlated with projected decision outcomes. For those that were not correlated, the value of information was zero, since the variable had no significant relationship with the outcome. This criterion excluded most
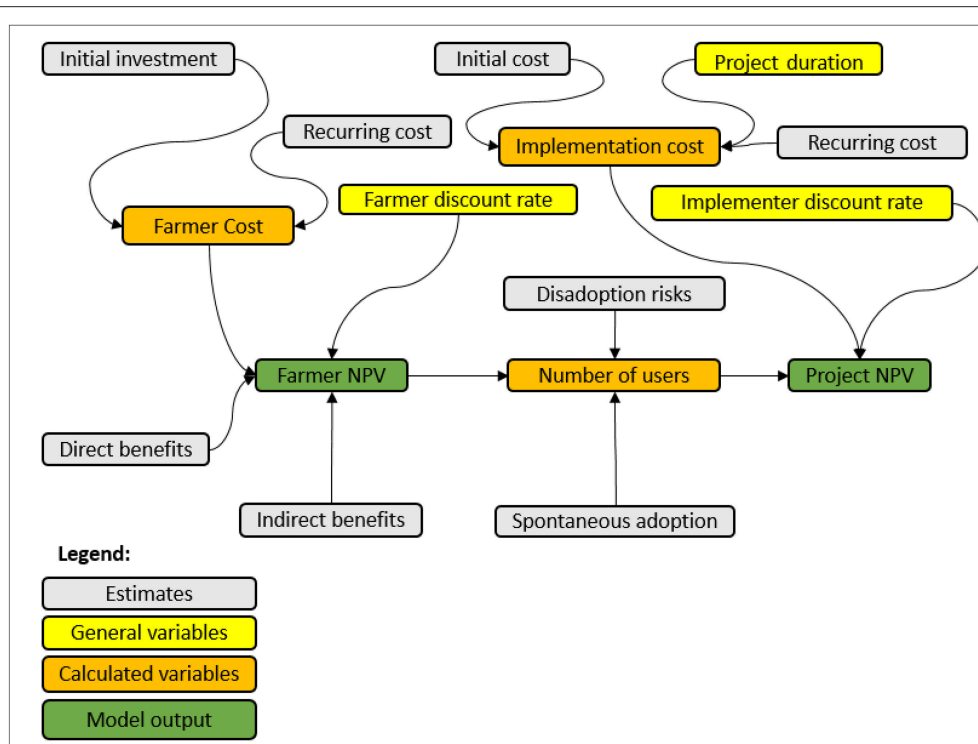
FIGURE 3 | Structure of the decision model.

variables, and discriminated well between informative and obviously uninformative variables.

We then sorted the dataset by values of each remaining input variable. In the sorted dataset, it was not possible to identify a clear threshold, where the expected decision outcomes transition from positive to negative. This was because the data was still too "noisy." To more clearly expose the effect of uncertain variables on the decision recommendation, we processed the dataset using a second-order low-pass Butterworth filter, with a critical frequency of one divided by one tenth of the number of values in the Monte Carlo output. This resulted in a smooth dataset, in which it was often possible to identify a threshold, where the sign of projected decision outcomes was reversed, e.g. for variable values above the threshold, positive outcomes would be expected, while values below the threshold would likely lead to negative outcomes (note that using a signal processing filter is a pragmatic way to solve the computational challenge of calculating the EVPI, but it introduces a small amount of inaccuracy into the threshold identification). With this threshold identified, the EVPI was computed as the sum of all outcomes with a sign that did not correspond to that of the expected value (e.g., all negative outcomes, when the analysis produced a positive expected value), multiplied by their respective chance of occurrence. This EVPI procedure was applied to all output variables.

RESULTS

Due to the economic and biophysical disparities between Awer and Witu, the two regions were modeled separately. The outcome

of the analysis was expressed as the net present value (NPV) of the intervention for a farmer and for the overall project. Emphasis was placed on the farmer NPV and overall project NPV, since the implementer's direct objective was to raise farmer incomes from beekeeping. Cash flows were also illustrated, and variables that had information value for the decision maker were identified.

Witu Community

In Witu, the median of the modeled distribution of average annual monetary income per farmer practicing beekeeping was Ksh 65 thousand, with 90% confidence that the actual income lay within the range of Ksh −36 thousand to 140 thousand for a farmer who would continuously practice beekeeping (**Figure 4**). The chance of a negative average monetary NPV for a farmer in this region was 24%.

The honey price, the amount of honey produced per hive, the number of hives per farmer, the number of harvesting seasons per year and the honey processing cost were the most important variables according to the PLS analysis. The distribution of total farmer NPV, including indirect benefits from beekeeping, e.g. through improved pollination of crops, had a median of Ksh 72 thousand, with a 90% confidence interval of Ksh-30–251 thousand.

Since both positive and negative farmer outcomes appeared plausible according to the simulation results, the EVPI analysis indicated that additional information about four variables could potentially change the recommendation emerging from the SIE process. These were the honey price, with a value of information of about Ksh 2750, followed by honey production per hive (Ksh
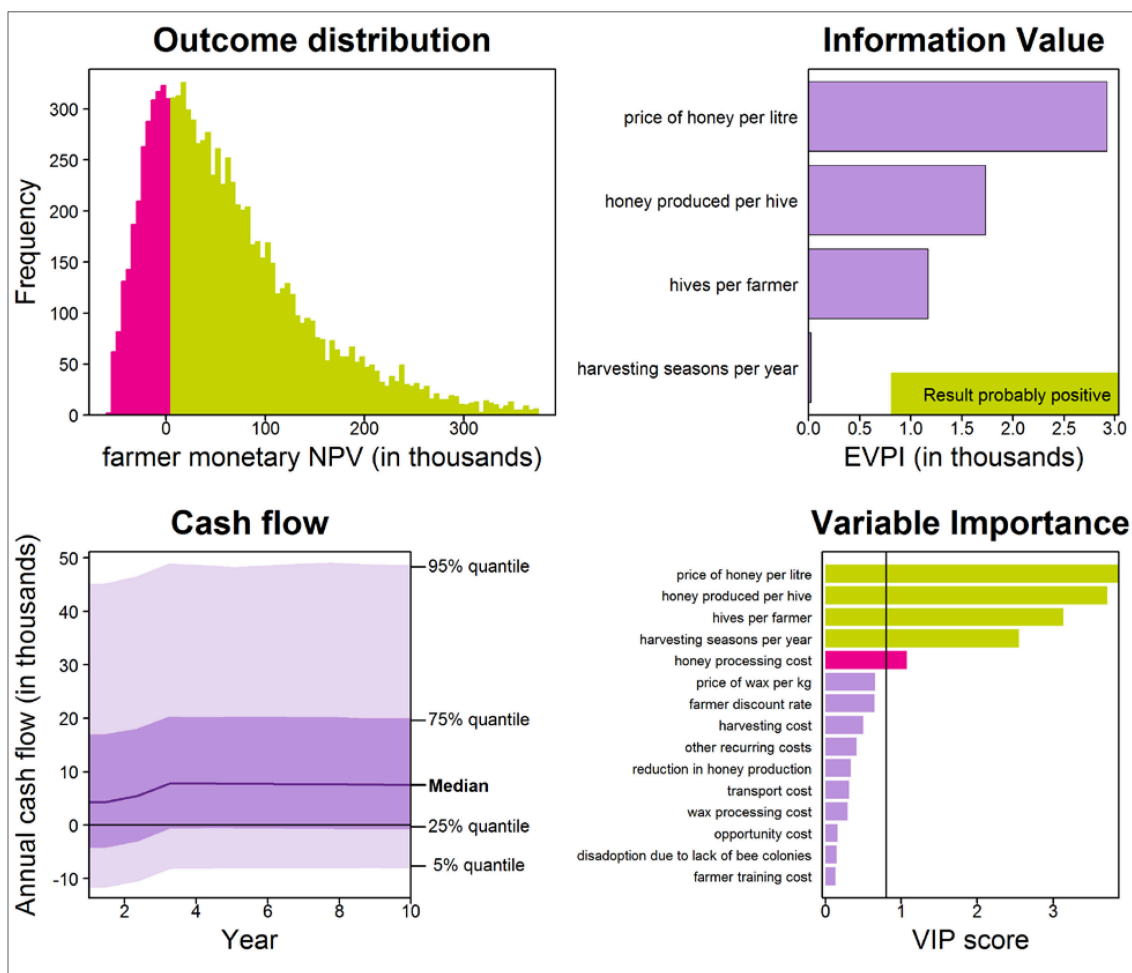
FIGURE 4 | Projected outcome distribution (top left), high-value variables (EVPI; top right), project cash flow (bottom left) and important variables (determined by PLS regression; bottom right) for a farmer practicing beekeeping in Witu, Kenya. Results were produced through Monte Carlo simulation (with 10,000 model runs) of project performance over 10 years. Red and green bars in the outcome distribution indicate positive and negative values, respectively. In the PLS plot, green bars indicate positive correlations of uncertain variables with the outcome variable, while red bars indicate negative correlations.

1750), the number of hives per farmer (Ksh 1250) and the number of harvesting seasons per year (Ksh 50).

The distribution of the projected NPV for the project in Witu had a median of Ksh 36 million, with 90% confidence that the actual NPV for the project lay within the range of Ksh-14–150 million. The model responded most sensitively to the amount of honey produced per hive. Seven other variables also had important impact on projected outcome values (**Figure 5**).

The distribution of projected decision outcomes included both positive and negative outcomes, and additional information on some individual uncertain variables had potential to change the decision recommendation. EVPI analysis indicated that information on the honey price in Witu was the most valuable to the project implementers, with a value of about Ksh 1.1 million. The amount of honey produced per hive (Ksh 0.6 million) and the number of hives per farmer (Ksh 0.4 million) were also of value to the project implementers.

Awer Community

In Awer, the distribution of projected average monetary income per farmer practicing beekeeping had a median of Ksh 130 thousand per year, with 90% confidence that the actual value lay within the range of Ksh 1.5–340 thousand per year. The chance of loss for a farmer in this region was 4.6%. Total farmer NPV including non-monetary benefits had a median of Ksh 140 thousand with a 90% confidence interval of Ksh 7.2–340 thousand. The price of honey was identified as the most uncertain variable by the PLS analysis, alongside five other variables (**Figure 6**).

EVPI analysis indicated that none of the variables had any information value for the farmer. This meant that, if the NPV was the main criterion for evaluating the attractiveness of the HVC intervention, then farmers would be well-advised to engage in beekeeping, and no additional information on any of the variables would change this recommendation.
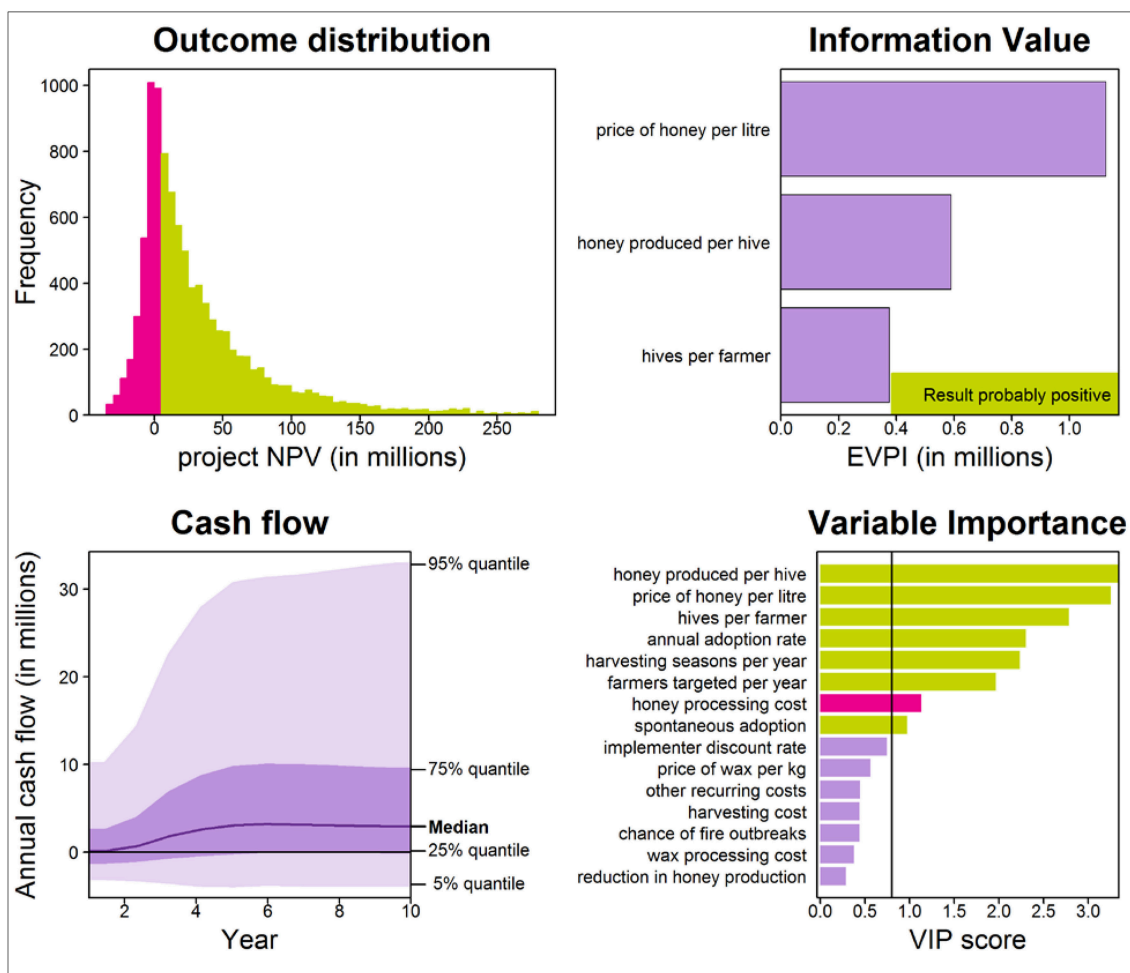
FIGURE 5 | Projected outcome distribution (top left), high-value variables (EVPI; top right), project cash flow (bottom left) and important variables (determined by PLS regression; bottom right) for the overall project in Witu, Kenya. Results were produced through Monte Carlo simulation (with 10,000 model runs) of project performance over 10 years. Red and green bars in the outcome distribution indicate positive and negative values, respectively. In the PLS plot, green bars indicate positive correlations of uncertain variables with the outcome variable, while red bars indicate negative correlations.

The overall project NPV had a median of Ksh 38 million per year, with 90% confidence that the actual value lay within the range of Ksh-0.9–120 million. PLS analysis indicated that the honey price was the most uncertain variable in the projection of project NPV in Awer (**Figure 7**). Although PLS indicated that projected project outcomes responded to variation in a number of variables, no additional information on any of these variables had the potential to change the recommendation that the project implementer should invest in honey value chains in the region.

DISCUSSION

Beekeeping is a high-risk, high-return venture that requires a well thought-out project design to maximize returns [20]. Introduction of the HVC intervention as an economic incentive for rural communities along the Kenya-Somalia border to conserve the environment may generate significant impact in terms of raising income (average of Ksh 65 thousand per farmer

for Witu and Ksh 130 thousand per farmer for Awer). However, the success of this intervention, especially for regions such as Witu, requires close monitoring by the implementing agency to minimize the chance of loss (24%).

In Witu, the most critical variable was the price of honey, which individual beekeepers have relatively little influence on. Current honey price, based on the baseline survey, is within the range of 3–7 hundred Ksh per liter. At this price, returns may be too low for a farmer who pursues beekeeping as the only source of income. For a farmer who invests in beekeeping as a supplementary source of income, the current honey price looks promising. Analyzing the impact of different factors that may affect the honey price in the future, e.g., packaging, honey quality, access to formal markets and exploitation by middlemen, will provide useful information to support farmers' decisions on investments in honey value chains. The amount of honey produced per hive is also a critical variable to farmers in Witu (**Figure 4**). Since productivity is a function that depends on many
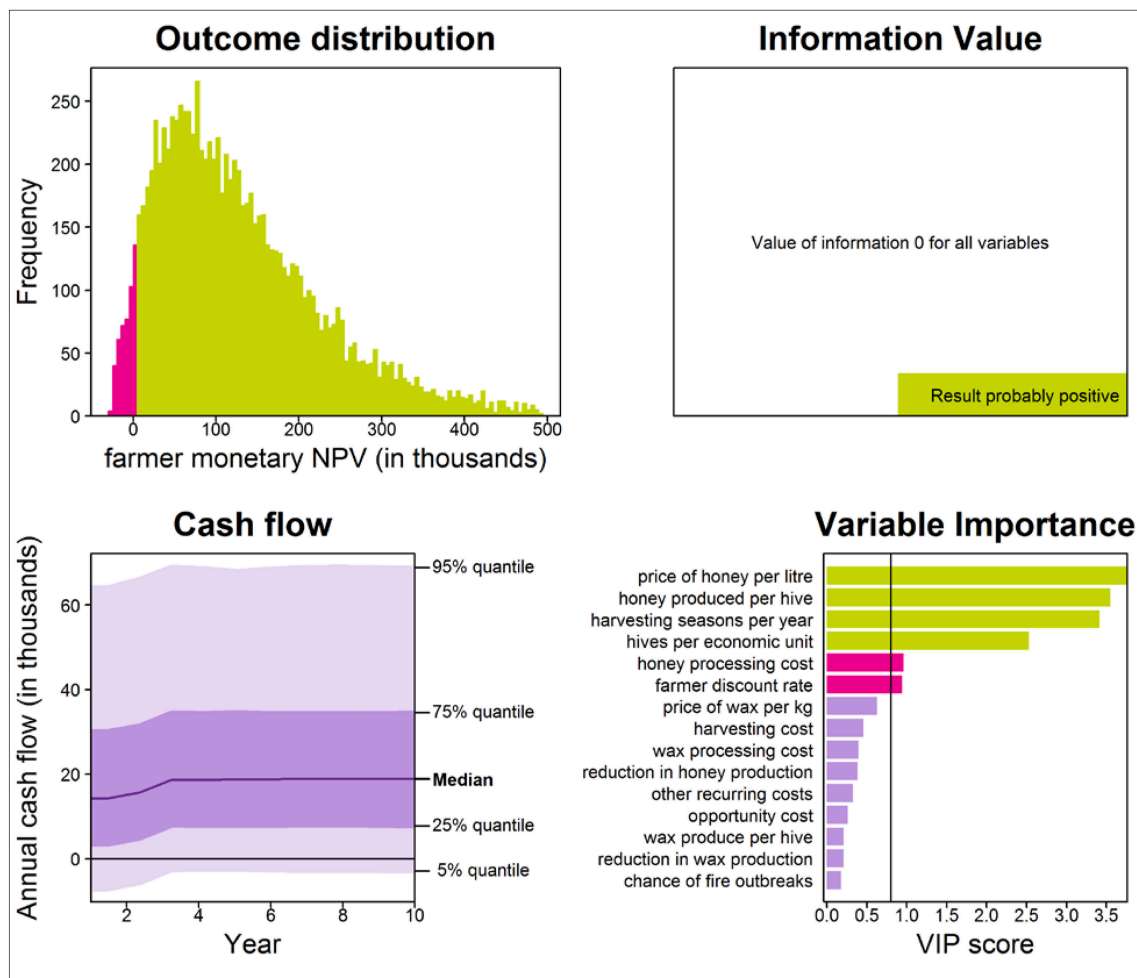
FIGURE 6 | Projected outcome distribution (top left), high-value variables (EVPI; top right), project cash flow (bottom left) and important variables (determined by PLS regression; bottom right) for a farmer practicing beekeeping in Awer, Kenya. Results were produced through Monte Carlo simulation (with 10,000 model runs) of project performance over 10 years. Red and green bars in the outcome distribution indicate positive and negative values, respectively. In the PLS plot, green bars indicate positive correlations of uncertain variables with the outcome variable, while red bars indicate negative correlations.

factors, such as good apiary management, access to extension services, technology deployed and bee forage availability, studies on these factors promise to reduce decision uncertainty for farmers. Modifications to the original project design that address the influence of these variables on the farmer NPV could reduce the chance of losses for beekeepers.

