The Stochastic-Quantum Correspondence

Jacob A. Barandes[1,*]

[1] *Jefferson Physical Laboratory, Harvard University, Cambridge, MA 02138*

(Dated: September 3, 2023)

This paper introduces an exact correspondence between a general class of stochastic systems and quantum theory. This correspondence provides a new framework for using Hilbert-space methods to formulate highly generic, non-Markovian types of stochastic dynamics, with potential applications throughout the sciences. This paper also uses the correspondence in the other direction to reconstruct quantum theory from physical models that consist of trajectories in configuration spaces undergoing stochastic dynamics. The correspondence thereby yields a new formulation of quantum theory, alongside the Hilbert-space, path-integral formulations, and quasiprobability formulations. In addition, this reconstruction approach opens up new ways of understanding quantum phenomena like interference, decoherence, entanglement, noncommutative observables, and wave-function collapse.

I. INTRODUCTION

The theory of stochastic processes describes the phenomenological behavior of systems with definite configurations that evolve probabilistically in time. Quantum theory is a comprehensive mathematical apparatus for making measurement predictions when taking into account the microscopic constituents of various kinds of physical systems, from subatomic particles to superconductors. At an empirical level, both theories involve probabilities, and at the level of formalism, both employ vectors and matrices.

There have been a number of previous attempts in the research literature to identify a fundamental relationship connecting stochastic-process theory and quantum theory [1–7]. The most well-known of these approaches are due to Bopp [8–10], Fényes [11], and Nelson [12, 13]. Altogether different are stochastic-collapse models [14, 15], in which a quantum system's wave function or density matrix is assumed to experience stochastic fluctuations through time.

The present paper, which is not continuous with those earlier developments, introduces an exact correspondence between a highly general class of stochastic systems and quantum theory. This *stochastic-quantum correspondence* takes the form of a simple 'dictionary' expressing any time-dependent stochastic matrix in terms of a suitable combination of Hilbert-space ingredients.

From a practical standpoint, the stochastic-quantum correspondence provides a systematic framework for constructing highly generic forms of stochastic dynamics, much as the Lagrangian or Hamiltonian formulations of classical mechanics provide systematic frameworks for constructing deterministic dynamics. Potential applications range from turbulence to finance, to name just two

* jacob_barandes@harvard.edu

examples. Importantly, the stochastic-quantum correspondence does not require assuming that the stochastic dynamics in question can be modeled as a Markov chain, nor does it require making a number of other frequently invoked approximations.

Taking a more foundational perspective, this paper also uses the stochastic-quantum correspondence to show that physical models based on configuration spaces combined with stochastic dynamics can replicate all the empirical predictions of textbook quantum theory—including interference, decoherence, entanglement, noncommutative observables, and wave-function collapse—without relying on the austere and metaphysically opaque Dirac-von Neumann axioms [16, 17]. In this alternative approach, a given system moves stochastically along a physical trajectory in a classical-looking configuration space, and the mathematical objects of the Hilbert-space formulation play a functional role akin to gauge-dependent variables.

At the very least, this approach therefore yields a new formulation of quantum theory, one that is based on a picture of stochastic systems evolving in configuration spaces. This new formulation therefore joins a list of ways to formulate quantum systems that include the traditional Hilbert-space formulation [16, 17], the path-integral formulation [18–20], and the quasi-probability formulation [21, 22]. As noted in [20], "there is a pleasure in recognizing old things from a new point of view," and "there is always the hope that the new point of view will inspire an idea for the modification of present theories, a modification necessary to encompass present experiments."

In addition to establishing these new results, this paper identifies several forms of gauge invariance that have not been described in the research literature thus far, analyzes the measurement process in detail, and describes the implications of the stochastic-quantum correspondence for dynamical symmetries and for formal enlargements or dilations of a system's Hilbert space. Taking advantage of having a concrete model of stochastic physical variables in hand, this paper also revisits and clarifies a number of important questions related to the status of nonlocality in quantum theory.

Considering the mathematical simplicity of this stochastic-quantum correspondence, it is surprising that it has apparently not shown up in the research literature before. To the author's knowledge, the only previous example that bears a suggestive resemblance to the approach taken in this paper, at least at the level of some of its equations, is the unpublished draft [23].[1] Although that reference argues that some stochastic processes can be modeled using a formalism similar to that of quantum theory, it does not establish that the resulting Hilbert-space representation is fully general. Nor does it attempt to show that the correspondence is bidirectional, so that quantum systems can be modeled by stochastic processes in configuration spaces.

Section II will start with the definition of a *generalized stochastic system*, and then introduce the key distinction between *divisible* and *indivisible* dynamics. Section III will construct the Hilbert-space representation for a given generalized stochastic system, formulate the dictionary that will ultimately connect generalized stochastic systems with quantum systems, and identify an important new class of gauge transformations that have not yet been described in the research literature. Section IV will establish the stochastic-quantum correspondence, which will involve reviewing the

[1] The author thanks Logan McCarty for finding this reference.

definition of *unistochastic matrices*, introducing the concept of a *division event*, and showing how some of the most characteristic features of any quantum system—like interference, entanglement, and decoherence—can be understood from the perspective of the underlying generalized stochastic system. Section V will provide a detailed treatment of the measurement process, which will entail introducing the notion of an *emergeable*, and then turn to a larger discussion of the measurement problem and the uncertainty principle. Section VI will describe further implications of the stochastic-quantum correspondence for quantum theory, focusing on symmetries, Hilbert-space dilations, and nonlocality. Section VII will conclude the paper with a brief discussion, as well as with several open questions to be addressed in future work.

II. STOCHASTIC PROCESSES

A. Generalized Stochastic Systems

The most general kind of stochastic process requires only a sample space, an initial probability distribution, and a time-dependent random variable. (For pedagogical treatments of the theory of stochastic processes, see [24–27].) Stochastic processes defined in this way lack an ingredient that plays the role of a *dynamical law*. This paper will be concerned with a slightly more narrowly defined construction that includes the notion of a dynamical law, and that can still describe a wide variety of physical systems.

A *generalized stochastic system* will be defined to consist of a *configuration space* $\mathcal{C}$ together with a dynamical law in the form of a *stochastic map* $\Gamma(t)$ that acts linearly on probability distributions over $\mathcal{C}$ at an initial time 0 to yield corresponding probability distributions over $\mathcal{C}$ at other times $t \neq 0$. For the purposes of this paper, the set of allowed times will be assumed to be isomorphic to a subset of the real line $\mathbb{R}$ containing 0, up to a choice of measurement units.

The formalism for a generalized stochastic system is easiest to express in the case in which $\mathcal{C}$ has a finite number N of configurations labeled by positive integers $1, \ldots, N$:

$$\mathcal{C} \equiv \{1, \ldots, N\}. \tag{1}$$

In that case, the system's *standalone probabilities* at the initial time 0 can be denoted by

$$p_j(0) \quad [\text{for } j = 1, \ldots, N], \tag{2}$$

the system's standalone probabilities at $t \neq 0$ can be denoted by

$$p_i(t) \quad [\text{for } i = 1, \ldots, N], \tag{3}$$

and the stochastic map consists of *conditional probabilities*

$$\Gamma_{ij}(t) \equiv p(i, t|j, 0) \quad [\text{for } i, j = 1, \ldots, N], \tag{4}$$

where $p(i,t|j,0)$ denotes the conditional probability for the system to be in its ith configuration at the time t, given that the system is in its jth configuration at the initial time 0. (Note that no assumption is made here about whether $t > 0$ or $t < 0$.) Being probabilities, these quantities satisfy the usual non-negativity and normalization conditions

$$p_j(0), p_i(t) \geq 0, \quad \sum_{j=1}^{N} p_j(0) = \sum_{i=1}^{N} p_i(t) = 1, \tag{5}$$

and

$$\Gamma_{ij}(t) \geq 0, \qquad \sum_{i=1}^{N} \Gamma_{ij}(t) = 1. \tag{6}$$

Then from *Bayesian marginalization*, one has the *linear* relationship

$$p_i(t) = \sum_{j=1}^{N} \Gamma_{ij}(t) p_j(0), \tag{7}$$

where the initial standalone probabilities $p_j(0)$ are assumed to be arbitrary and can therefore be freely adjusted without altering the conditional probabilities $\Gamma_{ij}(t)$.

Letting $p(0)$ denote an $N \times 1$ *probability vector* whose entries are given by the standalone probabilities $p_j(0)$, letting $p(t)$ denote the analogous $N \times 1$ probability vector with entries given by $p_i(t)$, and letting $\Gamma(t)$ denote the $N \times N$ time-dependent *transition matrix* consisting of the conditional probabilities $\Gamma_{ij}(t)$, one can naturally recast the linear Bayesian marginalization relationship (7) in matrix form as

$$p(t) = \Gamma(t) p(0). \tag{8}$$

The conditions (6) on the transition matrix $\Gamma(t)$ identify it, mathematically speaking, as a *(column) stochastic matrix*. On physical grounds, $\Gamma(t)$ will be assumed to satisfy the continuity condition that in the limit $t \to 0$, it approaches its initial value $\Gamma(0)$, which will be taken to be the $N \times N$ identity matrix $\mathbb{1}$:

$$\lim_{t \to 0} \Gamma(t) = \Gamma(0) \equiv \mathbb{1} \equiv \mathrm{diag}(1, \ldots, 1). \tag{9}$$

Next, consider a *random variable* $A(t)$ with (not necessarily distinct) real-valued *magnitudes* $a_1(t), \ldots, a_N(t)$ determined by the system's configuration $i = 1, \ldots, N$, and possibly also depending explicitly on the time t. Then the *expectation value* $\langle A(t) \rangle$ is defined as the statistical average or mean of the magnitudes of $A(t)$ over the system's standalone probability distribution at t:

$$\langle A(t) \rangle \equiv \sum_{i=1}^{N} a_i(t) p_i(t) = \sum_{i=1}^{N} \sum_{j=1}^{N} a_i(t) \Gamma_{ij}(t) p_j(0). \tag{10}$$

One can go on to define the standard deviation and various statistical moments of $A(t)$ by appropriate generalizations of this basic definition.

All these formulas can be extended to systems with continuous configuration spaces. For a system with a continuous configuration space $\mathcal{C}$, one uses standalone probability densities $p(y, 0)$ at the initial time 0 and $p(x, t)$ at $t \neq 0$, where x and y each symbolically denotes a set of real-valued coordinates. The Bayesian marginalization relationship (7) then becomes

$$p(x,t) = \int_{\mathcal{C}} d\mu(y)\, \Gamma(x,y,t) p(y,0), \tag{11}$$

where $d\mu(y)$ is a suitable integral measure over $\mathcal{C}$ and where the conditional probability density $\Gamma(x, y, t)$ naturally serves as an integral kernel. A random variable $A(t)$ then has magnitudes $a(x, t)$ labeled by x and t, and its expectation value (10) becomes

$$\left.\begin{aligned} \langle A(t) \rangle &\equiv \int_{\mathcal{C}} d\mu(x)\, a(x,t) p(x,t) \\ &= \int_{\mathcal{C}} d\mu(x) \int_{\mathcal{C}} d\mu(y)\, a(x,t) \Gamma(x,y,t) p(y,0). \end{aligned}\right\} \tag{12}$$

For ease of exposition, the discrete case will be assumed going forward.

Equations like (7), (9), and (11) may appear to single out the initial time 0 as a special time. Subsection IV F, however, will show that for systems in sufficiently strong contact with a repeatedly eavesdropping environment, the initial time 0 need not actually be a unique time, but will typically be only one of many times that play a similar role.

B. Traditional Approximations

In textbook treatments of stochastic processes, one often introduces various approximations or simplifications in defining a system's time-dependent transition matrix $\Gamma(t)$ to make it easier to construct and describe. A typical such approximation is to assume that the system is a *discrete-time Markov chain*, meaning that for some small but finite time scale δt, one can express the time-dependent transition matrix $\Gamma(n\,\delta t)$ at any integer number $n \geq 1$ of steps of duration δt as n powers of a constant stochastic matrix Γ:

$$\Gamma(n\,\delta t) = \Gamma^n. \tag{13}$$

Somewhat more generally, a convenient simplification is to assume that for any two times t and t' satisfying $t > t' > 0$, one has the composition law

$$\Gamma(t) = \Gamma\big(t \leftarrow t'\big)\Gamma\big(t'\big). \tag{14}$$

This simplification is known as *divisibility* [28], a term that seems to have first appeared in the research literature in a 2008 paper [29] by Wolf and Cirac in the context of quantum channels.[2]

[2] Note that this notion of divisibility is unrelated to the much older concept of *infinite divisibility*, which refers to

Here $\Gamma(t \leftarrow t')$ is likewise required to satisfy the mathematical requirements of being a stochastic matrix, in the sense that its entries are all non-negative and its columns each sum to 1, as in (6).

An even more special simplification is to take the transition matrix $\Gamma(t)$ to be a time-dependent *permutation matrix*, meaning a matrix whose columns are a permutation of the columns of the $N \times N$ identity matrix $\mathbb{1}$. In that case, $\Gamma(t)$ contains only 1s and 0s, so it does not contain nontrivial probabilities at all, and the system transitions deterministically from one configuration to another in its configuration space $\mathcal{C}$. In a suitable continuum limit $N \to \infty$, the time evolution reduces to smooth, deterministic dynamics.