Most model runs for project NPV in Witu indicate a positive outcome, although there is still a chance (23%) of not achieving the goal of raising incomes for farmers in the region. This is because beekeeping is a relatively new venture for most farmers. The value chain is not well developed and a lot of investment has to go into farmer training to increase productivity. The honey price, the amount of honey produced per hive and the number of hives per farmer are the most critical variables for the project implementer. It is therefore important for the implementer to work with farmers to acquire more information on these variables and how they will influence their respective NPVs.

In Awer, the HVC intervention can increase monetary income for beekeepers by about Ksh 48 thousand based on the baseline survey. This is slightly higher than in Witu (Ksh 26 thousand), but this is expected, since most farmers in the region were already practicing beekeeping. The project plan to support existing traditional beekeeping methods, establish a honey collection center and support value addition activities such as packaging, branding and processing of other bee products, was likely to be more effective in raising incomes, since it would focus on bridging gaps in the chain. In Witu, the focus is on introducing the HVC to the region.

The EVPI analysis indicated that no additional information on any of the variables could change the recommendation that farmers should invest in beekeeping (**Figure 6**). This is not surprising, since the region is characterized by high honey productivity and well-defined markets. The overall project NPV in Awer suggests that introducing HVC is economically
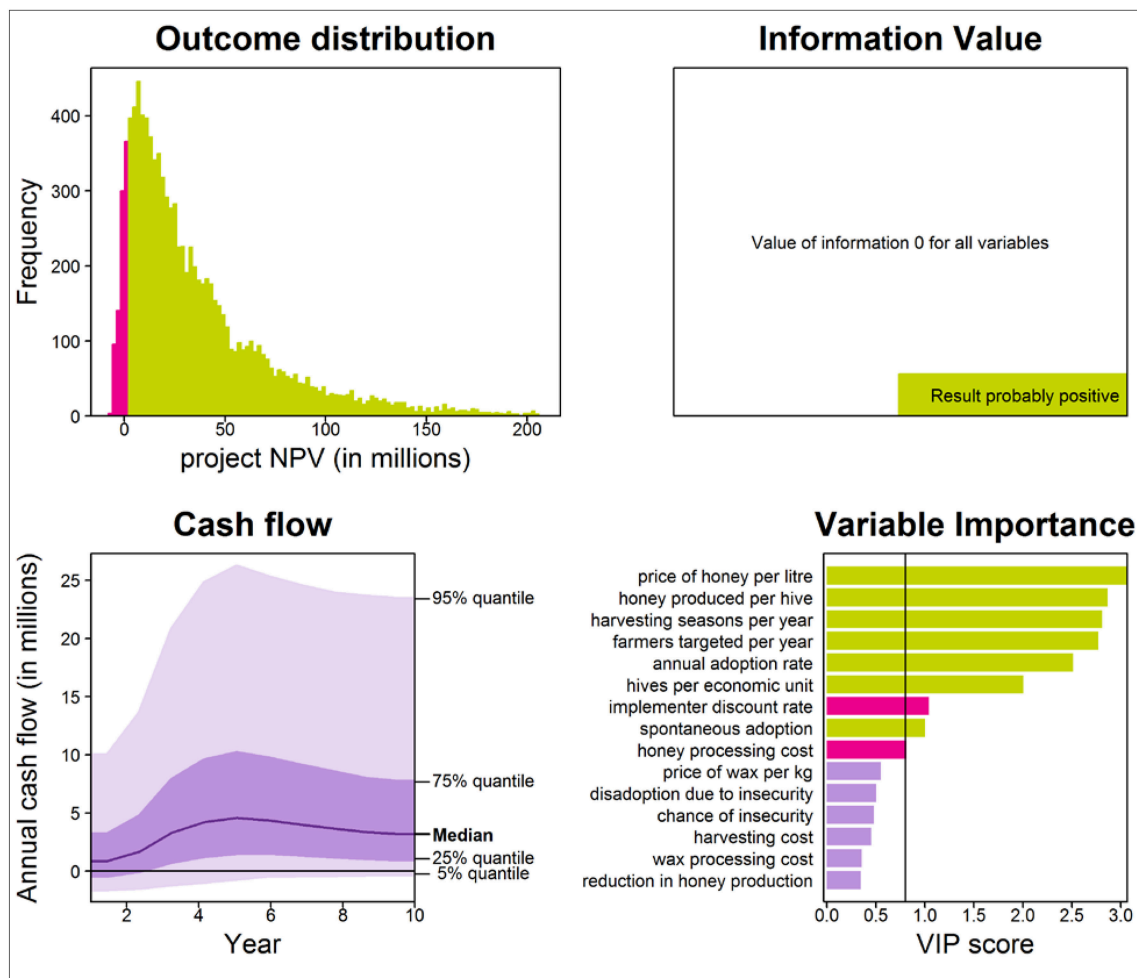
FIGURE 7 | Projected outcome distribution (top left), high-value variables (EVPI; top right), project cash flow (bottom left) and important variables (determined by PLS regression; bottom right) for the overall project in Awer, Kenya. Results were produced through Monte Carlo simulation (with 10,000 model runs) of project performance over 10 years. Red and green bars in the outcome distribution indicate positive and negative values, respectively. In the PLS plot, green bars indicate correlations of uncertain variables with the outcome variable, while red bars indicate negative correlations.

viable, even when fully considering the cost incurred by the implementing agency (**Figure 7**). The effort to improve certainty on the business case may not influence the decision to invest in HVC, since no additional information on any of the variables had value to the decision-maker.

Impact of the Intervention on Biodiversity Conservation Efforts

Biodiversity conservation is linked to income from beekeeping, because conserving the surrounding forests provides flowers for bees to forage on, thus ensuring high honey productivity throughout the year. However, raising incomes for communities within conservation areas, which is the goal of the HVC intervention, does not necessarily mean that these communities will conserve biodiversity [21]. Therefore, for IGAD-BMP to maximize its impact in terms of biodiversity conservation, the program must ensure that rural communities not only

generate sustainable incomes from honey production, but also perceive relationships between these incomes and biodiversity conservation, i.e., that high honey production correlates with success in preserving biodiversity. A study on how increased incomes from honey value chain interventions would affect biodiversity conservation outcomes could produce valuable information on the prospective impacts of the intervention on biodiversity conservation.

Recommendations for Project Monitoring

Although the intervention looks promising, a few areas need to be monitored to reduce the chance of negative outcomes. In Witu, the price of honey has to be monitored to ensure farm gate prices can provide sufficient income to bee-keeping farmers. To reduce the chance of negative outcomes as a result of low farm gate prices, there is a need to link farmers to formal markets, promote collective marketing and invest in value

addition. The amount of honey produced per hive should also be monitored, since it is critical in determining whether farmers investing in HVCs are likely to reap net benefits. Adequate farmer training and mitigation of risks, such as theft, land conflicts and fire outbreaks, will help reduce the chance of farmers abandoning honey production. Further, there is a need to identify the minimum number of hives per farmer that would provide substantial increase in income. The implementing agency would then focus on ensuring that most of the farmers have this minimum number to motivate them to continue contributing to HVCs.

In Awer, not much can be done in terms of project monitoring to reduce negative outcomes, since the chance of such negative outcomes was very low. However, the implementing agency can focus on monitoring how these positive outcomes affect biodiversity conservation efforts. There are no clear linkages between increase in income and biodiversity conservation. Clarifying and generating awareness on these linkages, e.g., through research and education on the need to conserve biodiversity to increase honey production, will help the implementing agency to better understand—as well as enhance the impact of—the project on biodiversity conservation.

Future Application of the SIE Approach

The costs, benefits and risks of biodiversity-based value chains, and agricultural interventions in general, typically have high levels of uncertainty, especially when considering the long-term and off-site effects of proposed interventions. In light of these uncertainties, the single most important task facing agricultural intervention planners is perhaps to determine the best way to make decisions [22]. Uncertainties and risks have to be quantified, policy preferences clarified and priority measurements for supporting their decisions identified.

The SIE approach can provide plausible solutions to this. The first critical step in the SIE process is to clearly define the decision to be taken. This step often does not receive sufficient attention [23], but it is instrumental in understanding what has to be measured. In this analysis, we spent about 30% of the entire effort on clarifying questions such as whether the implementing agency would cater for the full project implementation cost or whether these costs would be shared with farmers, who would be the actual beneficiaries of the project and whether the implementer was interested in return on investment. Clarifying these questions early in the implementation planning process enabled IGAD-BMP stakeholders to easily identify relevant benefits, costs, risks and external factors that would affect project performance.

The SIE approach allowed identification and prioritization of critical uncertainties in the project. Variables prioritized for further measurement in one of the regions (Witu) appeared easy to quantify with small efforts in data collection. In the other region (Awer), no measurements on any of the variables would change recommendations to farmers and the implementing agency to invest in HVCs. This demonstrates that the SIE approach has potential to substantially enhance the cost effectiveness of decision-supporting measurement campaigns for agricultural intervention planners. The approach presents a clear business case for or against investment in a project where outcomes are uncertain. This can provide critical information to investors aiming to support agricultural interventions with limited resources.

CONCLUSIONS

Investment in HVC as an economic incentive to conserve biodiversity requires thorough analysis of investment options to maximize returns to different stakeholders. This can be achieved through probabilistic stakeholder-disaggregated outcome projections for biodiversity-based value chain interventions. Clarifying the decision question turned out to be critical for the decision making process. This greatly facilitated the development of a decision's impact pathway, which could then be easily converted into a quantitative decision outcome projection model.

Tailoring interventions to meet economic, social and environmental requirements of rural communities is very important, so decision-makers need approaches that allow holistic ex-ante analysis of investment options. The cost-benefit and risk analyses for biodiversity-based value chains should consider all factors that are relevant for the implementation of the decision, even those that initially appear "intangible" or for which no data are available. Incorporation of local and expert knowledge into decision making using the SIE framework can significantly improve the quality of decisions.

Value of Information analysis can provide indications of what needs to be measured to support intervention decisions. While many uncertainties usually exist in all decisions that affect complex systems, only those uncertainties that are of value to the decision maker should be prioritized for further measurement. This can substantially reduce the cost of data collection aimed at informing decisions.

STUDY ETHICAL APPROVAL AND CONSENT

The analysis relied on information obtained from literature review, key informants and subject matter experts. This did not require ethical approval as per institutional and national guidelines, but a written and informed consent was obtained from key informants and subject matter experts. The baseline survey outcome in the study background was generated from an open access dataset, where the data was properly anonymized [24]. Written and informed consent was also obtained from all survey participants at the time of original data collection.

AUTHOR CONTRIBUTIONS

JW: The first author. YK: Contributed to the decision model development. YT: Contributed to the decision model development. GM: Contributed to the decision model development. CM: Contributed to the decision model development. GK: Contributed to the designing of the questionnaire for the baseline survey. Also provided insights on implementation of biodiversity based value chains. JD: Provided expertise on impact of biodiversity based value chains. Also contributed to the editing of the manuscript

before submission. JN: Provided vital insights in the framing of the investment decision. Also contributed to the variable estimation process. KS: Provided expertise in decision modeling. EL: Provided expertise in decision modeling and reviewed the manuscript before submission. Also provided insights on the future application of the SIE approach.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fams.2018.00006/full#supplementary-material

REFERENCES

1. Carwardine J, O'Connor T, Legge S, Mackey B, Possingham HP, Martin TG. Prioritizing threat management for biodiversity conservation. *Conserv Lett.* (2012) **5**:196–204. doi: 10.1111/j.1755-263X.2012.00228.x

2. De Leeuw J, Carsan S, Koech G, Yayé AD, Josephat N. *A Review of Best Practices for selected biodiversity value chains that promote pro-poor conservation in the Horn of Africa*. Nairobi: The World Agroforestry Centre (ICRAF) (2017).

3. Agera SIN. Role of beekeeping in the conservation of forests. *Global J Agric Sci.* (2011) **10**:27–32.

4. Figueiredo HS, Meuwissen MPM, Van der Lans IA, Oude Lansink AGJM. Identifying successful strategies for honey value chains in Brazil: a conjoint study. *Br Food J.* (2016) **118**:1800–20. doi: 10.1108/BFJ-01-2016-0026

5. Uaiene R, Arndt C, Masters W. Determinants of agricultural technology adoption in Mozambique. In: *National Directorate of Studies and Policy Analysis*. Maputo: Ministry of Planning and Development, Republic of Mozambique (2009). p. 1–31.

6. Arzberger P, Schroeder P, Beaulieu A, Bowker G, Casey K, Laaksonen L, et al. Promoting access to public research data for scientific, economic, and social development. *Data Sci J.* (2004) **3**:135–52. doi: 10.2481/dsj.3.135

7. Peterman RM, Anderson JL. Decision analysis: a method for taking uncertainties into account in risk-based decision making. *Hum Ecol Risk Assess Int J.* (1999) **5**:231–44. doi: 10.1080/10807039991289383

8. Luedeling E, Oord AL, Kiteme B, Ogalleh S. Malesu M, Shepherd KD, et al. Fresh groundwater for Wajir—ex-ante assessment of uncertain benefits for multiple stakeholders in a water supply project in Northern Kenya. *Front Environ Sci.* (2015) **3**:16. doi: 10.3389/fenvs.2015.00016

9. Hubbard DW. *How to Measure Anything: Finding the Value of Intangibles in Business*. Hoboken, NJ: Wiley (2014).

10. Luedeling E, Shepherd K. Decision-focused agricultural research. *Solutions* (2016) **7**:46–54.

11. Hubbard D. Applied information economics : a new method for quantifying IT value an executive overview. *Dec Anal.* (2004) **6**.

12. Koops MA. Reliability and the value of information. *Anim Behav.* (2004) **67**:103–11. doi: 10.1016/j.anbehav.2003.02.008

13. Thabrew L, Wiek A, Ries R. Environmental decision making in multi-stakeholder contexts: applicability of life cycle thinking in development planning and implementation. *J Clean Prod.* (2009) **17**:67–76. doi: 10.1016/j.jclepro.2008.03.008

14. Luedeling E, Göhring L. *decisionSupport: Quantitative Support of Decision Making under Uncertainty, CRAN archive.* Available online at: https://cran.r-project.org/web/packages/decisionSupport/ (2017).

15. Core Team R. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing (2017).

16. Velez-Pareja I. Including the risk in the analysis: montecarlo simulation (Análisis De Inversiones Bajo Riesgo: Simulación). *Soc Sci Res Netw.* (2007). doi: 10.2139/ssrn.986972

17. Rosenstock TS, Mpanda M, Rioux J, Aynekulu E, Kimaro AA, Neufeldt H, et al. Targeting conservation agriculture in the context of livelihoods and landscapes. *Agric Ecosyst Environ.* (2014) **187**:47–51. doi: 10.1016/j.agee.2013.11.011

18. Chen Y, Yang Y, Luo J. Uncertainty analysis of coalbed methane economic assessment with montecarlo method. *Pro Environ Sci.* (2012) **12**:640–5. doi: 10.1016/j.proenv.2012.01.329

19. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. In: *Chemometr Intell Lab Syst.* (2001) **58**:109–30. doi: 10.1016/S0169-7439(01)00155-1

20. Mujuni A, Natukunda K, Kugonza DR. Factors affecting the adoption of beekeeping and associated technologies in Bushenyi District, Western Uganda. *Livestock Res Rural Dev.* (2012) **24**.

21. Adams WM. Biodiversity conservation and the eradication of poverty. *Science* (2004) **306**:1146–9. doi: 10.1126/science.1097920

22. Lindenmayer DB, Gibbons P, Bourke M, Burgman M, Dickman CR, Ferrier S, et al. Improving biodiversity monitoring. *Aust Ecol.* (2012) **37**:285–94. doi: 10.1111/j.1442-9993.2011.02314.x

23. Rosenzweig P. *What Makes Strategic Decisions Different*. Boston, MA: Harvard Business Review (2013).

24. Wafula J, Koech G, Nyongesa J, Shepherd K, Luedeling E, De Leeuw J. Raw data for net income from bee-keeping in Lamu County, Kenya. *ICRAF Dataverse* (2017). doi: 10.7910/DVN/JNOEZM

The Amnesiac Lookback Option: Selectively Monitored Lookback Options and Cryptocurrencies

Ho-Chun Herbert Chang* and Kevin Li

Department of Mathematics, Dartmouth College, Hanover, NH, United States

This study proposes a strategy to make the lookback option cheaper and more practical, and suggests the use of its properties to reduce risk exposure in cryptocurrency markets through blockchain enforced smart contracts and correct for informational inefficiencies surrounding prices and volatility. This paper generalizes partial, discretely-monitored lookback options that dilute premiums by selecting a subset of specified periods to determine payoff, which we call amnesiac lookback options. Prior literature on discretely-monitored lookback options considers the number of periods and assumes equidistant lookback periods in pricing partial lookback options. This study by contrast considers random sampling of lookback periods and compares resulting payoff of the call, put and spread options under floating and fixed strikes. Amnesiac lookbacks are priced with Monte Carlo simulations of Gaussian random walks under equidistant and random periods. Results are compared to analytic and binomial pricing models for the same derivatives. Simulations show diminishing marginal increases to the fair price as the number of selected periods is increased. The returns correspond to a Hill curve whose parameters are set by interest rate and volatility. We demonstrate over-pricing under equidistant monitoring assumptions with error increasing as the lookback periods decrease. An example of a direct implication for event trading is when shock is forecasted but its timing uncertain, equidistant sampling produces a lower error on the true maximum than random choice. We conclude that the instrument provides an ideal space for investors to balance their risk, and as a prime candidate to hedge extreme volatility. We discuss the application of the amnesiac lookback option and path-dependent options to cryptocurrencies and blockchain commodities in the context of smart contracts.

Keywords: options pricing, lookback options, path-dependent options, Monte-Carlo methods, cryptocurrency, smart contracts

1. INTRODUCTION

Lookback options have been a prototypical example of exotic options within the financial literature [1]. The option gives the holder the right to buy or sell an underlying asset at any price attained within a specified "lookback" period. The payoff of a lookback call (put) is therefore the difference between the underlying price at maturity and the maximum (minimum) price attained.

The trader is thus able to capitalize on the underlying asset as if he sold it at the optimal time. By utilizing only the highest value of the underlying asset in determining payoff, lookback options capture the best case scenarios that people would like to sell at, but often miss due to uncertainty. By corresponding payoffs to the extreme movements of underlying asset prices, lookback options allow for investors to hedge against or invest in volatility.

Goldman et al. [2] [3] are one of the first who mentioned the lookback option within the financial literature. Their paper examines the properties of path-dependent European options under Black-Scholes assumptions. They laid out three primary motivations: to minimize regret, to live out the "fantasy" of buying low and selling high, and lastly, to use knowledge of the range to expand an investors opportunity set. This last condition is only applicable within the realistic market setting rather than a frictionless one. The nature of lookback options also lets investors guard against the behavioral flaws in people. The nature of lookback options also lets investors guard against the behavioral flaws in people. Because the payoffs of lookbacks are path dependent in a way to capture the effects of the best prices, lookbacks can hedge not only quantitative variations in the market, but irrational regret of the human mind.

Lookback options benefit the holder given greater volatility in the market. Thus, it is an effective instrument for hedging against large price movements and reducing the risk of destabilizing events that may cause markets to either rise or fall. These include political election outcomes or other unforeseen chance occurrences such as market anomalies [4]. Also among its potential uses is to hedge exposure in cryptocurrency markets, where extreme volatility has increasingly become a menace as investors look for ways to mitigate such risk.

However, the lookback option's strong reduction to risk exposure requires a sufficiently high premium, which reduces the option's demand. The proposition of this paper recognizes that one may not have to look back upon the entire associated time period; merely looking back upon a portion suffices to capture most of the max potential payoff of a full-time lookback option, most of the time. This conclusion allows for investors to adjust their risk using lookbacks of differing lookback period lengths by purchasing shorter time lookback options at a discount in order to increase return on investment. This strategy suggests the possibility of event-based trading whereby lookback periods may be chosen to correspond to market shaping events in hopes of extraordinary profits.