Absent these sorts of approximations or simplifications, one is confronted with the task of constructing an $N \times N$ time-dependent, generically non-Markovian, *indivisible* transition matrix $\Gamma(t)$ for a given configuration space $\mathcal{C}$, ideally in a systematic way. For small configuration spaces, it is easy to devise smoothly time-dependent, non-Markovian, indivisible examples, like the 2×2 transition matrix

$$\Gamma(t) \equiv \begin{pmatrix} e^{-t^2/\tau^2} & 1 - e^{-t^2/\tau^2} \\ 1 - e^{-t^2/\tau^2} & e^{-t^2/\tau^2} \end{pmatrix}, \tag{15}$$

where τ is a constant with units of time, or

$$\Gamma(t) \equiv \begin{pmatrix} \cos^2 \omega t & \sin^2 \omega t \\ \sin^2 \omega t & \cos^2 \omega t \end{pmatrix}, \tag{16}$$

where ω is a constant with units of inverse-time. However, it may not seem obvious how to construct smoothly time-dependent transition matrices $\Gamma(t)$ systematically beyond the 2×2 case, especially in the case of large ($N \gg 1$) configuration spaces.

A sufficiently general approach for accomplishing this task could have numerous practical applications in many scientific and technical fields. This paper will develop one such approach, and use it to provide a self-contained theoretical justification for why the Markov and divisibility approximations work so well in many real-world cases.

III. THE HILBERT-SPACE FORMULATION

A. The Schur-Hadamard Factorization

One of the goals of this paper will be to introduce a new and highly general framework for formulating time-dependent transition matrices $\Gamma(t)$, conceptually akin to the Lagrangian or Hamiltonian frameworks for formulating deterministic dynamics for mechanical systems.

The starting place will be to 'solve' the non-negativity condition $\Gamma_{ij}(t) \geq 0$ on the individual entries of the transition matrix $\Gamma(t)$ by expressing them in the following way:

$$\Gamma_{ij}(t) = |\Theta_{ij}(t)|^2. \tag{17}$$

---

a probability distribution that can be expressed as the probability distribution of a sum of any integer number of independent and identically distributed random variables.

This equation is not a postulate—it is a mathematical identity.

The $N \times N$ matrix $\Theta(t)$ introduced in (17) is guaranteed to exist, although it is not unique. Its entries $\Theta_{ij}(t)$ could be taken to be the real square roots of the corresponding quantities $\Gamma_{ij}(t)$, but they could also include complex numbers, quaternions, or even the elements of a more general normed algebra (although associativity is a very helpful property to require). To keep things simple, and with the eventual goal of reproducing the usual formalism of quantum theory, this paper will choose $\Theta_{ij}(t)$ to involve only at most the complex numbers.

On account of the normalization condition on the transition matrix $\Gamma(t)$ specified in (6), note that the matrix $\Theta(t)$ must satisfy the summation condition

$$\sum_{i=1}^{N} |\Theta_{ij}(t)|^2 = 1. \tag{18}$$

For now, no further conditions, such as unitarity, will be imposed on $\Theta(t)$, whose significance will soon become more clear.

There are several helpful ways to re-express the identity (17). To begin, one can introduce the *Schur-Hadamard product* $\odot$, which is defined for arbitrary $N \times N$ matrices X and Y as entry-wise multiplication [30–32]:

$$(X \odot Y)_{ij} \equiv X_{ij} Y_{ij}. \tag{19}$$

One can then regard (17) as expressing the transition matrix $\Gamma(t)$ as a *Schur-Hadamard factorization* of the complex-conjugated matrix $\overline{\Theta(t)}$ with $\Theta(t)$ itself:

$$\Gamma(t) = \overline{\Theta(t)} \odot \Theta(t). \tag{20}$$

Schur-Hadamard products are not widely used in linear algebra, in part because they are basis-dependent. For the purposes of analyzing a given generalized stochastic system, however, this basis-dependence is unimportant, because the system's configuration space $\mathcal{C}$ naturally singles out a specific basis, to be defined momentarily.

B. The Dictionary

As an alternative approach that will turn out to have significant ramifications, one starts by defining an N-member collection of $N \times N$ constant, diagonal projection matrices $P_1, \ldots, P_N$, which will be called *configuration projectors*. For each $i = 1, \ldots, N$, the configuration projector P_i consists of a single 1 in its ith row, ith column, and 0s in all its other entries. That is, P_i is defined as

$$P_i \equiv \mathrm{diag}(0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{th entry}}}{1}, 0, \ldots, 0), \tag{21}$$

with individual entries

$$P_{i,jk} = \delta_{ij}\delta_{ik}, \tag{22}$$

where δ_{ij} is the usual Kronecker delta:

$$\delta_{ij} \equiv \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases} \tag{23}$$

It follows immediately that the configuration projectors satisfy the conditions of *mutual exclusivity*,

$$P_i P_j = \delta_{ij} P_i, \tag{24}$$

and *completeness*,

$$\sum_{i=1}^{N} P_i = \mathbb{1}, \tag{25}$$

where again $\mathbb{1}$ is the $N \times N$ identity matrix. These two conditions are the defining features of a *projection-valued measure (PVM)* [33, 34], so the configuration projectors $P_1, \ldots, P_N$ naturally constitute a PVM.

Letting $\mathrm{tr}(\cdots)$ denote the usual matrix trace, one can then recast the mathematical identity (17) relating the entries of $\Gamma(t)$ with the entries of $\Theta(t)$ as

$$\boxed{\Gamma_{ij}(t) = \mathrm{tr}(\Theta^{\dagger}(t) P_i \Theta(t) P_j).} \tag{26}$$

This equation is a new result. It will turn out to serve as an important *dictionary* that translates between the formalism of generalized stochastic processes, as symbolized by $\Gamma_{ij}(t)$ on the left-hand side, and an expansive set of mathematical tools for constructing stochastic dynamics, as embodied by the right-hand side.[3]

C. The Hilbert Space

To understand what these mathematical tools are, the next step will be to introduce a set of $N \times 1$ column vectors $e_1, \ldots, e_N$, where e_i has a 1 in its ith component and 0s in all its other components:

$$e_1 \equiv \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \ldots, \quad e_N \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{27}$$

[3] Similar-looking formulas appear incidentally in equations (3)–(6) of [35] as an intermediate step in proving a lemma that the authors use for conceptually different purposes.

That is, e_i has components

$$e_{i,j} = \delta_{ij}. \tag{28}$$

It follows that the column vectors $e_1, \ldots, e_N$ form an orthonormal basis for the vector space of all $N \times 1$ column vectors, and $e_1, \ldots, e_N$ will be called the system's *configuration basis*. In particular,

$$e_i^\dagger e_j = \delta_{ij}, \quad e_i e_i^\dagger = P_i, \tag{29}$$

where P_i is the ith configuration projector, as defined in (21).

Hence, the right-hand side of the dictionary (26) is a trace over a *Hilbert space*, meaning a complete inner-product space over the complex numbers. More explicitly, the dictionary picks out a Hilbert space $\mathcal{H}$ that is isomorphic to the vector space $\mathbb{C}^N$ of $N \times 1$ column vectors with complex-valued entries, under the inner product $v^\dagger w$:

$$\mathcal{H} \cong \mathbb{C}^N. \tag{30}$$

The dictionary therefore provides a Hilbert-space formulation for constructing highly generic forms of stochastic dynamics.

Substituting the right-hand side of the dictionary (26) into the linear Bayesian marginalization relationship (7) between the system's standalone probabilities $p_j(0)$ at the initial time 0 and the standalone probabilities $p_i(t)$ at $t \neq 0$, one finds that

$$p_i(t) = \mathrm{tr}(P_i \rho(t)), \tag{31}$$

where $\rho(t)$ is an $N \times N$ time-dependent, self-adjoint, unit-trace, generically non-diagonal matrix defined as

$$\left.\begin{aligned} \rho(t) &\equiv \Theta(t) \left[\sum_{j=1}^{N} p_j(0) P_j \right] \Theta^\dagger(t) \\ &= \Theta(t) \mathrm{diag}(\ldots, p_j(0), \ldots) \Theta^\dagger(t), \\ &\qquad \text{with} \quad \rho^\dagger(t) = \rho(t) \\ &\qquad \text{and} \quad \mathrm{tr}(\rho(t)) = 1. \end{aligned}\right\} \tag{32}$$

Crucially, notice how the linearity of the Bayesian marginalization relationship (7) is ultimately responsible for the linearity of the relationship between the matrix $\rho(t)$ and its value $\rho(0)$ at the initial time 0, as expressed in (32).

Similarly, by substituting the formula (31) for $p_i(t)$ into the definition (10) of the expectation value of a random variable $A(t)$, one obtains

$$\langle A(t) \rangle = \mathrm{tr}(A(t) \rho(t)). \tag{33}$$

Here $A(t)$ is now understood to be the $N \times N$ time-dependent, diagonal matrix whose entries are the random variable's individual magnitudes $a_1(t), \dots, a_N(t)$:

$$A(t) \equiv \sum_{i=1}^{N} a_i(t) P_i = \mathrm{diag}(\dots, a_i(t), \dots). \tag{34}$$

In the special case in which the system's standalone probability distribution at the initial time 0 is pure, meaning that one of the system's configurations j is occupied with probability 1, the system's probability vector at the initial time 0 is equal to the jth vector e_j in the configuration basis (28):

$$p(0) = e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \leftarrow j\text{th entry} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad [\text{pure}]. \tag{35}$$

One can then define an $N \times 1$ column vector

$$\Psi(t) \equiv \Theta(t) e_j, \tag{36}$$

which is ultimately just the jth column of $\Theta(t)$. Due to the summation condition (18) on $\Theta(t)$, this column vector $\Psi(t)$ automatically has unit unit norm, in the sense that

$$\Psi^{\dagger}(t)\Psi(t) = 1. \tag{37}$$

Moreover, the ith component $\Psi_i(t)$ of $\Psi(t)$ is equal to the specific complex-valued matrix entry $\Theta_{ij}(t)$:

$$\Psi_i(t) = \Theta_{ij}(t). \tag{38}$$

This component $\Psi_i(t)$ is a purely law-like quantity, in the sense of being just another name for a part of $\Theta(t)$, which is itself just a way of encoding the system's dynamical law, as embodied by the system's transition matrix $\Gamma(t)$.

It follows from a short calculation that when the purity condition (35) holds at the initial time 0, the self-adjoint matrix $\rho(t)$ defined in (32) is rank-one and has factorization

$$\rho(t) = \Psi(t)\Psi^{\dagger}(t) \quad [\text{pure}]. \tag{39}$$

The probability formula (31) then simplifies to

$$p_i(t) = |\Psi_i(t)|^2, \tag{40}$$

and the formula (33) for the expectation value of a random variable $A(t)$ becomes

$$\langle A(t)\rangle = \Psi^\dagger(t)A(t)\Psi(t). \tag{41}$$

Looking at all these results, one notices a striking resemblance to mathematical objects and formulas that are familiar from textbook quantum theory. Specifically, one sees that $\Theta(t)$ plays the role of a *time-evolution operator*, $\rho(t)$ is a *density matrix,* $\Psi(t)$ is a *state vector* or *wave function*, and $A(t)$ represents an *observable.*[4] The probability formulas (31) and (40) have the same form as the *Born rule*, and (33) and (41) have the same form as the standard expressions for quantum expectation values. (For pedagogical treatments of quantum theory, see [36–40].)

These formulas are all expressed in what would conventionally be called the *Schrödinger picture.* It will also end up being useful to introduce the *Heisenberg picture*, which is defined according to

$$\left.\begin{aligned} \rho^H \equiv \rho(0), \quad \Psi^H \equiv \Psi(0), \\ A^H(t) \equiv \Theta^\dagger(t)A(t)\Theta(t), \end{aligned}\right\} \tag{42}$$

where $A^H(t)$ now includes both a possible *explicit* dependence on time through its magnitudes $a_i(t)$ as well as an *implicit* dependence on time through the time-evolution operator $\Theta(t)$. In the Heisenberg picture, the probability formula (31) becomes[5]

$$p_i(t) = \mathrm{tr}(P^H(t)\rho^H), \tag{43}$$

and the formula (33) for expectation values becomes

$$\langle A(t)\rangle = \mathrm{tr}(A^H(t)\rho^H). \tag{44}$$

Despite the similarity to expressions found in quantum theory, as well as the appearance of non-diagonal matrices, it is important to keep in mind that the system under investigation here is always fundamentally in a specific configuration $i = 1, \ldots, N$ in its configuration space $\mathcal{C}$ at any given time, and that the system's dynamics is completely captured by the transition matrix $\Gamma(t)$, whose entries are conditional probabilities $p(i, t|j, 0)$, in accordance with (4). The mathematical objects $\Theta(t)$, $\rho(t)$, $\Psi(t)$, and $A(t)$, despite being extremely useful, are difficult to assign direct physical meanings, in part because they are not uniquely defined by $\mathcal{C}$ or $\Gamma(t)$.

D. Gauge Transformations

To make this non-uniqueness more manifest, it will be helpful to introduce an analogy with the *Maxwell theory of classical electromagnetism.* (For pedagogical treatments of classical electromagnetism, see [41–43].)

[4] Note that for the purposes of this paper, the terms 'operator' and 'matrix' will be used interchangeably, as will the terms 'state vector' and 'wave function.'

[5] Note that for a generic time-evolution operator $\Theta(t)$, the Heisenberg-picture version $P_i^H(t) \equiv \Theta^\dagger(t)P_i\Theta(t)$ of a projector P_i will not likewise be a projector.

In classical electromagnetism, the electric and magnetic fields are physically meaningful quantities, but it is often very convenient to work instead in terms of scalar and vector potentials, which are not uniquely defined. All choices for the potentials that yield the same electric and magnetic fields are said to be related by *gauge transformations*. Any one such choice for the potentials is called a *gauge choice*, and the scalar and vector potentials themselves are called *gauge potentials* or *gauge variables*.