As the potential payoffs and insurance capabilities of an option increase, so does its premium. Thus, there is a desire for an instrument that limits risk exposure, but at a lower price. Here is where the proposed "Amnesiac Lookback Option" may find a tradeable niche, a variant of the conventional lookback option whose lookback time periods are restricted to a chosen subset of periods within a specified time interval at the time the contract is created. If premiums for lookbacks may be reduced, it could also open the door for their widespread use, particularly in high volatility cryptocurrency markets. The purpose of this study is therefore 2-fold: to propose a way to make the lookback option cheaper and more practical, and suggest the use of its properties to reduce risk exposure in cryptocurrency markets

and correct for informational inefficiencies surrounding prices and volatility. Additionally, we are interested in the following research questions. What periods of the amnesiac lookback option should be selected to maximize its final payoff? What is the probability that the amnesiac lookback option may produce greater profit than the standard lookback option? Lastly, would this option have practical applications in the cryptocurrency market?

The rest of the paper is organized in the following way: The rest of the introduction provides more context on the varieties of lookback options and applications to cryptocurrencies. Section 1 continues to discuss types of lookbacks and its application to the cryptocurrency market. Section 2 discusses existing literature on *partial lookback options*, lookback options with diluted premiums, and past efforts to price them. It then discusses our methods of simulation. Section 3 presents the results of three types of numerical simulations. First, a demonstration of the general properties of floating amnesiac lookback options for calls, puts and spread options are shown. Second, the payoff of fixed strike amnesiac lookback options is shown and compared to the fixed strike. Third, a study of equidistant monitoring vs. two forms of random monitoring are presented. Section 4 prices different cryptocurrency using the algorithm in conjunction with smart contracts.

1.1. Fixed vs. Floating Lookback Options

Lookbacks exist in many different varieties but can be classified into two broad categories: fixed and floating strikes. The strike price of fixed strikes is indicated in the initial contract, while for floating strike lookbacks the strike price is the minimum (call) or maximum (put) of the underlying asset. Additionally, there are spread lookback options with payoffs equal to the difference between the maximum and minimum prices attained by the underlying. The payoffs are summarized in **Table 1** where $M_0^T = \max\{S_i\}_{i=0}^T$ and $m_0^T = \min\{S_i\}_{i=0}^T$. T denotes the time of maturity.

Suppose an investor decides to take the long position for the next 2 months. However, the price of the stock drops unexpectedly within the last 4 days of closing by 10%. Not selling the stock early becomes a reason of regret. Similarly, the stock may drop in the span of 2 days before rising quickly, making the investor regret not buying the stock a little later. Thus the floating strike call lookback is useful for market exit and the fixed strike call for market entry.

1.2. The Partial and the Amnesiac

Due to the high premiums of lookback options and perhaps other factors, lookback options are not sold at particularly high volumes within OTC trades [5]. It has been suggested that partial lookback options may allow lookbacks to be more tradeable. One

---

TABLE 1 | Payoffs for floating and fixed strike lookback options.

Option name	Floating	Fixed (strike = K)
Call	$S_T - m_0^T$	$M_0^T - K$
Put	$M_0^T - S_T$	$K - m_0^T$
Spread	$M_0^T - m_0^T$	$M_0^T - m_0^T - K$

such way is to introduce a factor λ that decreases the effect or payoff of the lookback, proposed by Conze and Viswanathan [6]. For instance, the payoff of such an instrument is:

$$\lambda(\max\{S_i\}_1^N - S_T) \qquad (1)$$

If λ equals 1 then payoff is identical to a standard lookback option. Another form of the partial lookback, referred to as the fractional lookback, restricts the lookback periods to a continuous subset of days. Finally there is also the window lookback option, where separated continuous periods are monitored [7].

The amnesiac lookback option generalizes partial and discrete lookback options that are not linearly reliant on the payoff of the standard lookback option and is named such in response to Heynen and Kat's paper *Selective Memory*. The amnesiac lookback option is a lookback option whose lookback periods belong to a pre-specified subset of periods from a given time interval, determined at the moment of contract creation.

The amnesiac lookback option is thus defined as follows. Given a full, discrete lookback option with N total lookback periods, an amnesiac option is defined by a subset of periods of the full, discrete lookback option, denoted as A. The payoff scheme is the same as the standard lookback, while the extreme values of the option are restricted to A. This means, every full, discrete lookback option is in fact an amnesiac option of a full lookback with greater periods. This extends previous definitions of partial lookbacks to beyond equidistant intervals.

1.3. Application to Cryptocurrencies

Lookback options are extremely suited to hedge against volatility in general, whether the underlying asset surges in value, or whether the underlying asset declines in value. Unpredictability is therefore the prime feature of an asset that would drive demand and usage for a lookback option based upon that asset. In 1982, the Mocatta Metals Corporation issued one of the first "lookbacks," that allowed a trader to buy gold at the lowest price attained within a period. In the context of modern day financial markets, it would seem lookbacks could have high potential in hedging investments involving the popular digital gold cryptocurrencies. A pioneering innovation within currency markets, instruments such as Bitcoin and Ethereum may represent the future of monetary exchange. Given its relative technology security, explosion in value, and increasing acceptance of legitimacy, Carrick [8] has even suggested that cryptocurrencies like Bitcoin could be used as a complement to fiat currencies in emerging markets. At the same time, high volatility has become the premier feature of cryptocurrency markets, which has made investment risky. A Bitcoin was worth $2 in 2011—and exploded to $4,000 in 2017 [9]. During that time, volatility for daily returns would regularly exceed 10%, even skyrocketing to 16% [10]. Bitcoin once lost 75% of its value over 2 years, and then rising 2,000% within the next two.

Past literature has found evidence of time series characteristics and long memory behavior in Bitcoin markets, both in regards to pricing and in return volatility. Bariviera et al. [11] explores some of these inefficiencies by calculating Hurst exponents via the Detrended Fluctuation Analysis method for certain time windows of Bitcoin return and return volatility data. Bariviera finds evidence that from 2011 to 2014, daily Bitcoin returns had long-term positive autocorrelation with previous returns and that return volatility had long memory throughout the entire time period of 2011–2017 [11]. Bariviera et al. [11] utilize this same methodology to further show that intraday returns before 2014 exhibit long range memory as well. Additionally, Urquhart [12] finds that Bitcoin prices exhibit odd behaviors through the entire time period May 2012 to April 2017, with over 10 percent of prices ending with decimal digits of 00. One, two, three, five, and ten days before a round number from rising prices, returns are positive and statistically significant. Meanwhile, prices after a round number show no such predictable behaviors. Given that the lookback option utilizes the time series behavior of an asset's prices and volatility to determine its payoff, it is a suitable candidate for rectifying such market inefficiencies. The amnesiac lookback option is particularly appropriate because lookback period selection allows for targeted correction of unusual events, such as those found in Urquhart [12], and can more easily eliminate arbitrage opportunities and bring markets into more efficient states.

Ultimately, Bitcoin aims to be a type of fiat money; it has no value backed with consumable goods and its value comes from the minds of its investors and from the financial environment in which it occupies. Increases have been driven by hopes of future value, or in other terms, heavy speculation. Unexpected events therefore have rippling effects on the cryptocurrency market and create a volatility with few equal comparisons. China's decision to cease its bitcoin exchange in September of 2017 sent Bitcoin into a downward spiral, and as governments and regulators venture forth into the frontier that is the cryptocurrency market, events that will shock the market are sure to take place [13]. Even so, investors are increasingly accepting of cryptocurrencies as attractive investment propositions outside of speculation. Bouri et al. [14] find that Bitcoin can act as a hedge against market uncertainty in situations, specifically in short time horizons under extreme bear or bull market regimes, or when uncertainty is either very low or very high. Bouri et al. [15] overall minimizes the usefulness of Bitcoin as a general hedge, but still finds evidence that investors put money into Bitcoin when Asian stocks experience extreme down movements.

In July 2017, the U.S. Commodity Futures Trading Commission granted federal approval to LedgerX LLC for the ability to trade and exchange options based on cryptocurrency [16]. This happening marked the birth of the first federally regulated platforms of cryptocurrency options trading, and moved items such as Bitcoin, Ethereum, and even Dogecoin closer into the realm of financial legitimacy. Shortly after on December 1, 2017, the U.S. Commodity Futures Trading Commission announced the offering of self-certified derivatives from three financial firms: bitcoin futures from the Chicago Mercantile Exchange Inc. and the CBOE Futures Exchange and binary options from the Cantor Exchange [17]. In contrast with the effect of regulatory restrictions, this decision has driven up bitcoin prices from $5,000 prior to the announcement to above $11,000 by December 05, 2017 [18]. The issuance of new

financial instruments can thus feedback into the price of its underlying.

More importantly, the integration of smart contracts into blockchain technology has vastly expanded the possibilities of options trading, allowing traders great flexibility in designing their own options [19–21]. Smart contracts are computer protocol that allow for trade and exchange without the need for an intermediary, and because writers code their own contracts, blockchain technology can facilitate the trade of highly customized contracts often seen on the OTC market. Both Bitcoin and Ethereum, for instance, support programming languages that allow for the creation of custom smart contracts [22, 23]. Amnesiac lookback options could conceivably exist within this structure, and be traded as a smart contract.

Recent events lent credence to the prospect of a prolific smart contract exchange. In March 2018. The state of Tennessee signed into law a bill that recognizes smart contracts as having legal power [24], providing a pathway for other states to follow and further securing the legitimacy of these digital arrangements. Blockchain based smart contract firm Hedgy has also created irrefutable and unalterable "Smart Futures" that can enforce digital obligations and streamline settlements [25]. As more and more investors begin creating and trading new smart contracts, blockchain may 1 day even "democratize" the OTC market and open a plethora of new smart instruments to be exchanged.

Lookback options therefore emerge as an attractive and easily implementable product for currency speculators. As price fluctuation is the determining factor of payoff, such lookback instruments would be highly priced in the cryptocurrency market and be a key counterweight to the speculative risk of investors. Simple to understand and able to exist as smart contracts, lookback options and other path-dependent options could serve as a unique tool for traders with restricted access to dynamic trading strategies, and in the process introduce greater flows of capital from these sources to promising investments.

2. MATERIALS AND METHODS

2.1. Past Pricing Model for Lookbacks

Research with lookback options fall within pricing methods for path dependent exotic options, such as the barrier option which also rely on extreme value processes, and the Asian option which relies on the average price of the underlying. Goldman introduced the instrument into the financial literature in 1979 [2, 3]. Following Black, Scholes and Merton's pricing of vanilla options, Black-Scholes assumptions have been extended to price exotic options. This is defined by at least one asset with price S that moves under Geometric Brownian Motion with constant drift and volatility. The underlying price follows:

$$dS = S\mu dt + \sigma S dX \tag{2}$$

where the underlying has drift μ and volatility σ. In the risk-neutral measure, the stock price at a given time t and final time T is given below:

$$S_{t+1} = S_t e^{(r-D-\frac{\sigma^2}{2})\Delta t + \sigma\epsilon\sqrt{\Delta t}} \tag{3}$$

$$S_T = S_0 e^{(r-D-\frac{\sigma^2}{2})T + \sigma\epsilon\sqrt{T}} \tag{4}$$

r denotes the risk free rate, σ the volatility, D denotes dividends in the case of stocks, and $\epsilon \sim \mathcal{N}(0,1)$, the normal distribution.

Like the Vanilla European Option, an analytic pricing formula has been shown using martingale methods [26], with the price of a call given as:

$$C_t = S\Phi(a_1(S,m)) - me^{-r\tau}\Phi(a_2(S,m)) - \frac{S\sigma^2}{2r}(\Phi(-a_1(S,m))$$
$$-e^{-r\tau}(\frac{m}{S})^{\frac{2r}{\sigma^2}}\Phi(-a_3(S,m))) \tag{5}$$

$$P_t = -S\Phi(-a_1(S,M)) + Me^{-r\tau}\Phi(a_2(S,M)) + \frac{S\sigma^2}{2r}(\Phi(a_1(S,M))$$
$$-e^{-r\tau}(\frac{M}{S})^{\frac{2r}{\sigma^2}}\Phi(a_3(S,m))) \tag{6}$$

where M denotes the running maximum at time t, m denotes the running minimum, $\tau = T - t$ with T the time of maturity, and Φ the standard normal cumulative distribution function given as

$$\Phi(\alpha) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha} e^{-\frac{x^2}{2}}dx$$

With L being a dummy variable, the variables denote:

$$a_1(S,L) = \frac{\ln\frac{S}{L} + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$
$$a_2(S,L) = a_1(S,L) - \sigma\sqrt{\tau}$$
$$a_3(S,L) = a_1(S,L) - \frac{2r\sqrt{\tau}}{\sigma} \qquad \text{for } L > 0, S > 0$$

For fixed strikes, Conze and Vishwanathan [6] used known properties of maxima and minima distributions to price the lookback call, under the continuous case of lookback periods.

$$C_t = \begin{cases} S_0\Phi(b(T)) & -e^{rT}K(b(T) - \sigma\sqrt{T} + e^{-rT}\frac{\sigma^2}{2r}S_0\left(e^{rT}\Phi(b(T))\right. \\ & \left. -\left(\frac{S_0}{K}\right)^{\frac{-2r}{\sigma^2}}\Phi(b(T) - \frac{2r}{\sigma}\sqrt{T})\right) \qquad K \geq M \\ e^{rT}(M_T^0- & K) + S_0\Phi(b'(T)) - e^{-rT}M_T^0\Phi(b'(T) - \sigma\sqrt{T}) \\ & +e^{-rT}\frac{\sigma^2}{2r}S_0(e^{rT}\Phi(d'_T) - \left(\frac{S_0}{M_T^0}\right)^{\frac{-2r}{\sigma^2}}\Phi\left(b'(T)\right) \\ & -\left(\frac{2r}{\sigma}\sqrt{T}\right) \qquad K < M \end{cases}$$

where

$$b(T) = \frac{\ln\frac{S_0}{K} + (r - D + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

$$b'(T) = \frac{\ln\frac{S_0}{M_0^T} + (r - D + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \tag{7}$$

The similar case for the put is derived in their paper, and resembles the form of the float-strike option in Equation (5). These analytic solutions assume continuous monitoring and updates to the price of the underlying, while in reality most

options are monitored at discrete intervals. Thus, the analytic solutions are typically overpriced as compared to the prices of discretely monitored lookbacks, and has spurred two areas of study: how to either price discrete options more accurately, and how to correct errors of continuous monitoring.

Historically, notable papers that have priced discrete lookback options include those by Heynen and Kat using the Black-Scholes model [27], Babbs using a binomial valuation [28], and Cheuk and Vorst [29] using a binomial model. More recent pricing paradigms utilize stochastic volatility [30] and the Lévy Model [31] (more citations here). Cheuk and Vorst in particular consider the effects of observation frequency on price using equal lookback intervals.

The binomial options pricing model first described in Cox et al. [32] offers a conceptual method with which the price of an amnesiac lookback options may be determined. Under the assumptions of the binomial tree model, the price of an underlying stock S_t at discrete time t may strictly exhibit one of two behaviors. It may either increase in value to $S_t u$ with probability p, or decrease in value to $S_t d$ with probability q where $u = e^{\sigma \sqrt{t}}$, $d = 1/u$, $p + q = 1$, and sigma is a measure of the stock's volatility. The risk-neutral probability p is determined to equal $\frac{e^{(r\Delta t)} - d}{u - d}$ where r an exogenously determined risk free return. The lookback option can then be priced recursively using a lattice.

This pricing model was then extended to be more efficient by Conze and Vishwanathan [6], Babbs [28], and Cheuk and Vorst [29]. The binomial pricing method lends itself more easily to pricing at discrete intervals. With too few time intervals, the approximation for the standard lookback option becomes inaccurate. Yet as the number of periods in the binomial tree increases, so does the amount of computations. Past work involves increasing the efficiency of these computations using combinatorial properties of the binomial tree. Babbs [28] uses a change of numeraire approach, similar to Hull and White [33], to price both new and existing lookback options. Cheuk and Vorst formulate a modified tree, where the computation only relies on time and the difference between u and d. Let this variable be called j such that:

$$S_t = \left(\min_{t \leq T} S_t \right) u_n^j \qquad S_t = \left(\max_{t \leq T} S_t \right) u_n^{-j} \qquad (8)$$

for the call and the put respectively. Then the price for a floating strike lookback can be given as $C_n^{fl} = S_0 V_n(0,0)$, where:

$$V_n(0,0) = \sum_{j=0}^{n} (1 - u_n^{-j}) P(V_{0,n}^* = 0, V_n^*, n = j) \qquad (9)$$

The notation star such as in $V_{m,n}^*$ denotes for n fixed and $0 < m < n$. The probabilities are given by:

$$P(V_{0,n}^* = 0, V_n^*, n = j) = \sum_{k=j}^{l} \Lambda j, k, m (1 - q_n)^k q_n^{m-k}$$

$$\text{for} \qquad \Lambda j, k, m = \binom{m}{k-j} - \binom{m}{k-j-1} \qquad (10)$$

if $k > j$ and $\Lambda j, k, m = 1$ is $k = j$.

While both the modified tree model and the analytic formula give continuous valuation results, and it is well known the binomial tree model converges to the continuous price. the pricing formulae is still related to discrete lookback pricing. Cheuk and Vorst extended the total number of lookback periods, then selected an equal-distant subset to model the discrete monitoring periods. Similarly, the amnesiac lookback option is defined on a subset of monitoring periods.

Upon noting the discrepancy between continuous and discrete monitoring pricing assumptions, researchers continued to improve methods of pricing discrete options. Aitsahlia and Le Lai [34] uses the random walk duality to simplify recursive integration. Petrella and Kou [35] uses the Laplace transform to make pricing more efficient, and Broadie et al. [36] uses approximation adjustments to the continuous case to fairly price the option's discrete counterpart.

Contemporary researchers derive pricing under the framework of Lévy Processes, and as a natural extension, the Wiener-Hopf factorization is implemented as it gives the distribution of functionals within a random walk. Fusai et al. [37] combined this with Spitzer's identity for a general method of pricing discretely monitored exotic options, with an explicit formula for the fixed strike option. However, the method assumes equidistant monitoring windows. Boyarchenko and Levendorskii [38] similarly implement the Wiener-Hopf factorization but correct for the discrete monitoring dates using a Laplace transformation. Hieber [39] presents the general pricing of exotic options using the Fourier Transform and under Black-Scholes assumptions and regime switching model. This is not just useful for lookback options but for digital options and barrier options. However, the pricing scheme is for continuously monitored options. Feng and Linetsky [40] on the other hand, implements Hilbert transformations sequentially and presents an interesting method of computing maxima and minima, thereby applying it to the valuation of path-dependent options.

2.2. Method of Simulation

The definition of a amnesiac option in section 1 can be formalized as follows. Let N denote the total number of periods of a full lookback option. Let A denote a selected subset. Given a sequence of asset prices $\{S_i\}$, define M_0^T the maximum and m_0^T the minimum defined on the subsequence where $i \in A$. The payoff of the amnesiac option is then the same as **Table 1**.

To price the amnesiac lookback option, we utilize a Monte Carlo pricing simulation under the assumptions of the Black-Scholes world, including but not limited to, its modeling parameters, risk-free rate, dividends, and volatility. We assume geometric Brownian Motion under a risk-neutral measure as described in Equation (4):

$$S_T = S_0 e^{(r - D - \frac{\sigma^2}{2})T + \sigma \epsilon \sqrt{T}}$$

We assume dividends $D = 0$. We simulate this process using a Gaussian random walk. The algorithm is shown explicitly in Algorithm 1:

Algorithm 1: Algorithm for Pricing Amnesiac Option

p = number of selected days
V = vector with size equal to number of simulations
for *n* in *V* **do**
   A = generated list of indices of size p, using uniform distribution
   while *timestep* ≤ *total timesteps* **do**
      update asset price using Equation 2
      if *timestep is in A* **then**
         update maximum and minimum
      end
      update timestep
   end
   amnesiac call price = final value - minimum value
   amnesiac put price = maximum value - final value
   amnesiac spread price = maximum value - minimum value
end
Discount by e^{-rT}
Fair Price = Arithmetic mean of V for each of the three instruments
Standard Error (SE) = Standard Deviation / $\sqrt{\text{Number of Sims}}$
Confidence Interval = Fair Price ± (t)(SE)

---

where *t* is the corresponding *t* value for the confidence level. The updating of the asset price follows a Gaussian Random walk in Algorithm 2 where *dT* is the total time *T* divided by the number of selected days *p*:

Algorithm 2: The price evolution based on the Gaussian Random Walk.

S_i = current price
deterministic term: $\alpha = (r - d - \frac{\sigma^2}{2})dT$
stochastic term : $\beta = (\sigma)(\sqrt{dT})(\mathcal{N}(0,1))$
$S_{i+1} = S_i e^{\alpha + \beta}$

A denotes how the days are chosen. Our experiment features three forms of period selection: uniform random with fixed endpoints, completely random, and equal distant selection. For the amnesiac option with fixed endpoints, the first and last indices are first chosen, then indices $0 < i < N$ are chosen using the uniform distribution. For true random, indices are randomly selected under the uniform distribution with no guarantee of the first one being selected. For equal distant period selection, indices are chosen, and rounded down when not divisible. We make an assumption that on average, the population's selection of periods will be uniform random.

The parameters are shown in **Table 2**.

For each time step, the maximum and minimum value is updated. We analyze the pricing of floating and fixed amnesiac lookback options, along with the variants discussed in **Table 1**

TABLE 2 | Simulation parameters.

Starting stock price S_0	100
Volatility σ	0.2
Risk-free rate r	0.05
Total Periods N	100

(call, put, and spread). Then, the payoffs under equal interval sampling and random sampling are considered.

A useful function that appears in analysis is the general Hill Function. the Hill Function in the form:

$$V(X) = \frac{V_{Max} X^h}{K^h + X^h} \quad (11)$$

The Hill Function has its roots in chemistry, where it is commonly used to model the kinetics of substrate reaction rate, where the x-axis is the substrate concentration and the y-axis the rate of reaction. V_{Max} is the maximal saturation rate, K denotes the value of x required to attain half of V_{Max}, x is the density of a substance, and h is the hill coefficient, as an exponential of K and h.

Analogously, increasing lookback periods yield diminishing returns to payoff and the components of the Hill Function give intuition to the process itself. Firstly, V_{max} is theoretically the maximum payoff, or the price of the standard discrete lookback option. Secondly, as we increase the starting stock price, both V_{max} and C scale linearly. On the other hand, both K and h remain constant and depend on interest rate r and volatility σ. The Hill Curve can thus price amnesiac lookbacks independent of the initial asset price, keeping interest and volatility constant. If the price of the standard lookback option V_{max} is known, and interest rate and volatility incorporated into K and h, then pricing can be done very efficiently using the Hill Equation. Additionally, the curve can be adopted to any starting stock price, since the initial stock price only influences V_{max}. The ratio between the price evolution and initial price is simply $\frac{S_0 e^{(r-d-\frac{\sigma^2}{2})T + \sigma\sqrt{T}}}{S_0} = e^{(r-d-\frac{\sigma^2}{2})T + \sigma\sqrt{T}}$, hence when pricing under the same interest and volatility the payoff can simply be scaled.

The total number of simulations per parameter set is 2,000,000, the fair price being the arithmetic mean of the payoffs, with a target maximal standard error of 0.001.

3. RESULTS

We denote C_n and P_n as the amnesiac option with n chosen monitoring periods. When unspecified, the monitoring regime is assumed to be random period monitoring with fixed endpoints.
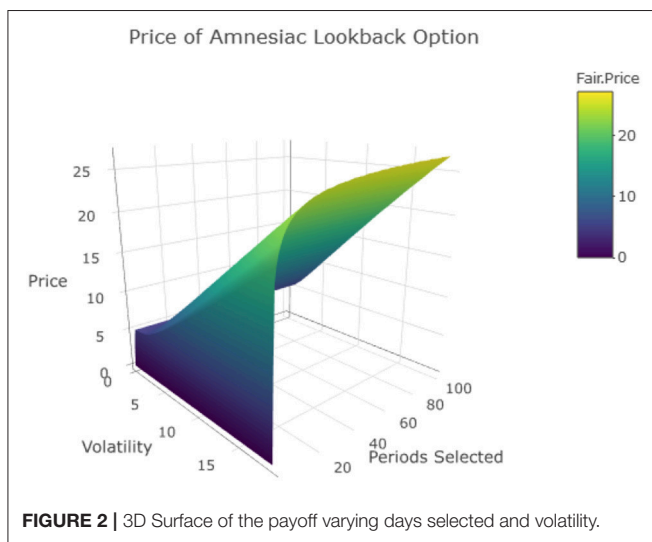
3.1. Floating Strike Amnesiac Option
To reiterate the definition of an amnesiac lookback option, its payoff scheme is as described in **Table 1**, where only the maxima and minima of a subset of periods are recorded. We use T and t to
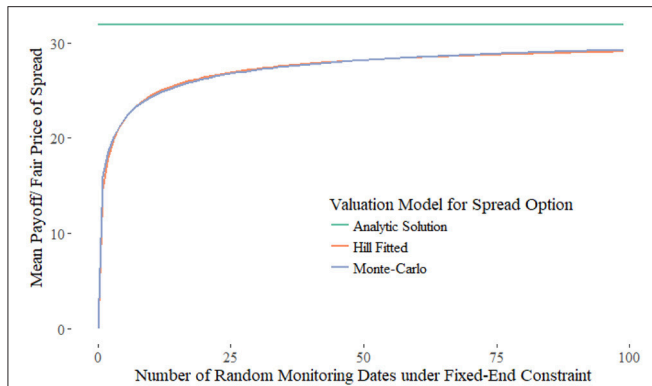
FIGURE 1 | Monte-Carlo simulation for the Spread Option results converge to the Hill-Function with $K = 1.18$, $h = 0.62$, and $V_{max} = 30.96$.

FIGURE 2 | 3D Surface of the payoff varying days selected and volatility.

denote the continuous case and N and n be the discrete case. Let N be the total number of periods and A the ordered set of indices:

$$A = \{n_a, n_b, , n_N\} \qquad \forall A \subset N$$

The amnesiac is defined on any k selected monitoring periods.

Figures 1, **2** show the payoff vs. the number of days selected. As the number of periods look backed on increases, the payoff approaches that of the standard lookback option. Yet it does not require many selected periods to capture most of the payoff of the full 100 selected days. This can be attributed to the diminishing marginal effect of additional selected periods; each period selected after the first contributes less additional payoff than the previous. For each data point, we took the mean of $2,000,000$ simulations, and given a calculated Standard Deviation of around 14.54, this gives an error of around 0.01 in absolute terms.

The sample errors are summarized in **Table 3**.

Figure 1 fits the Hill function as previously described, and **Figure 2** extends this to include volatility. For this particular case,

TABLE 3 | Average of sample errors from 2,000,000 simulations.

	Call	Put	Spread
Equal interval selection	0.01019	0.006877	0.007958
Uniform random with fixed endpoints	0.010181	0.006838	0.008000
Uniform random	0.009999	0.006756	0.007871

the formula is observed to be:

$$V(X) = \frac{30.96 X^{0.62}}{1.18^{0.62} + X^{0.62}}$$

Kernel density estimators further tell the differences of portfolios constructed from different amnesiac. **Figure 3** shows the distribution of the call payoffs. The peak clustered at zero denotes the likely-hood of the option not being exercised, which yields a payoff of zero. Comparisons between instruments vary case-by-case, so it is useful in comparing expected returns in a pairwise fashion.

Figure 4 shows the expected returns for the number of monitoring periods n equal to 99, 7, and 2. We define returns by the payoff minus the price of the option, which is the mean. Let x_i denote a single payoff from the simulation, then let the price of the call be defined as the average $\bar{X}$ then the returns are defined by:

$$R = x_i - \bar{X}$$

The payoff distribution of $n = 7$ for $R > 0$ is approximately the same as the distribution of $n = 99$. In fact, they converge in the negative region, hence their probability of $R > 0$ is approximately equal. Next, note that the downside of $n = 99$ is greater than that of $n = 7$. This means that the downside for C_7 is more limited than C_{99}, even if their conditional expectation for the downside is equal or even worse. This indicates that the amnesiac lookback, and partial lookbacks with less monitoring periods in general, can be used to limit the downside effects.

3.1.1. Three Types of Payoffs

Figure 5 shows the three payoff types as described in **Table 1** for floating strikes, denoted in the first column. The sum of the fair-price of the call and put is equal to the spread option, which is why it is often described as a "double lookback option."

Figure 5 also shows that the average maximal and minimal value is independent of the average final value. This observation can be illustrated as follows, given the payoffs in **Table 1** and **E** the expectation operator.

$$\begin{aligned} \mathbf{E}(call) + \mathbf{E}(put) &= \mathbf{E}(S_{max} - S_T) + \mathbf{E}(S_T - S_{min}) \\ &= \mathbf{E}(S_{max} - S_T + S_T - S_{min}) \quad (12) \\ &= \mathbf{E}(S_{max} - S_{min}) = \mathbf{E}(spread) \end{aligned}$$

These results show remarkable resemblance to the rate of convergence for exotic prices described in Dai et al. [41]. Indeed the prices produced by varying the number of total periods
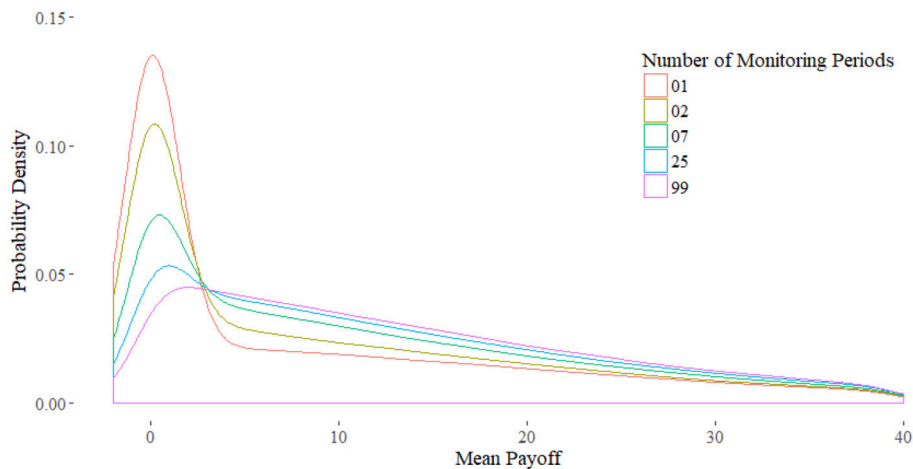
FIGURE 3 | Distribution of payoffs for the call with bandwidth = 1.
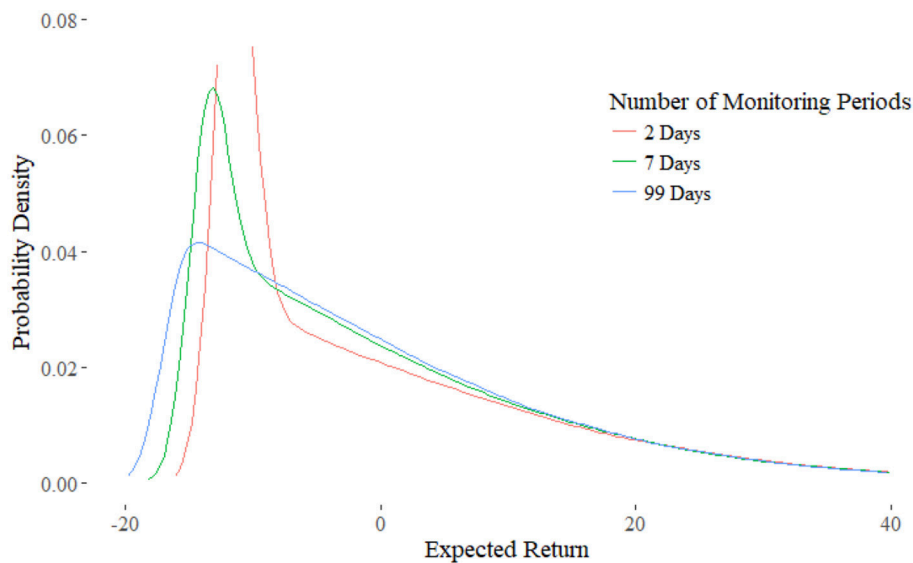
FIGURE 4 | Distribution of the Expected Returns. Right tails converge, while for greater *n* left tails shifts out.

N follows a similar curve, and produces an upper bound for the pricing. Furthermore, since the distribution of stocks is lognormal, if the interest rate $r = 0$, then the put will be of greater value than the call because the maximum will fluctuate absolutely more than the minimum. To see this suppose the simple case in the binomial model. Then for fixed n, $Se^{n\sqrt{\sigma}} - S$ is greater than $S - Se^{-n\sqrt{\sigma}}$. However, due to the positive risk-free rate the price of the call is greater as shown in **Figure 5**.

For illustration, consider the kernel densities of each of the three instruments shown in **Figure 6**. First, note the distribution of the spread option is lognormal, shown in the blue curve. In other words, the distribution of the range of a geometric Brownian motion is lognormal, due to the range of brownian motion being distributed normally.

Characteristic of the lognormal distribution is its fatter left tail. This corresponds to the shape of the put option, minus the peak created by the cluster of 0s. Additionally, the tail of the put is platykurtic in comparison to the tail of the call. In sum, once the 0 payoffs are eliminated, the distribution of the put corresponds to the left tail of the spread, and the distribution of the call corresponds to the inverted right tail.

3.2. Fixed Strike Options

The analysis of fixed strike amnesiac options first requires an analysis of how in-the-moneyness affects option payoff. In **Figure 7** the call price decreases as the strike price increases, as the more out-of-the-money the option gets, the more unlikely the maximum of the underlying will be greater than the strike. The
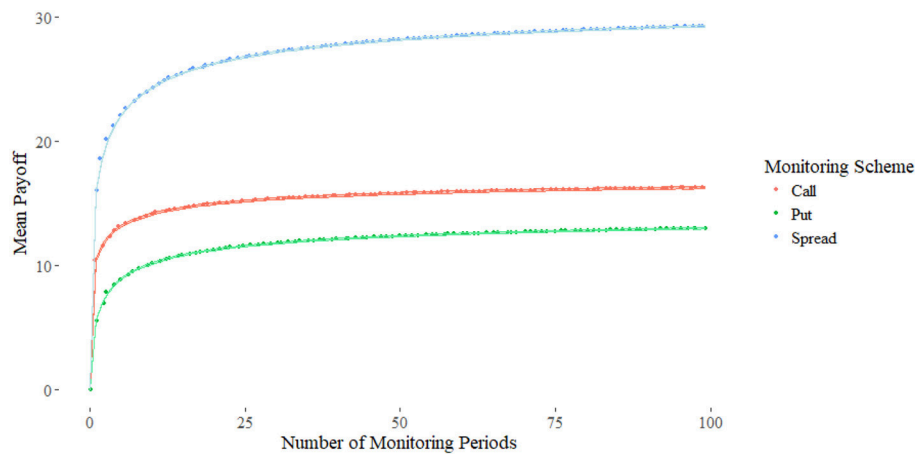
FIGURE 5 | The payoffs of the put (blue), call (orange), and spread (gray). The payoff of the put and call sum to the spread. The put payoff is lower than the call as the absolute difference of the maximal processes is greater than the minimal processes under Geometric Brownian Motion.
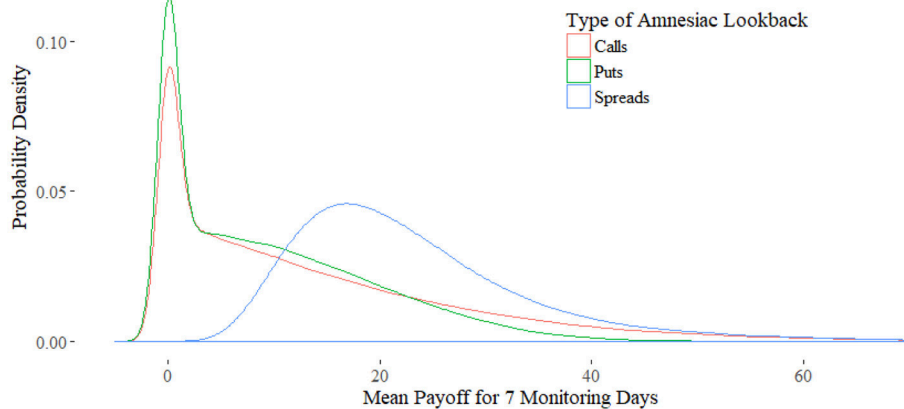
FIGURE 6 | Kernel Density Estimation of each Amnesiac type. The tail behavior of the spread option is captured by the distribution of the call and puts.

put option in **Figure 8** demonstrates the opposite relationship, as it relates to the minimum distribution. The payoff diagrams for the call show a linear relationship while in-the-money, then decaying exponentially afterwards. The put shows the inverse. Since the maximum (minimum) distribution for given number of periods is fixed, as the strike price decreases (increases) the difference is linear. This linear relationship is preserved as the number of selected periods in the amnesiac option are reduced, as seen in the diagrams.

When the number of monitoring periods is equal to 0, then the maximum is simply the last stock price at time N. Thus, the payoff is the final stock price minus the strike. The shape of this curve for selecting n periods is therefore a simple, vertical translation. Through these observations, we have demonstrated the bounds for the amnesiac option as shown in **Figure 9**. It is bounded by the standard lookback from above, and the vanilla European options from below.

3.3. Random vs. Equal Monitoring

One of the most important questions this paper addresses is if the choice of period matters in the payoff. What we find is randomly choosing lookback periods yields a worse payoff than equally spacing the days to look back on, as shown in **Figure 10**. The differences are shown in **Figures 11**, **12**.

There are two ways to monitoring randomly. We defined N equidistant points between time 0 and T, making it discrete. Suppose we are monitoring k points from the Gaussian Random walk. Under true random monitoring we would select k random points from the set of periods $\{0, 1, 2, ..., N-1\}$. In contrast, the fixed-end random monitoring selects the first period $n = 0$ then randomly chooses $k - 1$ periods from indices 1 through $N - 1$. Since the price of an option increases with the duration of the contract, fixing the first and last period ensures an equal time length for comparison. For the true random monitoring, the time to maturity would be the minimum of index.
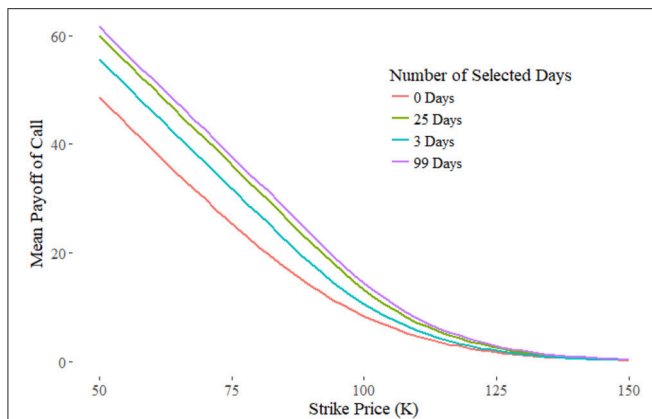
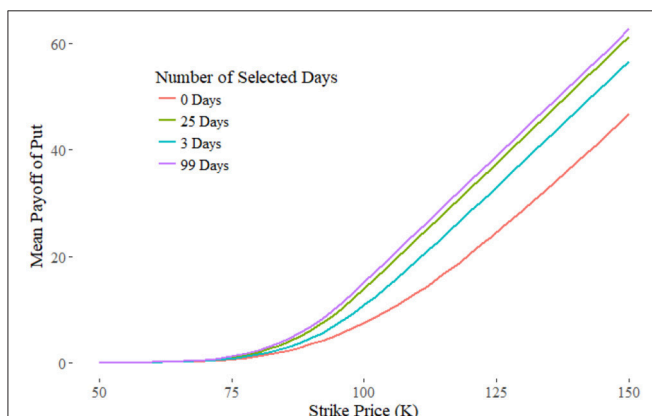**FIGURE 7 |** Amnesiac Call Payoff with fixed strike.

FIGURE 8 | Amnesiac Put Payoff with fixed strike.