Making a suitable gauge choice can greatly simplify many calculations, such as using Lorenz gauge to compute the electric and magnetic fields for delayed boundary conditions. Ultimately, however, no gauge choice is fundamentally more physically correct than any other gauge choice, and all calculations of physical predictions in classical electromagnetism must conclude with expressions that are written in terms of gauge-invariant quantities.

To set up the claimed analogy with electromagnetic gauge transformations, one starts by observing from the basic relationship $\Gamma_{ij}(t) = |\Theta_{ij}(t)|^2$ in (17) that the Schur-Hadamard product (19) of the time-evolution operator $\Theta(t)$ and a matrix of time-dependent phases $\exp(i\theta_{ij}(t))$ is a transformation of $\Theta(t)$ with no physical effects, and therefore corresponds to a genuine form of gauge invariance:

$$\Theta(t) \mapsto \Theta(t) \odot \begin{pmatrix} e^{i\theta_{11}(t)} & e^{i\theta_{12}(t)} & & \\ e^{i\theta_{21}(t)} & \ddots & & \\ & & & e^{i\theta_{NN}(t)} \end{pmatrix}. \tag{45}$$

This gauge transformation can be written equivalently at the level of individual matrix entries as

$$\Theta_{ij}(t) \mapsto \Theta_{ij}(t) e^{i\theta_{ij}(t)}. \tag{46}$$

To the author's knowledge, this kind of gauge invariance, which will be called a *Schur-Hadamard gauge transformation*, has not yet been described in the research literature, and is therefore a new result. It will turn out to play a key role in the analysis of dynamical symmetries that will be presented in Subsection VI A, and will be extended in an interesting way in the context of Hilbert-space dilations in Subsection VI B.

The Hilbert-space formulation has another form of gauge invariance, which appears to have first been written down in [44] in the context of transformations of the Schrödinger equation between inertial and non-inertial reference frames. Letting $V(t)$ be a time-dependent unitary matrix, the following transformation is also a gauge invariance of the Hilbert-space formulation, leaving all probabilities $p_i(t)$, expectation values $\langle A(t) \rangle$, and the transition matrix $\Gamma(t)$ as a whole unchanged:[6]

$$\left.\begin{aligned} \rho(t) &\mapsto \rho_V(t) \equiv V(t)\rho(t)V^\dagger(t), \\ \Psi(t) &\mapsto \Psi_V(t) \equiv V(t)\Psi(t), \\ A(t) &\mapsto A_V(t) \equiv V(t)A(t)V^\dagger(t), \\ \Theta(t) &\mapsto \Theta_V(t) \equiv V(t)\Theta(t)V^\dagger(0). \end{aligned}\right\} \tag{47}$$

[6] One should be mindful of the appearance of the initial time 0 in $V^\dagger(0)$ in the transformation rule for $\Theta(t)$.

If the unitary matrix $V(t)$ is taken to be time-*independent*, then the gauge transformation (47) is merely a change of orthonormal basis. However, if $V(t)$ depends nontrivially on time, and if one regards the system's Hilbert space at each moment in time as a *fiber* over a one-dimensional *base manifold* parameterized by the time coordinate t, then $V(t)$ represents a local-in-time, unitary transformation of each individual Hilbert-space fiber. In particular, any given time-dependent state vector $\Psi(t)$, regarded as a trajectory through the Hilbert space $\mathcal{H}$, can be mapped to any other trajectory by a suitable choice of time-dependent unitary matrix $V(t)$, so trajectories in $\mathcal{H}$ do not describe gauge-invariant facts.

IV. THE STOCHASTIC-QUANTUM CORRESPONDENCE

A. Kraus Decompositions

In the most general case, a time-evolution operator $\Theta(t)$ may not satisfy any nontrivial constraints apart from the summation condition (18). It will turn out to be helpful to find alternative ways of representing the $N \times N$ matrix $\Theta(t)$ in terms of more tightly constrained mathematical objects.

For $\beta = 1, \ldots, N$, let $K_\beta(t)$ be the $N \times N$ matrix defined to share its βth column with $\Theta(t)$, but with 0s in all its other entries:

$$K_\beta(t) \equiv \begin{pmatrix} 0 & \cdots & 0 & \Theta_{1\beta}(t) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \Theta_{N\beta}(t) & 0 & \cdots & 0 \end{pmatrix} \quad [\beta = 1, \ldots, N]. \tag{48}$$

The entries of $K_\beta(t)$ are given explicitly by

$$K_{\beta,ij}(t) = \delta_{\beta j}\Theta_{ij}(t). \tag{49}$$

Then the summation condition (18) on $\Theta(t)$ becomes the statement that the matrices $K_1(t), \ldots, K_N(t)$ satisfy the *Kraus identity*:

$$\sum_{\beta=1}^{N} K_\beta^\dagger(t) K_\beta(t) = \mathbb{1}. \tag{50}$$

These matrices are therefore called *Kraus operators* [45]. One can then write the dictionary (26) in an alternative form called a *Kraus decomposition*:[7]

$$\Gamma_{ij}(t) = \sum_{\beta=1}^{N} \mathrm{tr}(K_\beta^\dagger(t) P_i K_\beta(t) P_j). \tag{51}$$

Like all the other mathematical objects in the Hilbert-space formulation, the Kraus operators

[7] Conditional probabilities similar in form to (51) were studied in [46].

$K_1(t), \ldots, K_N(t)$ are not unique. Notice also that any number of $N \times N$ matrices satisfying the Kraus identity (50) are guaranteed to yield a valid transition matrix $\Gamma(t)$ via the Kraus decomposition (51). Said in another way, the preceding argument establishes the existence but not the uniqueness of Kraus operators for any given time-evolution operator $\Theta(t)$.

Kraus operators and Kraus decompositions play an important role in quantum information theory. In particular, they provide (non-unique) expressions for generalizations of unitary time evolution known as *quantum channels*, or *completely positive trace-preserving (CPTP) maps*.

B. Unistochastic Matrices

In general, Kraus operators need not have the specific form (48), and there need not be N of them. In the most minimal case in which a system's transition matrix $\Gamma(t)$ is determined by just a single Kraus operator $K_1(t)$, that Kraus operator will be denoted instead by $U(t)$. In that case, the general Schur-Hadamard factorization (20) specializes to

$$\Gamma(t) = \overline{U(t)} \odot U(t). \tag{52}$$

That is,

$$\Gamma_{ij}(t) = |U_{ij}(t)|^2. \tag{53}$$

Equivalently, in dictionary form (26), one has

$$\Gamma_{ij}(t) = \mathrm{tr}(U^\dagger(t) P_i U(t) P_j). \tag{54}$$

The Kraus identity (50), meanwhile, reduces to the statement that $U(t)$ is unitary,

$$U^\dagger(t) = U^{-1}(t), \tag{55}$$

and the system's transition matrix $\Gamma(t)$ is then said to be a *unistochastic matrix*. That is, a unistochastic matrix is a square matrix whose individual entries are the modulus-squares of the corresponding entries of a unitary matrix.