For illustration, suppose $k = 2$. For fixed-end random sampling the first period is chosen while the second can be any point between 1 and $N - 1$. More simply, both endpoints are fixed, and we are choosing a point between the beginning and end. Numerical results in **Figure 13** on average selection of the midpoint as more optimal. True random sampling produces weak payoffs due to shortened maturities. However, even when endpoints are always chosen to lookback on, the payoff for equidistant sampling is superior. Although this suggests the optimal strategy is to equally space the selected periods, this does not take into account sudden volatility swings from extreme, singular events.

The intuition for this toy experiment can be explained as follows. Since both endpoints are fixed, we are only interested in choosing the point in-between. If the point chosen is less than or equal to S_0 or S_N, then the payoff is the same as the case of choosing one point, which is S_0. For the third point S_i to be of value it must satisfy $S_i > S_0$ and $S_i > S_N$. Suppose n is very close to the first index 0, then $S_n = S_0 e^\epsilon$ for ϵ that falls in the upper half of the Gaussian distribution defined in Equation (2). The difference between S_0 and S_n is thus small, and the payoff will be close to that of the case of just choosing the first point. On the

other hand, suppose n is close to N. Then $S_N = S_n e^\epsilon$. Since the payoff is just the difference between the max and final value, the payoff is diminished.

Simply put, given that the chosen point S_n is greater than S_0 and S_N, then choosing too close to the beginning means not giving enough distance for the maximum to rise, and choosing too close to the end means diminished payoff due to proximity with S_N. Intuitively, time value is then the explanatory factor that determines how much marginal payoff a selected period gives, where the value of a point is proportional to the time value associated with the distance between it and the closest adjacent lookback period.

We now demonstrate the above claim mathematically. For simplicity we fix the first period and denote $j = k - 1$ as the number of lookback periods. Then let N denote the total number of equally spaced periods of the standard lookback with total time T, and let A be the set of selected periods. Define function γ that gives the minimum distance

$$\gamma(i) = \min\{d(i, k) \quad \forall k \in A, k \neq i\} \quad (13)$$

Then the totals of time value is given as
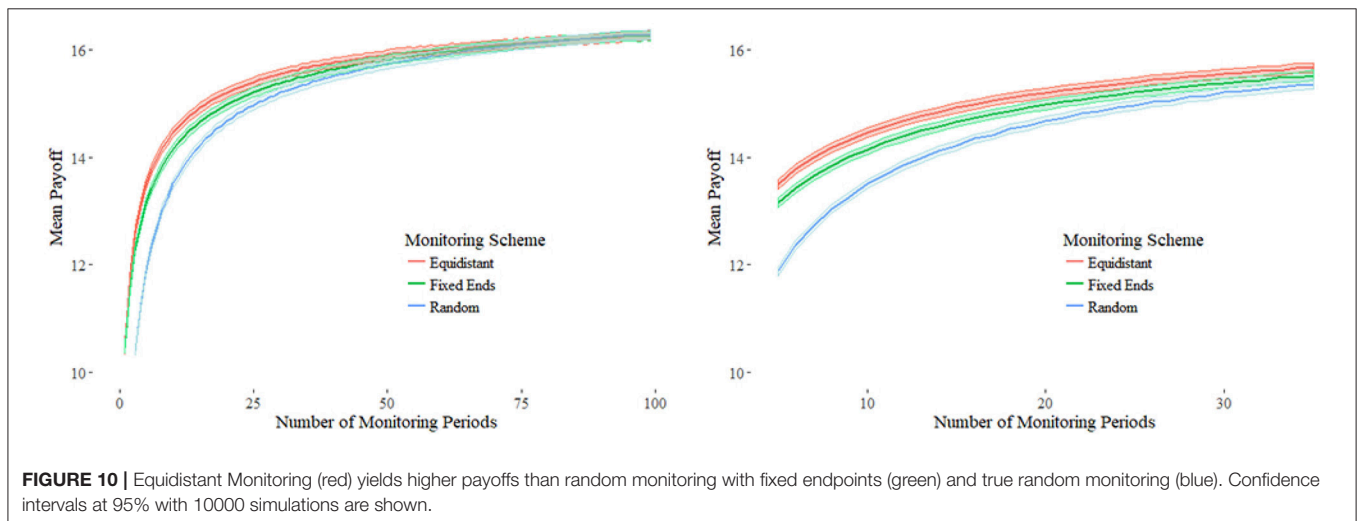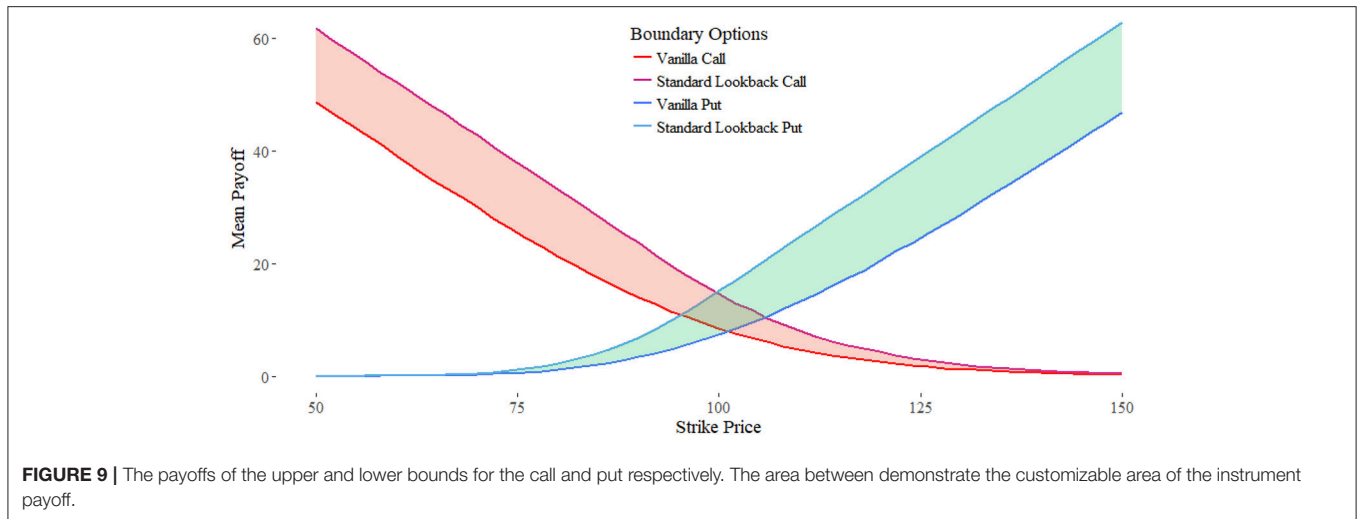
$$\Gamma(A) = \sum_{i \in A} \gamma(i) \quad (14)$$

where the value of Γ for equidistant sampling is $\frac{j}{j+1}$. We show this value is maximal.

Consider the base case of $n = 1$. If $i_0 = \frac{N}{2}$ the only element in A_0, then $\Gamma(A_0) = \frac{N}{2}$. Compare to this to A_1 whose only element $i_1 = \frac{N}{2} + \epsilon$ for $\epsilon \in \mathbb{Z}^+, \epsilon < \frac{N}{2}$. Since $d(0, i_1) > \frac{N}{2}$, then $d(i_1, N) = \frac{N}{2} - \epsilon < \frac{N}{2}$. Thus $\Gamma(A_1) = \frac{N}{2} - \epsilon < \Gamma(A_0)$. By symmetry, any i_1 greater or less than $\frac{N}{2}$ is worth less. Cases for larger j can be shown in a similar fashion by assuming the first point greater than $\frac{N}{j}$. Thus our claim that lookback period value is proportional to the time value of the next closest period remains true.

What this conclusion implies in real situation trading is that it would be inefficient to cluster lookback points around a presumed volatile event in order to capture a price extrema, even if volatility spikes seem certain, because the marginal value of each period chosen beyond the first in the locality diminishes rapidly due to the decreasing proximity of neighboring points. Prices are less likely to change a significant amount between two points in time the closer the two points in time are.

From the framework of analysis, suppose we sub-divide Geometric Brownian Motion into N subintervals defined by select indices. Then we are interested in the endpoints $[i, i + 1]$ as our monitoring dates. If these subintervals are of equal length, then the probability of the true maximum being captured is equal, since within the subinterval, the maximum is equally likely to fall on any point. Thus, the probability of the maximum being captured by the neighborhood around the endpoints diminish as subintervals get larger. The maximal neighborhoods are when endpoints are equally spaced.

This has been explored in the risk theory and queuing theory literature, which assumes expected maxima calculations on

FIGURE 9 | The payoffs of the upper and lower bounds for the call and put respectively. The area between demonstrate the customizable area of the instrument payoff.

FIGURE 10 | Equidistant Monitoring (red) yields higher payoffs than random monitoring with fixed endpoints (green) and true random monitoring (blue). Confidence intervals at 95% with 10000 simulations are shown.

equidistant sampling. For further references consider papers by Janssen and Van Leeuwaarden [42, 43] on equidistant sampling of Brownian Motion and Gaussian Random Walks, and Alfi's work on exacting roughness in finite random walks [44]. Under equidistant sampling, the expected maximum shown using the Spitzer Identity which gives the joint distribution of partial sums and their maximal sums for a collection of random variables:

$$\mathbb{E} \max_{n=0,...,N} B\left(\frac{n}{N}\right) = \sum_{n=1}^{N} \frac{1}{n} \mathbb{E} \max\{\frac{n}{N}\} \qquad (15)$$

Now let the process X_n be defined as

$$X_n = \sum_{i=1}^{n} \delta_{xi} \qquad (16)$$

After computation [44], the expected maximum value can be expressed as the following

$$\mathbb{E}(M_k) = \lim_{s \to 1} \frac{\mathbb{E}(s^{M_k}) - 1}{\ln(s)} = \sum_{i=1}^{k} \frac{\mathbb{E}(|X_i|)}{2i} \qquad (17)$$

Adapting these results to Gaussian random walks with unequal steps may produce approximations for the payoff of lookbacks and other exotic options.

If the events of the world are analogous to random occurrences, then nonrandom selection of lookback periods is akin to event-based trading strategies. Specifically in regards to the amnesiac lookback option, such a strategy leads to lower payoffs. This conclusion implies that given general market efficiency, event-based period selection strategies of lookback periods should be unprofitable.

FIGURE 11 | Put option difference between equidistant monitoring and random monitoring payoffs.

FIGURE 13 | Payoff for $k = 2$ with floating strike. The payoffs of always choosing the midpoint is greater than random sampling.

as a ratio of the initial price in **Figure 14**. Data is taken from Coinmarketcap.

The last row in **Table 5** shows a Bitcoin amnesiac option priced at the implied volatility of Bitcoin calls in December. Interestingly, it is less than the minimum volatility of Bitcoin itself. When the implied volatility is less than historic volatility, this indicates that market sentiment benefits option buyers. In other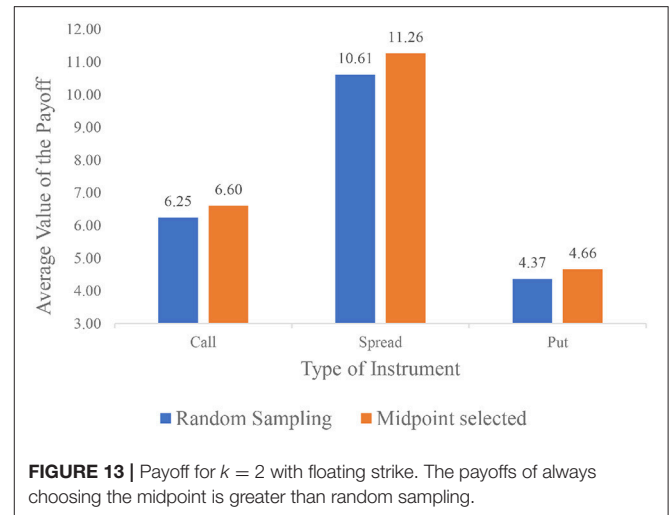 words, the market expects lower volatility in the future, and as a result, using historical volatility as a proxy is prone to overpricing.

To verify this is what we expect, we observe a linear correlation between the price of the amnesiac lookback and its volatility, yielding a correlation coefficient is $r^2 = 0.99$. An interesting point is the difference in shape between the call and the put curves. As seen in **Figure 14**, the put curves slower than the call, due to the right vs. left tail of the lognormal distribution, especially when volatility is high in the case of ripple. This means that the amnesiac put may be useful if a trader anticipates a down market without needing to worry when to sell, while enjoying a greater range of values to choose from due to the slower convergence to the standard lookback option.

One can conceive of using Algorithm 1 to price an option, particularly for Ethereum which was developed for Smart Contracts via the Ethereum Virtual Machine (EVM). The scripting language it uses is very expressive and Turing complete, and the EVM acts as a distributed computer that executes commands based on the resource *gas*. Smart contracts thus become a good way to buy contracts that previously only existed in the OTC market. In this case, a transaction function transfers Ether to the writers account at the time of purchase. At the time of maturity, if the option is exercised, the smart contract executes the necessary transactions as determined by the option automatically. The buyer sends in a selection of monitoring periods A, then the input parameter could be the number of periods which is used to generate a subset of pseudorandom monitoring periods, or direct simulation of the selected periods. The latter produces a more accurate price but is also computationally more expensive and as a result, raises

FIGURE 12 | Spread option difference between equidistant monitoring and random monitoring payoffs.

4. DISCUSSION

These results can be immediately applied to pricing cryptocurrency. Under the Black-Scholes model, the implied volatility is usually calculated using the current market price of the option. However, because this is a new option being proposed, and since both futures and vanilla options are still new derivatives on the market, it is not possible at this time to even price them with rudimentary Monte-Carlo simulations.

Instead, we use the historical volatility as a proxy and use the 3-month treasury bond as the risk-free rate. **Tables 4**, **5** show the prices of amnesiac calls and puts based on historical data in a 3 and 2 month window, respectively. These are then shown

Cryptocurrency	S_0	vol	C_1	C_7	C_{92}	P_1	P_7	P_{92}
Bitcoin	10234	0.476	1980.4	2668.5	3193.8	1851.1	2995.1	4010.5
Ethereum	447	0.543	98.25	131.27	155.42	92.18	150.53	203.79
Iota	1.33	0.971	0.499	0.6309	0.7329	0.4851	0.8569	1.215
Litecoin	88.3	0.719	24.931	32.593	37.992	23.778	39.979	54.991
Monero	180.14	0.66	47.359	62.356	73.601	45.327	75.637	103.417
Neo	33	0.976	12.542	15.82	18.097	12.062	21.382	30.372
Ripple	0.251	0.529	0.0534	0.0716	0.0857	0.0505	0.0822	0.1108

Cryptocurrency	S_0	vol	C_1	C_7	C_{62}	P_1	P_7	P_{62}
Bitcoin	9170.54	0.65	2310.4	2971.4	3427.5	2262.9	3534.6	4606.8
Ethereum	1036.79	0.65	258.31	333.17	387.42	255.4	399.21	520.3
Iota	1.91	1.17	0.7452	0.9186	1.0435	0.7607	1.2645	1.7152
Litecoin	142.3	0.89	48.918	61.125	69.393	49.18	79.927	106.707
Monero	240.97	0.733	73.741	93.272	106.413	72.831	116.508	154.464
Neo	126	0.96	42.336	53.004	60.053	42.099	68.197	90.952
Ripple	0.963	1.24	0.4667	0.5577	0.6197	0.4826	0.8426	1.1787
Bitcoin (IV)	9170.54	0.65	1730.6	2272.2	2667.3	1657	2528.9	3244

the transaction fees of the smart contract. The choice becomes a question of cost vs. profit, and once statistical distributions on people's preferences of monitoring periods are collected, the pricing can be made even more accurate at lower computational costs.

5. CONCLUSION

Lookback options are immensely useful instruments for the use of hedging against risks associated with high volatility and notably effective at canceling investor regret as well. Although the literature has explored lookback options and many concepts of the partial lookback, previous work has mainly involved lookbacks with continuous monitoring, and equidistant lookback period selection. We suggest the use and exploration of a new form of partial lookback, the amnesiac lookback option, that can look back upon any discretely selected period prior to expiration.
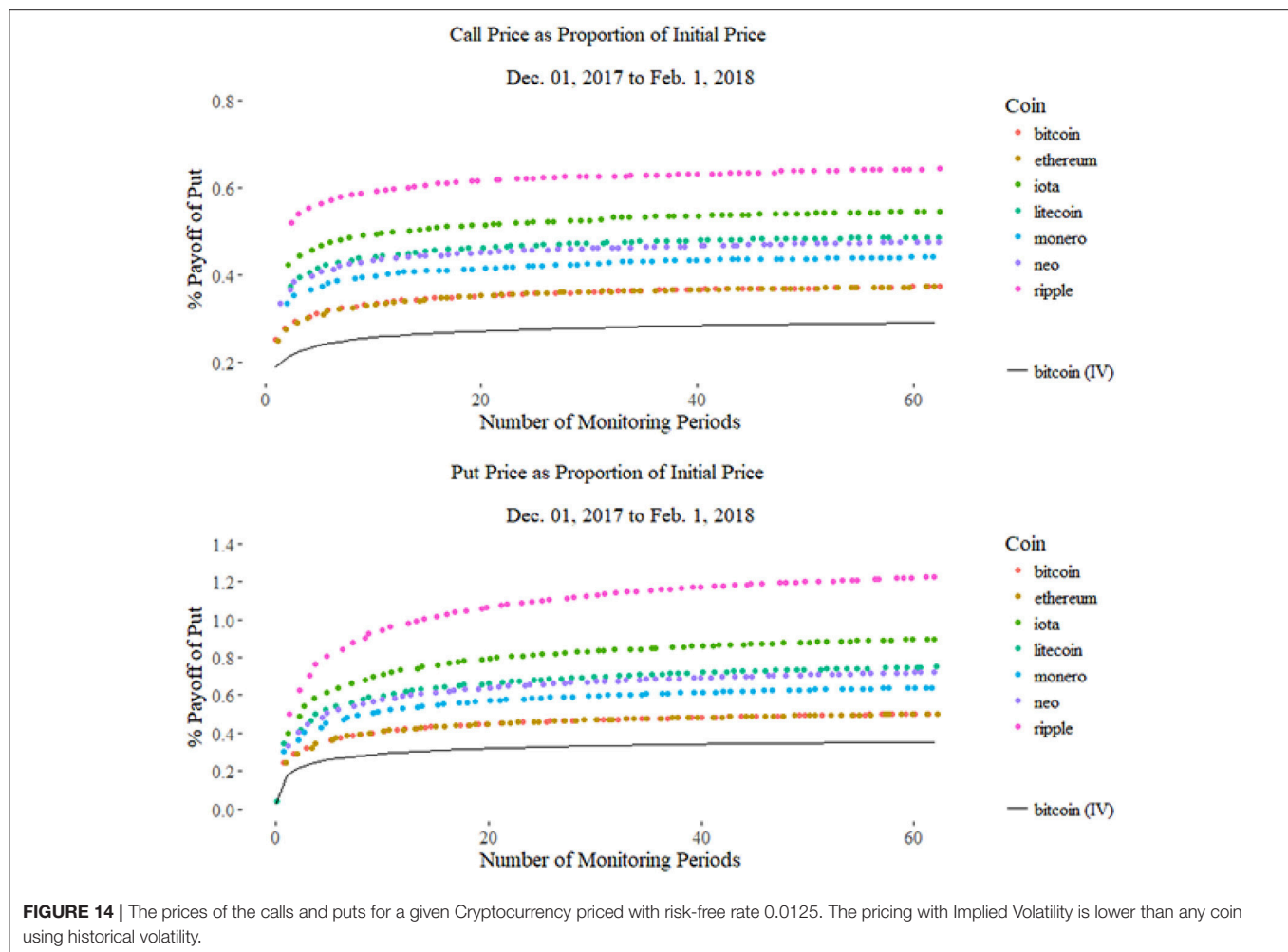
We discuss the practical applications that amnesiac lookback options may have in the trading of cryptocurrencies such as Bitcoin and especially Ethereum. The cryptocurrency market has been defined by its high volatility trading and plethora of exogenous shocks that regularly disrupt the market. In such a high risk and speculative market, it is imperative that investors have some way of mitigating their risks and exposures. In this realm, lookback options are a simple and effective way to do so, while amnesiac lookback options allow an investor to fully customize the amount of risk he is willing to take on or forego by adjusting the number of lookback periods, and potentially where to place them, as shown in **Figure 1**. This is particularly true for put options under high volatility. Past studies have also found long memory in Bitcoin returns and return

volatility, which path-dependent options such as the amnesiac lookback are able to correct by closing arbitrage opportunities, if traded on rationally. This opportunity is greatly enhanced by Smart Contracts, which decentralizes and automates highly customizable derivatives like path-dependent options.

Using Monte Carlo simulations under Black-Scholes assumptions, we find the Hill Function to be a suitable model for pricing the payoffs of the amnesiac lookback option. We also determine the bounds of the amnesiac lookback consistent with prior results on partial lookbacks, that is the standard lookback from above, and the corresponding vanilla European option (call or put) from below.

Our work additionally discusses the merits of equal spacing vs. random sampling of lookback periods. In this regard, we find that equally spaced lookback periods yield the greatest payoffs under Black-Scholes assumptions. Since the formula for modified trees and lattice methods produce analytic results for equidistant sampling, and equidistant sampling is a strict upper bound for random sampling, we can conclude that if days are not chosen equally then the instrument will be prone to being overpriced. More importantly, our results are more realistic and related to the designated function of these options in actual markets than some previous works.

The fact that how periods are selected affect the pricing suggest multiple paths of further inquiry. Since this is phenomenon under Black-Scholes assumptions, different models such as introducing stochastic volatility and Lévy processes may produce different results. Tying this to real commodities, statistical analysis of Cryptocurrency data can benchmark the efficiency of such an option. Simulating extreme, volatile events within the context of the Gaussian random walk would bridge the

FIGURE 14 | The prices of the calls and puts for a given Cryptocurrency priced with risk-free rate 0.0125. The pricing with Implied Volatility is lower than any coin using historical volatility.

empirical and axiomatic formulations, and is ultimately the purpose of having a choice in monitoring periods. Understanding period selection strategies will yield new perspectives of using path-dependent options for hedging risk, and expand trading strategies with the growth of novel and volatile blockchain commodities.

AUTHOR CONTRIBUTIONS

H-CC wrote the code for the numerical simulation, data processing, and the mathematical finance literature review. KL contributed economic analyses, interpretation, literature review and extension to cryptocurrencies. Both contributed to the conception and design of the option's amnesiac feature, interpretation of results, and general authorship the paper.

REFERENCES

1. Shreve SE. *Stochastic Calculus for Finance I: The Binomial Asset Model*. Vol. 11. New York, NY: Springer Science & Business Media (2004).
2. Goldman MB, Sosin HB, Gatto MA. Path dependent options:"Buy at the low, sell at the high". *J Finan*. (1979) **34**:1111–27.

3. Goldman MB, Sosin HB, Shepp LA. On contingent claims that insure ex-post optimal stock market timing. *J Finan*. (1979) **34**:401–13. doi: 10.1111/j.1540-6261.1979.tb02102.x
4. Schwert GW. Anomalies and market efficiency. *Handb Econ Finan*. (2003) **1**:939–74. doi: 10.1016/S1574-0102(03)01024-0
5. Kat HM, Heynen RC. Selective memory. *Risk* (1994) **7**:73–6.

6. Conze A, Viswanathan. Path dependent options: the case of lookback options. *J Finan.* (1991) **46**:1893–907. doi: 10.1111/j.1540-6261.1991. tb04648.x

7. Buchen P, Konstandatos O. A new method of pricing lookback options. *Math Finan.* (2005) **15**:245–59. doi: 10.1111/j.0960-1627.2005.00219.x

8. Carrick J. Bitcoin as a complement to emerging market currencies. *Emerg Markets Finan Trade* (2016) **52**:2321–34. doi: 10.1080/1540496X.2016.1193002

9. The Guardian. *Don't Dismiss Bankers' Predictions of a Bitcoin Bubble – They Should Know.* Guardian News and Media (2017). Available online at: https://www.theguardian.com/business/2017/sep/17/jamie-dimon-bitcoin-bubble-he-would-know-banking

10. Buy Bitcoin Worldwide. *The Bitcoin Volatility Index.* Available online at: https://www.buybitcoinworldwide.com/volatility-index/

11. Bariviera AF, Basgall MJ, Hasperué W, Naiouf M. Some stylized facts of the Bitcoin market. *Phys A* (2017) **484**:82–90. doi: 10.1016/j.physa.2017.04.159

12. Urquhart A. Price clustering in Bitcoin. *Econ Lett.* (2017) **159**:145–8. doi: 10.1016/j.econlet.2017.07.035

13. Times TE. *Bitcoins Lose Lustre in the Face of Flak* (2017). Available online at: http://economictimes.indiatimes.com/wealth/personal-finance-news/bitcoins-lose-lustre-in-the-face-of-flak/articleshow/60709423.cms

14. Bouri E, Gupta R, Tiwari AK, Roubaud D. Does Bitcoin hedge global uncertainty? Evidence from wavelet-based quantile-in-quantile regressions. *Finan Res Lett.* (2017) **23**:87–95. doi: 10.1016/j.frl.2017. 02.009

15. Bouri E, Molnár P, Azzi G, Roubaud D, Hagfors LI. On the hedge and safe haven properties of Bitcoin: is it really more than a diversifier? *Finan Res Lett.* (2017) **20**:192–8. doi: 10.1016/j.frl.2016.09.025

16. Russo C. *Bitcoin Options Will Be Available This Fall.* Bloomberg (2017). Available online at: https://www.bloomberg.com/news/articles/2017-07-24/bitcoin-options-to-become-available-in-fall-after-cftc-approval

17. U S Commodity Futures Trading Commision. *CFTC Statement on Self-Certification of Bitcoin Products by CME, CFE and Cantor Exchange.* CFTC (2017). Available online at: https://www.cftc.gov/PressRoom/PressReleases/pr7654-17

18. Roberts D. *Why Nasdaq, CME, CBOE All Want in on Bitcoin Futures.* Yahoo! (2017). Available online at: https://finance.yahoo.com/news/nasdaq-cme-cboe-want-bitcoin-futures-183256191.html

19. Szabo N. Smart contracts: building blocks for digital markets. In: *EXTROPY: The Journal of Transhumanist Thought.* Marina Del Rey, CA (1996).

20. Szabo N. The idea of smart contracts. In: *Nick Szabo's Papers and Concise Tutorials* (1997). Available online at: http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/idea.html

21. Buterin V. A next-generation smart contract and decentralized application platform. In: *White Paper* (2014). Available online at: https://www.weusecoins.com/assets/pdf/library/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf

22. Wright A, De Filippi P. *Decentralized Blockchain Technology and the Rise of Lex Cryptographia* (March 10, 2015). doi: 10.2139/ssrn. 2580664

23. Omohundro S. Cryptocurrencies, smart contracts, and artificial intelligence. *AI Matt.* (2014) **1**:19–21. doi: 10.1145/2685328.2685334

24. De N. *Smart Contracts Now Recognized Under Tennessee Law* (2018). Available online at: https://www.coindesk.com/blockchain-bill-becomes-law-tennessee/

25. Peters G, Panayi E, Chapelle A. Trends in cryptocurrencies and blockchain technologies: a monetary theory and regulation perspective. *J. Financ. Perspect.* (2015) **3**. Available online at: https://ssrn.com/abstract=3084011

26. Musiela M, Rutkowski M. *Martingale Methods in Financial Modelling, volume 36 of Applications of Mathematics: Stochastic Modelling and Applied Probability.* New York, NY: Springer (1997).

27. Kat HM, Heynen RC. Lookback options with discrete and partial monitoring of the underlying price. *Appl Math Finan.* (1995) **2**:273–84. Available online at: http://www.tandfonline.com/doi/abs/10.1080/13504869500000014

28. Babbs S. Binomial valuation of lookback options. *J Econ Dyn Cont.* (2000) **24**:1499–525. doi: 10.1016/S0165-1889(99)00085-8

29. Cheuk TH, Vorst TC. Currency lookback options and observation frequency: a binomial approach. *J Int Money Finan.* (1997) **16**:173–87. doi: 10.1016/S0261-5606(96)00052-6

30. Leung KS. An analytic pricing formula for lookback options under stochastic volatility. *Appl Math Lett.* (2013) **26**:145–9. doi: 10.1016/j.aml.2012.07.008

31. Dia EHA, Lamberton D. Connecting discrete and continuous lookback or hindsight options in exponential Lévy models. *Adv Appl Probabil.* (2011) **43**:1136–65. doi: 10.1239/aap/1324045702

32. Cox JC, Ross SA, Rubinstein M. Option pricing: a simplified approach. *J Finan Econ.* (1979) **7**:229–63. doi: 10.1016/0304-405X(79)90015-1

33. Hull JC, White AD. Efficient procedures for valuing European and American path-dependent options. *J Derivat.* (1993) **1**:21–31. doi: 10.3905/jod.1993.407869

34. Aitsahlia F, Le Lai T. Random walk duality and the valuation of discrete lookback options. *Appl Math Finan.* (1998) **5**:227–40. doi: 10.1080/135048698334655

35. Petrella G, Kou S. Numerical pricing of discrete barrier and lookback options via Laplace transforms. *J Comput Finan.* (2004) **8**:1–38. doi: 10.21314/JCF.2004.114

36. Broadie M, Glasserman P, Kou SG. Connecting discrete and continuous path-dependent options. *Finan Stochast.* (1999) **3**:55–82. doi: 10.1007/s007800050052

37. Fusai G, Germano G, Marazzina D. Spitzer identity, Wiener-Hopf factorization and pricing of discretely monitored exotic options. *Eur J Operat Res.* (2016) **251**:124–34. doi: 10.1016/j.ejor.2015.11.027

38. Boyarchenko S, Levendorski S. Efficient Laplace inversion, Wiener-Hopf factorization and pricing lookbacks. *Int J Theor Appl Finan.* (2013) **16**:1350011. doi: 10.1142/S0219024913500118

39. Hieber P. Pricing exotic options in a regime switching economy: a Fourier transform method. *Rev Derivat Res.* (2017) 1–22. doi: 10.1007/s11147-017-9139-1

40. Feng L, Linetsky V. Computing exponential moments of the discrete maximum of a Lévy process and lookback options. *Finan Stochast.* (2009) **13**:501–29. doi: 10.1007/s00780-009-0096-x

41. Dai TS, Liu LM, Lyuu YD. Linear-time option pricing algorithms by combinatorics. *Comput Math Appl.* (2008) **55**:2142–57. doi: 10.1016/j.camwa.2007.08.046

42. Janssen AJEM, Van Leeuwaarden J. Cumulants of the maximum of the Gaussian random walk. *Stochast Process Their Appl.* (2007) **117**:1928–59. doi: 10.1016/j.spa.2007.03.006

43. Janssen AJEM, Van Leeuwaarden JSH. On Lerch's transcendent and the Gaussian random walk. *Ann Appl Probabil.* (2007) **17**:421–39. doi: 10.1214/105051606000000781

44. Alfi V, Coccetti F, Marotta M, Petri A, Pietronero L. Exact results for the roughness of a finite size random walk. *Phys A* (2006) **370**:127–31. doi: 10.1016/j.physa.2006.04.020

On the Aerodynamic Forces on a Baseball, With Applications

Gerardo J. Escalera Santos[1]*, Mario A. Aguirre-López[2], Orlando Díaz-Hernández[1], Filiberto Hueyotl-Zahuantitla[1,3], Javier Morales-Castillo[4] and F-Javier Almaguer[2]

[1] Facultad de Ciencias en Física y Matemáticas, Universidad Autónoma de Chiapas, Tuxtla Gutiérrez, Mexico, [2] Facultad de Ciencias Físico-Matemáticas, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Mexico, [3] Cátedra CONACyT, Mexico City, Mexico, [4] Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Mexico

The aerodynamic forces acting on a baseball are those produced by the contact between the ball and the air, and are defined by the initial conditions of the pitch. It is well known that such forces determine the changes from the typical parabolic ballistic trajectory, either in the direction of the movement of the ball (drag force), or producing a lift or lateral deflection (Magnus and seam forces). The drag and Magnus effects have been widely studied and there are many references about their nature and the trajectory they produce, which is predictable. This has led to most baseball research being related with spinning pitches. On the other hand, there is the unpredictable motion of a knuckleball, whose erratic trajectory accompanied by a poor understanding of the forces produced by the asymmetry of the seams had markedly limited research about it until the beginning of this century. However, nowadays interest in the knuckleball is resurfacing. Data collected by wind tunnel experiments and real pitches have motivated researchers to analyze the phenomenon and build models that try to predict the motion of the ball. In this work we aim to provide the reader some basic ideas on aerodynamic forces through a combination of experimental results, phenomenological and dimensional analysis, with special focus on new advances on the seam effects of a knuckleball pitch. In addition, we discuss possible ways to extend the existing models about the seam forces. Finally, we summarize from the literature some methods regarding the reproduction and reconstruction of baseball trajectories from aerodynamic forces and discuss their application as well.

Keywords: baseball, knuckleball, seams, drag force, lift force, Magnus force, ball games

1. INTRODUCTION

Aerodynamics and the flight of baseballs are very interesting phenomena from the point of view of sports, engineering and science. The different types of pitches are denoted by the different initial configurations that the pitcher gives to the ball by means of his hand. Each initial combination of velocity and spin produces a way of interaction between the air and the ball which results in curveballs, fastballs, sliders, and knuckleballs, among others. Moreover, such interaction can be affected by other factors, such as when the ball is dented or when it has liquid on its surface. This can cause an erratic movement in the expected trajectory of the ball, and be the difference between a strike or a home run.

For engineering, the evolution of a baseball during its flight means an open door to many possible applications. In turn, for science it means the study of the process in which a subsonic flow interacts with a solid rough sphere, having the characteristic that for zero or lower spin values of the spin the trajectory is very erratic and becomes unpredictable, whereas for higher values the trajectories are smooth and predictable. In this way, the trajectories of a baseball are commonly classified by spinning pitches and non-spinning pitches.

The predictability of both types of trajectories seems contradictory at first glance, since one would expect the spinning balls to have a more complex movement because more forces act on them. This is not the case, however, and the main cause of this is the role of the seams of the baseball. When the interaction between the seams and the air is insignificant, more predictable and stable trajectories occur, and vice-versa. This last statement can be better understood by looking at the Reynolds' transport theorem, which establishes [1–3]

$$\frac{\partial}{\partial t} \int_{\mathcal{V}} \rho \mathbf{V} d\mathcal{V} + \int_{\mathcal{S}} \rho \mathbf{V} \left(\mathbf{V} \cdot \mathbf{n} \right) dA = \int_{\mathcal{S}} \mathbf{T} \cdot \mathbf{n} d\mathcal{S} + \int_{\mathcal{V}} \rho \mathbf{b} d\mathcal{V} \quad (1)$$

for a fluid-containing volume of space with volume $\mathcal{V}$ and surface $\mathcal{S}$, where $\mathbf{V}$ and ρ denote the velocity and density of the fluid, respectively, and $\mathbf{n}$ is a unit normal vector pointing outside $\mathcal{S}$. In turn, the right-hand side denotes the sum of all forces acting on the volume of space, which can be classified according to the intensive properties as body (first term) and surface (second term) forces, with $\mathbf{b}$ and $\mathbf{T}$ being the body forces vector per unit mass and the stress tensor, respectively. Among body force gravity, Coriolis and centrifugal forces can be found, whereas the forces generated at the ball-air boundary include pressure, normal and shear stress, etc.

In spite of gravity, body forces are weak and are commonly neglected in calculations of ball sports as reported in calculations by Robinson and Robinson [4] and Aguirre-López et al. [5]. On the other hand, the surface forces can be classified according to the direction and nature of their origin: drag, Magnus, lift, side and other forces [6–8]. Magnus force can be defined as that caused by the spin of the ball, therefore it is present only in spinning pitches. In turn, we define lift and side forces as those caused only by the motion of the baseball through the air without rotation (non-spinning pitches); then it depends on the orientation of the ball because of the asymmetry of the seams. A proposed curve of how lift and Magnus forces behave when varying the magnitude of the spin is shown in **Figure 1**. The lift and side forces produced by the seams decrease with the increase of the angular velocity (which is expected according to Watts and Sawyer [11], Mehta [12], and Cross [9]) while by definition, the Magnus force approaches zero when the angular velocity vanishes [8]. In this way, the non-spinning window is related to erratic and unstable trajectories whereas the spinning window is related to predictable and stable trajectories. Finally, there is a window between both cases, in which the superposition of forces is more significant and the ball can continue with the erratic movements of non-spinning pitches or can draw a smooth, slow
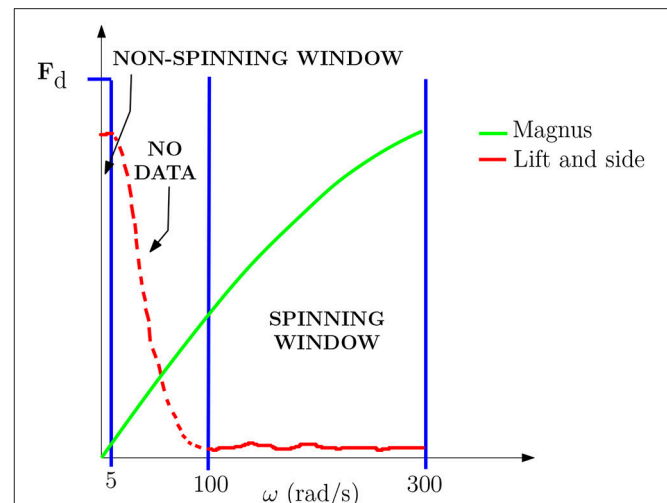
FIGURE 1 | Scheme of the maximum value of lift, side and Magnus forces when varying the angular velocity ω for a given speed V. The y−axis is scaled according to the maximum drag force for the considered velocity. The curves for lift, side and Magnus forces are drawn according to experimental data [8–10]. The missing curve for lift and side forces between spinning and non-spinning windows is because there are no data reported in such range.

and predictable trajectory. From our understanding, this is the main reason why pitchers do not throw balls within this range.

The curves in **Figure 1** also show the evolution of the study of the types of pitches at this time. Until the beginning of this century, spinning pitches covered most baseball research; therefore, books and compilations of works related to the subject avoided the phenomena present in non-spinning pitches [7]. Nowadays, the interest in non-spinning pitches is resurfacing and there is more information available in books of the present decade, like Cross [6]. However, the new developments and methods for studying non-spinning and intermediate windows are very disconnected, so a new compilation of work on the matter is necessary.

The purpose of this work is to introduce the reader to the diverse existing methodologies for studying the aerodynamic forces in the flight of a baseball, and to discuss the possible ways to complement such studies for future research and applications. All this is based on a literature review of the subject. Each topic begins with a brief derivation of the mathematical model of the respective force; then the model is compared with the results obtained from experimental measures, and a discussion of this is achieved. We have special interest in discussing how the seams affect the aerodynamics of the ball; therefore, the lift force and non-spinning pitches will get more attention.

This review is structured according to the classification of pitches mentioned above. In this way, we begin by presenting the most common mathematical model to describe the movement of a baseball with a concise review of the spinning pitches in section 2. Next, section 3 deals fully with the advances on non-spinning pitches and presents some comments on the boundary layer and the wake of a baseball in motion. Finally, we present a compilation of the existing uses regarding the aerodynamics

of baseballs, and outline the research trends of the subject in section 4.

2. SPINNING PITCHES

In accordance with **Figure 1**, we define a spinning pitch as any throw, excepting knuckleballs, that has an initial velocity in the range $\mathbf{V} = [(V_x, V_y, V_z)] \in [(-3, 30, -3), (3, 50, 3)]$ m/s and an initial spin of $|\boldsymbol{\omega}| = [100, 310]$ rad/s, which are the values at which baseballs are thrown by professional pitchers when fixing the $y-$axis in mound-home direction and $z-$axis perpendicular to the floor, according to the right-hand rule [13]. In this way, some examples of spinning pitches are the curveball, slider, change-up, and the variants of the fastball. The hallmark of this type of pitches is the Magnus force. We begin by describing the drag force in section 2.1, which is present in all types of pitches, with the aim of establishing some basic concepts to introduce the Magnus force.