Note that a unitary time-evolution operator $U(t)$ will not generically remain unitary under arbitrary Schur-Hadamard gauge transformations (45). Hence, writing a unistochastic transition matrix $\Gamma(t)$ in terms of a unitary time-evolution operator $U(t)$ corresponds to making a gauge choice—or, somewhat more precisely, to partially fixing the gauge freedom (45).

Unistochastic matrices were first introduced in 1954 by Horn [47], who originally called them 'ortho-stochastic' matrices. The modern term 'unistochastic matrix' was introduced by Thompson in 1989 [48, 49]. The term *orthostochastic matrix* now refers to a square matrix whose entries are the modulus-squares of the corresponding entries of a *real orthogonal* matrix.

Every orthostochastic matrix is unistochastic. Importantly, however, the reverse statement is not generally true, meaning that the complex numbers generically play a necessary role in

formulating a unistochastic transition matrix $\Gamma(t)$ in terms of a unitary time-evolution operator $U(t)$. Even when the complex numbers are not strictly necessary for writing down a unitary time-evolution operator $U(t)$, such as if the time-evolution operator can be taken to be real and orthogonal, it is still very convenient to employ the complex numbers for a given Hilbert-space representation, so that one can take advantage of the many highly useful constructs that show up in standard treatments of quantum theory.

It follows immediately from the dictionary formula (54) that every unistochastic transition matrix is *doubly stochastic*, or *bistochastic*, which means that summing over any of its rows *or* any of its columns always yields 1:

$$\sum_{i=1}^{N} \Gamma_{ij}(t) = \sum_{j=1}^{N} \Gamma_{ij}(t) = 1. \tag{56}$$

C. Unistochastic Systems

A generalized stochastic system whose transition matrix $\Gamma(t)$ is a unistochastic matrix will be called a *unistochastic system*.

To provide a simple example, note that every permutation matrix is, in particular, a unitary matrix. Moreover, because the entries of a permutation matrix are all 1s and 0s, they are individually invariant when one computes their modulus-squares, so every permutation matrix is *also* a unistochastic matrix. It follows that a discrete, deterministic system whose dynamics is defined by a permutation matrix is a special case of a unistochastic system.

Remarkably, as will be shown in Subsection VI B, *every* stochastic map can be expressed in terms of a unitary time-evolution operator on a suitably enlarged or *dilated* Hilbert space. As a consequence, *every* generalized stochastic system can be regarded as a subsystem of a unistochastic system. This statement can be formalized as a theorem, called the *stochastic-quantum theorem* [50], and implies that the study of generalized stochastic systems can essentially be reduced to the study of unistochastic systems. Hence, assuming a unistochastic system is not as special a condition as it might initially seem.

Assuming a unistochastic system based on a unitary time-evolution operator $U(t)$ that is a differentiable function of the time t, one can define a corresponding self-adjoint generator $H(t)$, called the system's *Hamiltonian*, according to

$$H(t) \equiv i\hbar \frac{\partial U(t)}{\partial t} U^{\dagger}(t) = H^{\dagger}(t). \tag{57}$$

Here the factor of i ensures that the $N \times N$ matrix $H(t)$ is self-adjoint, and, for present purposes, the *reduced Planck constant* $\hbar$ is a fixed quantity introduced for purposes of measurement units. Ultimately, the specific numerical value of $\hbar$ in any given set of units must be determined empirically by comparison with experiments.

In terms of the Hamiltonian, the system's density matrix $\rho(t)$ then evolves in time according to

the *von Neumann equation*,

$$i\hbar\frac{\partial\rho(t)}{\partial t} = [H(t), \rho(t)], \tag{58}$$

its state vector $\Psi(t)$ (if it exists) evolves according to the *Schrödinger equation*,

$$i\hbar\frac{\partial\Psi(t)}{\partial t} = H(t)\Psi(t), \tag{59}$$

its Heisenberg-picture random variables $A^H(t)$ evolve according to the *Heisenberg equation of motion*,

$$\frac{dA^H(t)}{dt} = \frac{i}{\hbar}[H^H(t), A^H(t)] + \left(\frac{\partial A(t)}{\partial t}\right)^H, \tag{60}$$

and its expectation values $\langle A(t)\rangle$ evolve according to the *Ehrenfest equation*,

$$\frac{d\langle A(t)\rangle}{dt} = \frac{i}{\hbar}\mathrm{tr}([H(t), A(t)]\rho(t)) + \left\langle\frac{\partial A(t)}{\partial t}\right\rangle. \tag{61}$$

The matrix $H^H(t)$ appearing in the Heisenberg equation of motion (60) is the Hamiltonian expressed in the Heisenberg picture (42). Note also that the brackets $[X, Y]$ that naturally show up in these equations are genuine *commutators* $XY - YX$, not Poisson brackets, and involve products of non-diagonal matrices that do not generally commute with each other under multiplication.

The emergence of these famous equations from a physical model based on a stochastically evolving trajectory in a configuration space $\mathcal{C}$ is a surprising new result.

Intriguingly, if the system's time-evolution operator $\Theta(t) = U(t)$ is indeed unitary, then under the unitary gauge transformation defined by (47), the Hamiltonian transforms precisely as a *non-Abelian gauge potential*:

$$\left.\begin{aligned} H(t) &\mapsto H_V(t) \\ &= V(t)H(t)V^\dagger(t) - i\hbar V(t)\frac{\partial V^\dagger(t)}{\partial t}. \end{aligned}\right\} \tag{62}$$

This transformation behavior makes clear that a Hamiltonian is not a gauge-invariant observable, even though it may happen to coincide with various observables according to particular gauge choices. (For pedagogical treatments of non-Abelian gauge theories, see [51, 52].)

Furthermore, one can write the Schrödinger equation (59) as

$$\mathcal{D}(t)\Psi(t) = 0. \tag{63}$$

Here $\mathcal{D}(t)$ is a *gauge-covariant derivative* defined according to

$$\mathcal{D}(t) \equiv \mathbb{1}\frac{\partial}{\partial t} + \frac{i}{\hbar}H(t), \tag{64}$$

which maintains its form under the unitary gauge transformations (47), in the sense that

$$V(t)\left[\mathbb{1}\frac{\partial}{\partial t}+\frac{i}{\hbar}H(t)\right](\cdots)=\left[\mathbb{1}\frac{\partial}{\partial t}+\frac{i}{\hbar}H_V(t)\right][V(t)(\cdots)]. \tag{65}$$

These formulas make manifest that the Hilbert-space formulation of a generalized stochastic system is ultimately a collection of gauge-dependent quantities, or gauge variables. Hence, although a Hilbert-space formulation may be extremely useful for constructing stochastic dynamics or for carrying out calculations, one might rightly be suspicious about trying to assign direct physical meanings to its mathematical ingredients.