2.1. The Drag Force

Drag or friction is the force that resists the movement of an arbitrary object. The classical way to derive drag force is by considering that the study of a ball moving through a static medium is equivalent to the one of a static ball with fluid in motion; therefore, $\mathbf{V}$ corresponds to the velocity of the fluid, and thus a mathematical formula for the drag of the ball can be obtained from the momentum conservation equation [2, 14]

$$\frac{\partial \mathbf{V}}{\partial t} + (\mathbf{V} \cdot \nabla)\mathbf{V} = -\nabla \left(\frac{p}{\rho}\right) + \mathbf{g} \qquad , p \equiv \text{pressure,} \quad (2)$$

which, in turn, is computed by considering that the air is a Newtonian and incompressible fluid in the Reynold's transport theorem (1); see Ferziger 1996 [2] for the computing of Equation (2). The conservation of mass equation also reduces to

$$\nabla \cdot \mathbf{V} = 0. \quad (3)$$

Then, assuming a steady flow the Bernoulli's equation results

$$\frac{1}{2}V^2 + \frac{p}{\rho} \equiv \text{constant,} \quad (4)$$

where we used $\frac{1}{2}\nabla V^2 = \mathbf{V} \times (\nabla \times \mathbf{V}) + (\mathbf{V} \cdot \nabla)\mathbf{V}$ avoiding the gravitational term [14]. From Bernoulli's Equation (4) the reader can observe that the greatest pressure occurs at points where the velocity equals zero. These points are commonly found on the surface of solid bodies and are usually called stagnation points. For the case of a sphere ball, one stagnation point is present on the front side. The pressure at this point is

$$p_{\max} = p_0 + \frac{1}{2}\rho V^2 \quad (5)$$

where p_0 is the pressure of the fluid at infinity [14, 15]. In this way, the second term is related to the force opposing the motion of the ball, the drag force, and thus drag can be calculated by the difference of pressure between the front and back sides of the

ball [2]. However, there are other phenomena affecting the drag. At the corresponding Reynolds number (Re) of a typical throw ($10^4 - 10^5$) the flow is not steady but turbulent [8, 10, 13, 16, 17]; then the point of separation of the boundary layer moves away from the front of the ball when the ball's velocity increases, making the wake smaller and avoiding a momentum in the reverse direction on the rear of the ball [18, 19].

The approximation of the drag force is done by introducing the sectional transverse area A and a dimensionless coefficient C_d into the second term of the right-hand side in the Equation (5), such that

$$\mathbf{F}_d = -\frac{1}{2}\rho C_d A V \mathbf{V}. \quad (6)$$

where $C_d = C_d(V)$, and then C_d acts in (6) as a fraction of the area A interacting with the air, or in other words, a measure of how aerodynamic the ball is [18, 20, 21].

As a consequence, the study of the drag coefficient (C_d) is so important that it becomes the most effective formula to approximate the drag; therefore, the drag force ($\mathbf{F}_d$) is commonly expressed in terms of C_d. There are two ways to explain the behavior of C_d. On the one hand, there is the drag crisis phenomenon for smooth spheres mentioned in Landau and Lifshitz [14], which indicates that a crisis in the value of C_d can occur at Re$\sim 10^5$ (at velocities of 30–40 m/s for a typical baseball). Data shown in Frohlich [22] and Cross [23] fit to such model. On the other hand, from the famous curve of Adair [7] to the recent data obtained by Naito [24], most of the experimental measures suggest a model without drag crisis [6, 17, 25, 26]. This predominance maintains for other sphere balls [23, 27–29] and gives rise to most of the used models to compute the drag force without including the drag crisis phenomenon. There are two models of special interest: the model by Cross [6], from which the drag coefficient can be obtained if the instantaneous ball's velocity at a fixed time is known, and the curve of Adair [7], which approaches to

$$C_d(V) = 0.29 + 0.22 \left[1 + e^{(V-32.37)/5.2}\right]^{-1}, \quad (7)$$

according to Aguirre-López etal. [5]; see **Figure 2**.

Finally, it is important to mention that the drag has an oscillating dependence on the orientation of the seams. However, for spinning pitches such oscillations average to zero and so they are not considered here. The dependence of C_d on the seams' orientation will be discussed in section 3.1.

2.2. The Magnus Force

Magnus force is the essential characteristic of a spinning pitch, as we mentioned before. The Magnus effect is observed as smooth deflections in the trajectory of a ball. All of us have a clear empirical knowledge of the Magnus effect: large deflections are reached by increasing the spin frequency of the ball. Although this statement is true, the direction of the deflection varies for different configurations of linear velocity $\mathbf{V}$, the angular velocity $\boldsymbol{\omega}$ and the ball properties. In fact, for any viewer, the expected direction of the ball's deflection goes on ($\boldsymbol{\omega} \times \mathbf{V}$),

FIGURE 2 | Model of drag and Magnus coefficients. Adair's drag model [7] and the approximation of Aguirre-López et al. [5] showing the sigmoidal decreasing of drag when increasing the velocity of the ball. In turn, the Magnus model (10) presented in Robinson [8] shows that the Magnus coefficient increases when the angular velocity increases. Modified from Aguirre-López et al. [5].
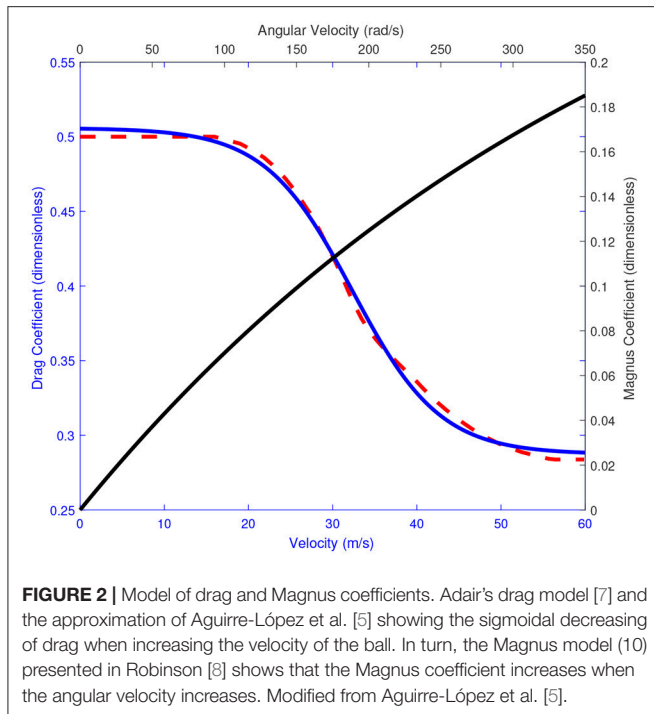
as illustrated in **Figure 3**. However, a reverse direction of the Magnus force has been reported for smooth balls like those used in soccer [27], and also in smooth spheres simulating baseballs in experiments by Briggs [30], Cross and Lindsey [31]. This phenomenon is commonly called the "anti-Magnus effect" and it is possible only for a range of Re and spin when one side of the smooth ball remains in a laminar flow while the opposite side becomes turbulent. Then, a low pressure region is originated in the turbulent side because it is generally farther to the ball surface than the laminar layer. Thus, the ball moves to the region with lower pressure by conservation of momentum, as illustrated in **Figure 3**. For a detailed explanation of the causes of the reverse in the direction of Magnus force, the reader is referred to Cross and Lindsey [31]. For a general understanding of the phenomenon, the reverse-Magnus occurs at Re$\sim 10^5$ combined with a low speed due to Magnus force $R\omega$ (where R is the radius of the ball) compared with the ball's velocity V, such that the spin factor $S = R\omega/V$ is in the range 0–0.6.

However, the baseball is not smooth and there are no reported studies of an anti-Magnus effect in baseballs. This is because the seams of the ball give to it some roughness that accelerates the separation of the boundary layer in both the up and down sides of the ball. Indeed, the Magnus effect in baseballs' flight arises because one side of the ball offers larger friction than its opposite side, which means that the speed of the main flow of air is larger on the former and as a consequence the lower-pressure region locates at the opposite side, according to the anti-Magnus phenomenon [6–8]; see the schemes in **Figure 3**. Therefore, the nature of the Magnus force is similar to that of the drag force, due to a difference of pressure. For this reason, Magnus force is

commonly written in a similar way as drag force (Equation 6), namely,

$$F_{\mathrm{M}} = \frac{1}{2}\rho A C_{\mathrm{M}} V^2, \tag{8}$$

for an arbitrary direction of motion, with C_M being the Magnus coefficient. Moreover, in the same system of coordinates to drag force, Equation (8) becomes

$$\mathbf{F}_{\mathrm{M}} = \frac{1}{2}\rho A C_{\mathrm{M}} \sin\phi V^2 \hat{\mathbf{n}}, \tag{9}$$

where the unit vector $\hat{\mathbf{n}} = \frac{\boldsymbol{\omega}\times\mathbf{V}}{|\boldsymbol{\omega}\times\mathbf{V}|}$ gives the direction of the resulting linear momentum, ϕ is the angle between $\boldsymbol{\omega}$ and $\mathbf{V}$ so that $(V\sin\phi)$ is the component of V that contributes to the force [6, 8, 10]. It is important to remark that equation (9) has been widely used to reproduce Magnus force of sport balls and other areas of aerodynamics [32–34], and it has been proved experimentally (for sport balls) only for $\phi = 0, 90$ and $180°$ [8, 10].

The Magnus coefficient C_M is a function of ω and V for arbitrary magnitudes of such variables because Equation (9) depends on both the instantaneous and spin velocities [6, 10, 31]. However, Nathan [17] has reported that C_M behaves independently of V for ω and $\mathbf{V}$ values in the range at which spinning pitches are thrown. Moreover, the value of C_M in such range is similar to the Magnus coefficient for other sport balls, and there are some models to calculate its value [8, 30, 35]. **Figure 3** shows the model (10):

$$C_{\mathrm{M}}(\omega) = 3.19 \times 10^{-1} \left[1 - e^{-2.48\times10^{-3}\omega} \right], \tag{10}$$

which has been used to compute the Magnus force in numerical simulations [5, 8].

2.3. Discussion and Potential Research Trends

In the following items, we summarize and discuss the highlights of the spinning pitches, and also sketch out the potential research trends of the subject:

- Both drag and Magnus forces are commonly expressed in terms of their coefficients, which have information about the object taken from experiments.
- The drag coefficient (7) decreases in a sigmoidal way when the ball's velocity increases. It maintains a value of around $C_d = 0.5$ up to velocities of 20 m/s, then decreases to $C_d = 0.3$ in the range of 20–50 m/s (which corresponds to the range of spinning pitches), and maintains that value for larger velocities. In this way, the drag coefficient must be considered as a function of the velocity when computing baseball trajectories.
- There are different ways to measure drag force or estimate the drag coefficient that have not been reported, for instance, by Computational Fluid Dynamics (CFD) or by analyzing the von Karman vortex trails generated by the ball [2, 15].
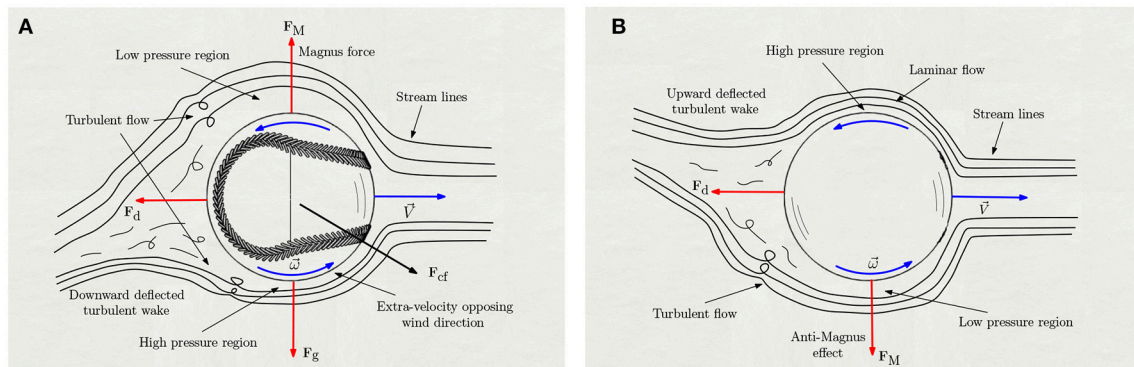
FIGURE 3 | Scheme of the possible ways of causing a Magnus force. **(A)** The common Magnus force. Modified from Aguirre-López et al. [5]. **(B)** The reverse Magnus force.

Both of them could be promising areas of opportunity to characterize the drag, and the aerodynamics of a baseball, in a more comprehensive way.

- Model (9) is the most common formula to approximate the Magnus force. It considers the angular velocity ω as a constant in time (despite torquing forces). This is acceptable since such forces are very small, as mentioned in Ward-Alaways [10].
- The exponential behavior of the Magnus coefficient (10) denotes the effect of ω in the Magnus force. In addition, the model does not depend on the velocity V inside the range of initial conditions for spinning pitches, which simplify the estimation of the Magnus effect.
- The effects of drag and Magnus forces on a spinning pitch can be decoupled and studied separately according to Aguirre-López et al. [5], which could have a lot of applications, such prediction, reconstruction and clustering of trajectories; see section 4.
- Other aspects exist that affect a thrown baseball, such as wind, humidity, etc. However, the study of these effects is more complex and there are few investigations about them [6, 8].

3. NON-SPINNING PITCHES

Non-spinning pitches consist of a specific type of throw: the knuckleball. It contains all combinations for the linear and angular velocities in the ranges $V \in [20, 40]$ m/s and $\omega \in [0, 50]$ rad/s, respectively [36]. Knuckleball pitches are the most interesting throws for aerodynamics because the ball can no longer be considered a sphere since the effect of the seams is significant. However, this is more complex to understand and, therefore, knuckleball studies are fewer than those for spinning pitches.

In this section we discuss the advances in drag force for non-spinning pitches (section 3.1). Then, the description of the models for lift and side forces is presented in section 3.2. An introduction to the modeling of the seams and the boundary layer observations is shown in section 3.3. A compilation of the knuckleball's research and trends is discussed in section 3.4.

3.1. The Drag Force in Non-spinning Pitches

Drag force is different in knuckleballs than in spinning pitches. The drag is approximately constant for a specific velocity in a spinning ball, however, in a knuckleball it is not. As mentioned in section 2.1, the Re of a baseball pitch corresponds to an unsteady flow and then an experimental coefficient must be introduced in the model for the drag force (6). However, when the ball does not spin some vortices are shed from the ball and then the drag oscillates in time. Ferziger and Perić [2] discusses such an effect for a smooth cylinder simulated by CFD. The drag on the cylinder oscillates periodically with a frequency according to the appearance of the vortices such that it has one maximum and one minimum during the formation and shedding of each vortex. Such vortex shedding has also been reported for baseballs in Texier et al. [37], whose effect in lift and side forces will be discussed in section 3.2.

In addition to the variation in time, the drag changes when varying the orientation of the ball. Investigations carried out in the present decade show structured oscillations of drag coefficient despite the turbulence present in the phenomenon. For example, the experiment by Higuchi and Kiura [38] with different configurations of the ball, namely, the four-seam (4S), the two-seam (2S) and an arbitrary orientation of the ball[1]. They found the largest variation in oscillations for the 4S orientation, which is about twice as large as the case of the 2S orientation and around four times that of the arbitrary orientation. The shape of the oscillations in the 4S orientation is maintained for ball velocities in the range of 16–30 m/s.

Similar average variations for C_d have been reported by Alam et al. [25] in studies of the drag force for Major League Baseball (MLB) and National Collegiate Athletic Association (NCAA) baseballs and softballs. They reported lower variations in C_d values for NCAA than for MLB baseballs at different ball orientations, which suggests a dependence on the height of the seams (1.5 mm for NCAA and 1mm for MLB balls). At first

---

[1]A detailed explanation of the typical baseball orientations can be found in Borg and Morrisey [36].

glance, the idea seems to be solid because a larger "extra-obstacle" should cause a corresponding "extra-drag" in the fluid. Even more, this is supported by the results of Kensrud et al. [39], who analyzes hit balls with different heights of seams and found that balls with smaller seams reach larger distances, which indicates a lower drag.

Finally, it is important to remark that commonly C_d decreases with increasing V for baseballs, softballs, cricket balls and smooth spheres. In the case of baseballs, the value of C_d decreases from ~ 0.6 to ~ 0.4 dimensionless units [25, 32].

3.2. The Lift and Side Forces

The main reason for which knuckleball pitches are much more unpredictable than spinning pitches is because of lift and side forces[2]. Similar to the drag and Magnus forces, lift force is produced by an imbalance in pressure and then it is proportional in magnitude to the square of the ball velocity [11, 40], so that

$$F_L = \frac{1}{2}\rho C_L A V^2. \tag{11}$$

However, the behavior of the lift coefficient (C_L) is not like C_d or C_M but it varies with the angle of attack θ of the ball. **Figure 4** shows how unpredictable a knuckleball can be, even for the most symmetric ball orientations (4S and 2S). Results of Borg and Morrisey [36] show four cycles in C_L for the 4S orientation, each one with a period of 90°, with a semi-sinusoidal behavior and ringlets at the end of a cycle. This means that a variation of $\sim 22.5°$ in smooth-angle zones may or may not produce a maximum/minimum lift; instead, a variation of only $\sim 10°$ in the zone of ringlets can produce any type of motion. In addition to such complexity for a strictly non-spinning ball, in practice it is difficult for a ball to travel without rotating since a little spin is induced by contact from the air with the seams [12, 37]. As a consequence, the ball's trajectory could have an apparent erratic motion whereas the map of balls passing through the home plate in a real pitch is seen as random when varying θ [40]. All of this makes it difficult to compute a model that may accurately reproduce the trajectory of a knuckleball with any orientation.

It is important to mention that there is controversy in the causes that originate the lift force. It is evident that the asymmetry of the seams plays a fundamental role in causing such force. According to Watts and Sawyer [11] and Mehta [12], there are two possible ways to produce a lift force on a baseball: for many years, the classical hypothesis stated that the lift is produced not only by the seams but by the shedding of vortices that occurs at the rear side of the ball. All of these interact in a complex way, as mentioned in Ferziger and Perić [2], Watts and Sawyer [11], and Mehta [12]. On the other hand, and according to Texier et al. [37], there is the possibility that the lift force may be originated only by the perturbations at the front side of the ball. This means that the seams produce the total lift of the ball, especially those located at the separation or critical points of the boundary layer at 52, 140, 220, and 310° [12, 36]; this will be addressed in detail in section 3.3. Texier calculated that the force produced by the

---

[2]From this point on we will use the term "lift force" for referring to both lift and side forces because both are equivalent, excluding gravity.

vortices at Strohual numbers (St) of St~ 0.2 is significantly lower than the magnitude of the lift measured in experiments, so that it practically does not contribute to lift. More information about lift force in ball sports can be found in investigations by Mehta [12], Hong et al. [42], and Murakami et al. [43].

3.3. The Seams and the Boundary Layer

The manner in which the seams affect the boundary layer is very sensitive to small changes between seams. As commented by Borg and Morrisey [36], when a seam is located near the natural separation angle (the angle of separation of a smooth ball), it can induce turbulence and consequently provoke a delay in the separation of the boundary layer; in turn, the seam can force the separation to occur and cause an advance in the separation angle. Such effects are seen in the sudden changes in the values of the lift coefficient for a 4S orientation. For a smooth ball, a sinusoidal shape of C_L would be expected, in a similar way to the lift coefficient reported for a smooth cylinder by Ferziger and Perić [2]. However, **Figure 4** suggests a quasi-periodic behavior for C_L for 4S balls, with fast changes in magnitude and direction at 52, 140, 220 and 310°, as observed by Watts and Sawyer [11]. This is because the separation point is located around such degrees and then it advances or delays with a little variation in the stitch position.