Notice that if one picks the gauge-transformation matrix $V(t)$ to be the adjoint of the unistochastic system's time-evolution operator $U(t)$,

$$V(t)\equiv U^{\dagger}(t), \tag{66}$$

then the Hamiltonian precisely vanishes:

$$H_V(t)=0. \tag{67}$$

This choice of gauge is nothing other than the definition (42) of the Heisenberg picture. Unitary gauge transformations (47) can therefore be viewed as generalized changes of time-evolution picture.[8]

D. Interference

The appearance of the Schrödinger equation (59) is an important signal that the dictionary (26) is more than just a tool for using Hilbert-space methods to craft highly general forms of stochastic dynamics. It also suggests that generalized stochastic systems might have the resources to replicate the features of quantum theory more broadly.

As another hint pointing in this direction, one starts by noting that an arbitrary time-dependent transition matrix $\Gamma(t)$ is generically *indivisible*, in the sense that it does not satisfy the divisibility property (14) at arbitrary times. To see what goes wrong with divisibility, suppose that at some time t', the transition matrix $\Gamma(t')$ has a matrix inverse $\Gamma^{-1}(t')$, and define a new $N\times N$ matrix $\tilde{\Gamma}(t\leftarrow t')$ according to

$$\tilde{\Gamma}\big(t\leftarrow t'\big)\equiv\Gamma(t)\Gamma^{-1}\big(t'\big). \tag{68}$$

As an immediate consequence, one then has

$$\Gamma(t)=\tilde{\Gamma}\big(t\leftarrow t'\big)\Gamma\big(t'\big), \tag{69}$$

[8] The fact that one can set $H_V(t)=0$ for all t is a manifestation of the fact that the *fiber bundle* in this case, consisting of copies of the system's Hilbert space fibered over a one-dimensional base manifold parameterized by the time t, has vanishing curvature.

which resembles the divisibility property (14). However, it follows from an elementary theorem of linear algebra that the inverse of a stochastic matrix can only be stochastic if both matrices are permutation matrices, and therefore do not involve nontrivial probabilities.[9] Hence, the matrix $\tilde{\Gamma}(t \leftarrow t')$ defined in (68) is not generically stochastic, so (69) does not express a genuine form of divisibility.

There is an alternative—and far-reaching—way to understand the generic indivisibility of a time-dependent transition matrix $\Gamma(t)$. To this end, suppose that $\Gamma(t)$ happens to be unistochastic, and let $U(t)$ be a unitary time-evolution operator for $\Gamma(t)$. Then for any two times t and t', one can define a *relative* time-evolution operator

$$U\big(t \leftarrow t'\big) \equiv U(t)U^{\dagger}\big(t'\big), \tag{70}$$

which yields the composition law

$$U(t) = U\big(t \leftarrow t'\big)U\big(t'\big). \tag{71}$$

At the level of the unistochastic transition matrix $\Gamma(t)$, one has from the Schur-Hadamard factorization (52) that

$$\left.\begin{aligned} \Gamma(t) &= \overline{U(t)} \odot U(t) \\ &= \overline{[U(t \leftarrow t')U(t')]} \odot \big[U\big(t \leftarrow t'\big)U\big(t'\big)\big], \end{aligned}\right\} \tag{72}$$

which cannot generally be expressed in the form $\Gamma(t \leftarrow t')\Gamma(t')$ for any stochastic matrix $\Gamma(t \leftarrow t')$, due to the presence of cross terms.

Indeed, examining individual matrix entries, one finds more explicitly that

$$\left.\begin{aligned} \Gamma_{ij}(t) = &\sum_{k=1}^{N} |U_{ik}\big(t \leftarrow t'\big)|^2 |U_{kj}\big(t'\big)|^2 \\ &+ \sum_{k \neq l} \overline{U_{ik}(t \leftarrow t')U_{kj}(t')} U_{il}\big(t \leftarrow t'\big)U_{lj}\big(t'\big). \end{aligned}\right\} \tag{73}$$

With $\Gamma_{kj}(t')$ defined according to (53) as usual,

$$\Gamma_{kj}\big(t'\big) = |U_{kj}\big(t'\big)|^2, \tag{74}$$

and defining

$$\Gamma_{ik}\big(t \leftarrow t'\big) \equiv |U_{ik}\big(t \leftarrow t'\big)|^2, \tag{75}$$

[9] Proof: Let X and Y be $N \times N$ matrices with only non-negative entries and with $Y = X^{-1}$, so that $XY = \mathbb{1}$. Then, in particular, the first row of X must be orthogonal to the second-through-Nth columns of Y. Because Y is invertible, the columns of Y must all be linearly independent, so the first row of X must be orthogonal to the $(N-1)$-dimensional subspace spanned by the second-through-Nth columns of Y. Because the entries of X and Y are all non-negative by assumption, the only way that this orthogonality condition can hold is if precisely one of the entries in the first row of X is nonzero, with a 0 in the corresponding entry in each of the second-through-Nth columns of Y. Repeating this argument for the other rows of X, one sees that X can only have a single nonzero entry in each row. If X is a stochastic matrix, then each of these nonzero entries must be the number 1, so X must be a permutation matrix. Because the inverse of a permutation matrix is again a permutation matrix, it follows that Y must likewise be a permutation matrix. QED

which is manifestly unistochastic, one sees that the discrepancy between $\Gamma(t)$ and its would-be division $\Gamma(t \leftarrow t')\Gamma(t')$ is given by

$$\left.\begin{aligned} &\Gamma_{ij}(t) - \big[\Gamma\big(t \leftarrow t'\big)\Gamma\big(t'\big)\big]_{ij} \\ &= \sum_{k \neq l} \overline{U_{ik}(t \leftarrow t')\Psi_k(t')} U_{il}\big(t \leftarrow t'\big)\Psi_l\big(t'\big), \end{aligned}\right\} \tag{76}$$

where $\Psi(t') \equiv \Theta(t')e_j$ is the system's state vector at the time t', in keeping with the general definition of state vectors in (36).

Remarkably, the right-hand side of (76) gives the general mathematical formula for quantum interference, despite the absence of manifestly quantum-theoretic assumptions. One sees from this analysis that interference is a direct consequence of stochastic dynamics not generally being divisible. More precisely, interference is nothing more than a generic discrepancy between *actual* indivisible stochastic dynamics and *hypothetically* divisible stochastic dynamics.

In particular, quantum-mechanical interference does not imply that matter has a physically wavelike nature, contrary to frequent claims in textbook treatments like [53]. Indeed, from the perspective of the present discussion, the notion that quantum-mechanical interference ever necessitated assigning matter a physically wavelike quality was merely an unfortunate accident of history, arising from the fact that many early empirical examples of interference in quantum systems happened to resemble the behavior of interfering waves propagating in three-dimensional physical space.