Experiments by Higuchi and Kiura [38] show that a variation of only one degree (36 to 37°) in the stitch position causes a sudden separation. Moreover, they reported that the balls are more susceptible to hysteresis (including induced rotation) at the zones of separation. For 4S balls and Re above 1.5×10^5, the ball is sensitive to the initial rotation, namely, spins of 5 rad/s become 10.5 rad/s, increasing linearly and having a spin limit of 18.9 rad/s even for Re above 2×10^5. In turn, for 2S balls they found that the oscillation frequency is constant over Re$\in [1.9 \times 10^5, 4.6 \times 10^5]$. As a consequence of the induced rotation, the phenomenon becomes more unpredictable because the separation point moves forward or backward at every moment of time. This is the reason why the throws inside the intermediate window in **Figure 1** are the most difficult to study. We invite the reader to consult the research of Higuchi and Kiura [38] for detailed observations of the boundary layer.

To end the collection of the advances on knuckleballs, it is important to mention the phenomenological model proposed by Aguirre-López et al. [41] for computing the lift coefficient. It consists of computing a super-imposition of the forces produced by the vortex shedding and each stitch, so that

$$C_L(\theta) = a_0 \sin(4\theta - \pi) + a_1 \sum_{i=1}^{n} \left[\sin\left(\frac{||\mathbf{s}_i - \mathbf{p}||\pi}{2d} + \pi/2 \right) \cdot \mathrm{sgn}\left(p^* - s_i^* \right) \right], \tag{12}$$

where $C_L = C_L(\theta)$ is now a function of the angle of attack of the ball, the first term in the right-hand side is the force caused by the vortex shedding and the term $\sum(\cdot)$ is the sum of forces produced by the seams, $\mathbf{p}$ is the stagnation point, $\mathbf{s}_i$ is the position of the i-th stitch, s_i^* and p^* are the $z-$components of $\mathbf{s}_i$ and $\mathbf{p}$, respectively, and a_0 and a_1 are weight coefficients.
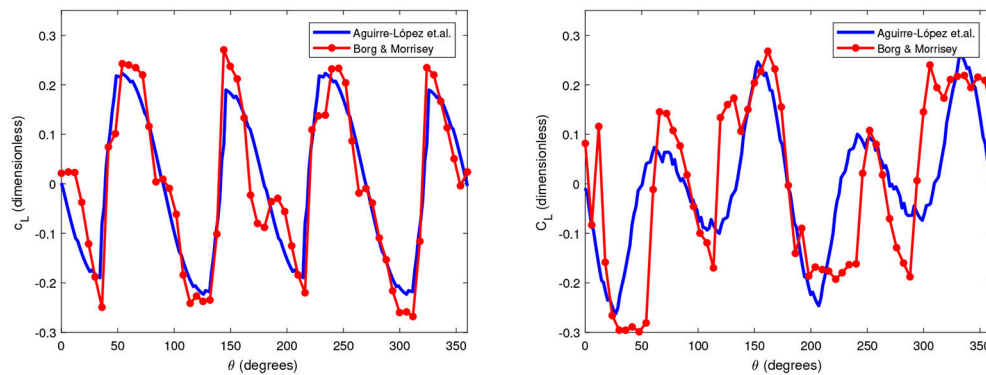
FIGURE 4 | Experimental data [36] and model [41] of C_L as a function of the attack angle θ for 4S (A) 2S (B) balls. Modified from Aguirre-López et al. [41].

Therefore, each stitch in model (12) produces a force whose magnitude decreases smoothly when the stitch moves away from the stagnation point and takes into account the symmetry on the z-axis by introducing the sign function sgn(·). Thus, despite the fact that model (12) does not consider effects of hysteresis and high sensitivity to perturbations at the separation point, it fits the experimental data of Borg and Morrisey [36] for 4S and 2S orientations, as shown in **Figure 4**. Model (12) opens the door to future research on how the seams and the vortex shedding affect the lift force.

3.4. Discussion and Potential Research Trends

Here we summarize and discuss the highlights of the non-spinning pitches and outline potential research trends of knuckleballs pitches as follows:

- The structure of the oscillations of the drag coefficient depends on the orientation of the ball, whereas the height of the seams increases the magnitude of the drag.
- The value of C_D decreases from 0.6 to 0.4 units when increasing V.
- The lift coefficient oscillates every 90°, with a quasi-periodic behavior for 4S balls, which is related to the effect of the seams. In fact, values of C_L for 2S balls oscillate every 180°, with an inversion every cycle [36, 38].
- The origin of lift force is not well understood. On the one hand, seams could cause the total lift, and on the other hand, a sum of both the seams and the vortex shedding could be the source of it [11, 36, 37].
- Observations on the boundary layer suggest that the lift force is more susceptible to perturbations at some angles, including 52, 140, 220, and 310° for 4S balls. Hysteresis is partly responsible for this [38].
- We consider that simulations using CFD techniques could disentangle the causes that produce the lift force.
- In addition to the last point, CFD simulations could help to improve the model (12) or propose a variation of it that involves the susceptible zones of the separation point, and extend the model to arbitrary orientations.

4. APPLICATIONS

As the reader may suppose, there are many ways to make useful the information compiled in sections 2 and 3. And, indeed, baseball studies have been the basis of numerous technologies on the matter, specifically those ones about spinning pitches. We finalize this work with a brief summary of the main applications of the aerodynamic forces on baseballs. Section 4.1 is focused on the studies related to baseball's trajectories. In turn, section 4.2 talks about the complementary applications including the best known of them: the PITCHf/x algorithm.

4.1. Prediction and Reconstruction of Trajectories

4.1.1. The Simulation Problem
The most simple use of aerodynamic forces is the simulation or prediction of trajectories. A simulation of a baseball pitch is frequently carried out by using the Equations (6) and (9) along with gravity to compute a model of forces as:

$$m\dot{\mathbf{V}} = \mathbf{F}_d + \mathbf{F}_M + m\mathbf{g}, \qquad (13)$$

which can be solved numerically by Runge-Kutta-4 or other integration methods, whereas drag and Magnus coefficients can be computed by Equations (7) and (10) or similar approximations [6, 44]. The simulations could become more realistic by including eventual forces in the model (13). An example of this is the model of Robinson and Robinson [8], which adds a constant in wind to the ball velocity so that $\mathbf{V}'$ is redefined as ($\mathbf{V}' = \mathbf{V} + \mathbf{W}$), where $\mathbf{W}$ is the wind velocity.

In turn, simulations of baseball trajectories are commonly applied to some sport and technology areas such as in video games [45], baseball machines [10] as well as for instruction for baseball players. The last one is the main reason for which research on knuckleballs is a topic of special interest.

4.1.2. The Reconstruction Problem
The counterpart of simulation is the reconstruction of trajectories. In these works, there is no possibility to give

the initial conditions of a pitch and obtain the trajectory but the trajectory must be extracted, tracking or reconstructed only by a set of 3D or 4D (space plus time) points belonging to the trajectory. Such data points are commonly recorded by baseball broadcast videos and/or images of real games; therefore, the trajectory obtained is used for the replay in television.

Various methodologies have been reported for extracting or tracking the trajectory. Most of them use types of diverse filters like color, position, size and shape [46–48], and others [49, 50] select possible trajectories. Then the chosen trajectories are compared with the model (13) so that, if the resulting trajectory does not agree with the model, then it is discarded and a new one is needed. Takahashi et al.'s investigation [47] also deals with classifying the type of pitch by relating a total of 36 features, including the shape and speed calculated from the ball trajectory data and the ball speed from the screen display. They report an accuracy of $\sim$ 89% with their methodology.

The methodologies that deal with a "direct" reconstruction are based on the use of the equations of motion. Shum and Komura [51] and Miyata et al. [52] use color filtering for detecting 2D candidate trajectories. Then, Shum and Komura estimate the depth of the ball in the scene by introducing a model (13). In turn, Miyata et al. [52] chose one candidate by fitting a uniformly accelerated motion model [similar to model (13)] and finally, they use multiple cameras calibrated temporally and geometrically to obtain a 3D trajectory. On the other hand, Aguirre-López et al. [5] developed an algorithm that directly solves the model (13) in two interrelated parts by decoupling the Magnus force from the equations of motion, using the Newton-Raphson method when knowing three points of the trajectory. They reported absolute error values of $\sim$ 0.1 mm between simulated and reconstructed trajectories. Finally, Kagan and Nathan [53] have developed a software called the *trajectory calculator*, which is similar in operation to that of Aguirre-López et al. [5] but with simpler assumptions. The results are less accurate but it is a good tool to start in the subject. The *trajectory calculator* can be downloaded directly at [54].

4.2. The PITCHf/x Algorithm and Clustering

The second part of the application deals with problems related to the classification of trajectories, among which the PITCHf/x algorithm is the most popular and accurate reported method in research and in the world of baseball. The algorithm (including new versions and software packages) has been consolidated as a powerful tool in the area of pitch classifications [55, 56].

The PITCHf/x algorithm consists of two parts. The first one involves reconstructing trajectories by estimating the coefficients C_d, C_M and the spin axis ϕ (the angle between $y-$axis and $\boldsymbol{\omega}$) using non-linear least-squares fitting with the Levenberg-Marquardt algorithm. Nathan [55] reports very good adjustments; indeed, root-mean-square deviations of the fitted trajectory of around 1 mm in each dimension. The second (and the main) part of the work deals with the classification of pitches. The classification is based on ($\mathbf{V}$ vs. ϕ) and ($\boldsymbol{\omega}$ vs. ϕ) graphics. As a result, the types of pitches are arranged in clusters in polar scatter plots and scatter plots of the deflection of the ball at home. Pane [57] carried out an interesting cluster analysis from the results of Nathan [55] based on PITCHf/x. The research on the topic continues.

ETHICS STATEMENT

We express our gratitude to Elsevier for the permission to reuse a modified version of the copyrighted **Figures 2**, **3a**, **4** in the present paper. Order numbers: 4511461494464 and 4511510076028.

AUTHOR CONTRIBUTIONS

GE and OD-H contributed to the design and discussion of this document. MA-L contributed to the design and writing of this document. FH-Z contributed to the discussion and writing of this document. JM-C and F-JA contributed to the discussion of this document and as advisor.

ACKNOWLEDGMENTS

REFERENCES

1. White FM. *Fluid Mechanics*. New York, NY: McGraw Hill (2003).

2. Ferziger JH, Perić M. *Computational Methods for Fluid Dynamics*. Berlin: Springer (1996).

3. Fitzpatrick R. *Theoretical Fluid Mechanics*. Available online at: http://farside.ph.utexas.edu/teaching/336L/Fluid.pdf

4. Robinson G, Robinson I. Are inertial forces ever of significance in cricket, golf and other sports? *Phys Scripta* (2017) **92**:043001. doi: 10.1088/1402-4896/aa634e

5. Aguirre-López MA, Morales-Castillo J, Díaz-Hernández O, Escalera Santos GJ, Almaguer F-J. Trajectories reconstruction of spinning baseball pitches by three-point-based algorithm. *Appl Math Comput.* (2018) **319**:2–12. doi: 10.1016/j.amc.2017.01.016

6. Cross R. *Physics of Baseball & Softball*. New York, NY: Springer (2011).

7. Adair RK. *The Physics of Baseball*. New York, NY: Harper-Collins Publishers (2002).

8. Robinson G, Robinson I. The motion of an arbitrarily rotating spherical projectile and its application to ball games. *Phys Scripta* (2014) **88**:018101. doi: 10.1088/0031-8949/88/01/018101

9. Cross R. *Information About Knuckleballs*. Available online at: http://www.physics.usyd.edu.au/ cross/KNUCKLEBALLS.htm

10. Ward-Always L. *Aerodynamics of the Curve-Ball: an Investigation of the Effects of Angular Velocity on Baseball Trajectories*, Philosophy Doctoral Thesis, University of California (1998).

11. Watts RG, Sawyer E. Aerodynamics of a knuckleball. *Am J Phys.* (1975) **43**:11. doi: 10.1119/1.10020

12. Mehta RD. Aerodynamics of sports balls. *Annu Rev Fluid Mech.* (1985) **17**:151–89. doi: 10.1146/annurev.fl.17.010185.001055

13. MLB Games (2015). *Statistics of 2015 MLB Games*. Available online at: http://m.mlb.com/news/article/160896926/statcast-spin-rate-compared-to-velocity

14. Landau LD, Lifshitz EM. *Fluid Mechanics*. Elsevier (1987).

15. Prandtl L, Tietjens OG. Applied hydro- and aeromechanics. USA: Dover (2016).

16. Official Baseball Rules (2015). Available online at: http://mlb.mlb.com/mlb/downloads/y2015/official_baseball_rules.pdf

17. Nathan AM. The effect of Spin on the Flight of a Baseball. *Am J Phys.* (2008) **76**:119–24. doi: 10.1119/1.2805242

18. Watts RG, Moore G. The drag force on an american football. *Am J Phys.* (2003) **71**:791–3. doi: 10.1119/1.1578065

19. Luthander S, Ryberg A. Experimentelle Untersuchungen uber den Luftwiderstand bei einer um ein mit der Windrichtung parallelen Achse roteirenden Kugel. *Phys Z* (1935) **36**:552–8.

20. Daish CB. *The Physics of Ball Games*. London: The English Universities Press (1972).

21. Hennekam W. The speed of a cyclist. *Phys Educ.* (1990) **25**:141–6. doi: 10.1088/0031-9120/25/3/001

22. Frohlich C. Aerodynamic drag crisis and its possible effect on the flight of baseballs. *Am J Phys.* (1984) **52**:325–34. doi: 10.1119/1.13883

23. Cross R. Effects of turbulence on the drag force on a golf ball. *Eur J Phys.* (2016) **37**:054001. doi: 10.1088/0143-0807/37/5/054001

24. Naito K, Hong S, Koido M, Nakayama M, Sakamoto K, Asai T. Effect of seam characteristics on critical Reynolds number in footballs. *Mech Eng J.* (2018) **17**:00369.

25. Alam F, Ho H, Smith L, Subic A, Chowdhury H, Kumar A. A study of baseball and softball aerodynamics. *Proc Eng.* (2012) **34**:86–91. doi: 10.1016/j.proeng.2012.04.016

26. Kensrud JR, Smith LV. *In situ* drag measurements of sports balls. *Proc Eng.* (2010) **2**:2437–42. doi: 10.1016/j.proeng.2010.04.012

27. Kray T, Franke J, Frank W. Magnus effect on a rotating soccerball at high Reynolds numbers. *J Wind Eng Indust Aerodyn.* (2014) **124**:46–53. doi: 10.1016/j.jweia.2013.10.010

28. Cross R. Vertical impact of a sphere falling into water. *Phys Teacher* (2016) **54**:153–5. doi: 10.1119/1.4942136

29. Goff JE. A review of recent research into aerodynamics of sport projectiles. *Sports Eng.* (2013) **16**:137–54. doi: 10.1007/s12283-013-0117-z

30. Briggs LJ. Effect of spin and speed on the lateral deflection (curve) of a baseball; and the magnus effect for smooth spheres. *Am J Phys.* (1959) **27**:589–96. doi: 10.1119/1.1934921

31. Cross R, Lindsey C. Measurements of drag and lift on smooth balls in flight. *Eur J Phys.* (2017) **38**:044002. doi: 10.1088/1361-6404/aa6e44

32. Sayers AT. Hill A. Aerodynamics of a cricket ball. *J Wind Eng Indust Aerodyn.* (1999) **79**:169–82. doi: 10.1016/S0167-6105(97)00299-7

33. Goodwill SR, Chin SB, Haake SJ. Aerodynamics of spinning and non-spinning tennis balls. *J Wind Eng Indust Aerodyn.* (2004) **92**:935–58. doi: 10.1016/j.jweia.2004.05.004

34. Lorentz RD. *Spinning Flight. Dynamics of Frisbees, Boomerangs, Samaras, and Skipping Stones*. New York, NY: Springer (2006).

35. Robinson G, Robinson I. Reply to Comment on 'The motion of an arbitrarily rotating spherical projectile and its application to ball games'. *Phys Scripta* (2014) **89**:5. doi: 10.1088/0031-8949/89/6/067002

36. Borg JP, Morrisey MP. Aerodynamics of the knuckleball pitch: experimental measurements on slowly rotating baseballs. *Am J Phys.* (2014) **82**:921. doi: 10.1119/1.4885341

37. Texier BD, Cohen C, Quéré D, Clanet C. Physics of knuckleballs. *N J Phys.* (2016) 18:073027. doi: 10.1088/1367-2630/18/7/073027

38. Higuchi H, Kiura T. Aerodynamics of knuckle ball: flow-structure interaction problem on a pitched baseball without spin. *J Fluids Struct.* (2012) **32**:65–77. doi: 10.1016/j.jfluidstructs.2012.01.004

39. Kensrud JR, Smith LV, Nathan A, Nevins D. Relating baseball seam height to carry distance. *Proc Eng.* (2015) **112**:406–11. doi: 10.1016/j.proeng.2015.07.216

40. Nathan AM. Analysis of knuckleball trajectories. *Proc Eng.* (2012) **34**:116–21. doi: 10.1016/j.proeng.2012.04.021

41. Aguirre-López MA, Díaz-Hernández O, Almaguer F-J, Morales-Castillo J, Escalera Santos GJ. A phenomenological model for the aerodynamics of the knuckleball. *Appl Math Comput.* (2017) **311**:58–65. doi: 10.1016/j.amc.2017.05.001

42. Hong S, Chung C, Nakayama M, Asai T. Unsteady aerodynamic force on a knuckleball in soccer. *Proc Eng.* (2010) **2**:2455–60. doi: 10.1016/j.proeng.2010.04.015

43. Murakami M, Seo K, Kondoh M, Iwai Y. Wind tunnel measurement and flow visualisation of soccer ball knuckle effect. *Sports Eng.* (2012) **15**:29–40. doi: 10.1007/s12283-012-0085-8

44. Clark JM, Greer ML, Semon MD. Modeling pitch trajectories in fastpitch softball. *Sports Eng.* (2015) **18**:157–64. doi: 10.1007/s12283-015-0176-4

45. Syamsuddin MR, Kwon YM. Simulation of baseball pitching and hitting on virtual world. *Phys Scripta* (2011) 663–667.

46. Chu WT, Wang CW, Wu JL. Extraction of baseball trajectory and physics-based validation for single-view baseball video sequences. *Int Cong Multimedia Expo* (2006) 1813–16. doi: 10.1109/ICME.2006.262905

47. Takahashi M, Fujii M, Yagi N. Automatic pitch type recognition from baseball broadcast videos. *Int Symp Multimedia* (2008) 15–22.

48. Chang P, Han M, Gong Y. Extract highlights from baseball game video with hidden Markov models. In: *International Congress of Imaging Processing, Vol. 1,* (2002). p. 609–12.

49. Hung MH, Hsieh CH. Event detection of broadcast baseball videos. *Trans Circ Syst Video Technol.* (2008) **18**:1713–26.

50. Guéziec A. Tracking pitches for broadcast television. *Computer* (2002). doi: 10.1109/2.989928

51. Shum H, Komura T. A spatiotemporal approach to extract the 3D trajectory of the baseball from a single view video sequence. In: *International Conference on Multimedia and Expo, Vol. 3* (2004) p. 1583–86. doi: 10.1109/ICME.2004.1394551

52. Miyata S, Saito H, Takahashi K, Mikami D, Isogawa M, Kimata H. Ball 3D trajectory reconstruction without preliminary temporal and geometrical camera calibration. In: *IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2017).

53. Kagan D, Nathan AM. Statcast and the baseball trajectory calculator. *Phys Teach.* (2017) 55:134. doi: 10.1119/1.4976652

54. Nathan AM. *The Trajectory Calculator*. Available online at: http://baseball.physics.illinois.edu/trajectory-calculator-old.html

55. Nathan AM. *Analysis of PITCHf/x Pitched Baseball Trajectories*. (2008). Available online at: https://www.researchgate.net/publication/228563555_Analysis_of_PITCHfx_Pitched_Baseball_Trajectories

56. Sievert C. Taming PITCHf/x Data with XML2R and pitchRx. *R J.* (2014) **6**:5–19.

57. Pane MA. *Trouble With the Curve: Identifying Clusters of MLB Pitchers using Improved Pitch Classification Techniques*, Philosophy Doctoral Thesis, Carnegie Mellon University (2013).

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: info@frontiersin.org | +41 21 510 17 00

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