These historical examples were clearly special cases. *Multiparticle* systems have Schrödinger waves that propagate through high-dimensional configuration spaces, as was noted by Schrödinger himself in his early work on wave mechanics [54]. For more abstract systems, like *qubits*, there fail to exist continuous configuration spaces for Schrödinger waves altogether.

This new way of thinking about quantum-mechanical interference has implications for the interpretation of the famous *double-slit experiment*. Recall that in the double-slit experiment, an emitter sends particles one at a time toward a wall with two slits in it, and a detection screen on the other side of the wall records the particle's eventual landing site. In the usual 'classical' description of the experiment, one asks first which slit the particle enters, and then, *conditioning on the answer*, one then *restarts* the dynamics with that slit as the new initial condition. Over many repetitions of the experiment, the detection screen records a blend consisting of the statistical distribution of landing sites from particles passing through the *upper* slit with the statistical distribution of landing sites from particles passing through the *lower* slit. In the case of quantum-mechanical particles like electrons, however, one instead finds that the landing sites form a wavelike interference pattern, and the conclusion is supposedly that each particle is really a Schrödinger wave of some kind, or that the particle fails to go through one slit or the other.[10]

According to the approach laid out in this paper, the particle really does go through a specific slit in each run of the experiment. The final interference pattern on the detection screen is due to the

[10] The exposition in [53] ends up at precisely such a conclusion: "It is *not* true that the electrons go *either* through hole 1 or hole 2." [Emphasis in the original.] This conclusion, however, does not logically follow from the empirical appearance of interference effects, but also implicitly depends on the hidden assumption that the behavior of an electron in a double-slit experiment can be described by divisible dynamics.

generic indivisibility of time evolution for quantum systems. One cannot divide up the particle's evolution into, firstly, its transit from the emitter to the slits, and then secondly, conditioned on which slit the particle enters, the particle's transit from the slits to the detection screen. The interference that shows up in the double-slit experiment may be surprising, but that is only because indivisible stochastic dynamics can be highly nonintuitive. In the historical absence of a sufficiently comprehensive framework for describing indivisible stochastic dynamics, it was difficult to recognize just how nonintuitive such dynamics could be, or what sorts of empirically appearances it could produce.

In response to this last point, one might suggest that Schrödinger waves nonetheless offer a superior means of explaining why the double-slit experiment yields the results that it does. Unfortunately, such hopes are dashed as soon as one considers sending *two* particles toward the slits on each run of the experiment. A two-particle system's Schrödinger wave evolves in a *six-dimensional* configuration space, which is arguably not more intuitive than indivisible stochastic dynamics.

Of course, if one regards the quantum-mechanical particles that make up matter as arising more fundamentally from underlying *quantum fields*, then the wavelike properties of those quantum fields ensure that particles of matter have wavelike properties as well, and therefore exhibit *wave-particle duality*. That said, there is nothing about the analysis of the double-slit experiment alone that calls for positing quantum fields. The necessity of quantum field theory comes from other theoretical and empirical considerations. (For a modern motivation, see [55].) One should also keep in mind that quantum fields are conceptually distinct from Schrödinger wave functions.

E. Implications of Interference

The fact that interference shows up in a sufficiently generic stochastic model means that relative phase factors in state vectors have clear empirical signatures, even in the absence of the usual axioms of textbook quantum theory. These empirical manifestations of relative phases are strong evidence that it should be possible to carry out measurements on a much wider set of observables than those that are represented by diagonal matrices (34) in a generalized stochastic system's configuration basis. Indeed, Subsection V B will show that non-diagonal, self-adjoint matrices will turn out to be candidate observables as well.

Thinking more broadly, this overall analysis means that if one is given an indivisible generalized stochastic system, then there will generically be a quantitative discrepancy between the system's actual behavior—as predicted theoretically or measured empirically—and predictions made for the system based on the nearest divisible or Markovian approximation to the system's stochastic dynamics. Again, this discrepancy is precisely interference.

One way to understand this discrepancy is to note that under a divisibility approximation, one can assign definite probabilities to each of the system's possible trajectories by iteratively applying transition matrices, according to the composition law $\Gamma(t) = \Gamma(t \leftarrow t')\Gamma(t')$ from (14). Because iteratively applying transition matrices is not possible for a generalized stochastic system with indivisible dynamics, such systems do not generically have well-defined probabilities for entire possible trajectories.

In the Hilbert-space formulation of a generalized stochastic system, one can assign complex-valued quantities called *amplitudes* to the system's possible trajectories, using the fact that unitary time-evolution operators can be composed or applied iteratively, as in $U(t) = U(t \leftarrow t')U(t')$ from (71). These amplitudes form the conceptual basis for the *path-integral formulation* of quantum theory [18–20]. From the standpoint of the stochastic-quantum correspondence, which gives an alternative formulation of quantum theory, the fact that these amplitudes 'interfere' with each other does not mean that they all physically occur in some sort of literal superposition, or that the system simultaneously takes all such paths in reality, but is merely an artifact of the indivisible dynamics of the underlying generalized stochastic system.

Collectively, the foregoing observations lead to the concrete prediction that interference should arise in a much broader class of contexts than just for quantum systems. Indeed, given any probabilistically evolving system with indivisible or non-Markovian dynamics, one can now interpret any discrepancies between the behavior of such a system and the behavior of its nearest divisible or Markovian approximation as manifestations of interference. One could therefore imagine experimentally measuring interference effects for essentially any system that can be modeled using indivisible or non-Markovian stochastic dynamics.

F. Division Events and the Markov Approximation

Why do discrete-time Markov chains (13) provide such a good approximation to so many stochastic processes in the real world? One intuitively reasonable explanation is that when a system is not isolated from a noisy and intrusive environment, delicate correlations from one time to another 'wash out' over short time scales as those correlations leak out into the environment.

Deriving this intuitive picture from first principles in a more precise way might appear to be a difficult task. Indeed, such a derivation would seem to require finding a more general framework for describing a non-Markovian process, and then showing that such a process becomes approximately Markovian in the appropriate physical circumstances. Fortunately, this paper provides just such a framework.

To set things up, one starts by introducing a composite system $\mathcal{SE}$ consisting of a *subject system* $\mathcal{S}$ together with an *environment* $\mathcal{E}$. The configurations of the subject system's configuration space $\mathcal{C}_{\mathcal{S}}$ will be labeled by $i = 1, \ldots, N$, and the configurations of the environment's configuration space $\mathcal{C}_{\mathcal{E}}$ will be labeled by $e = 1, \ldots, M$, where $M \geq N$. The configuration space of the composite system is then the Cartesian product[11]

$$\mathcal{C}_{\mathcal{SE}} = \mathcal{C}_{\mathcal{S}} \times \mathcal{C}_{\mathcal{E}}, \tag{77}$$

meaning that each element of $\mathcal{C}_{\mathcal{SE}}$ is a simple ordered pair of the form (i, e). One then singles out N configurations of the environment by labeling them as $e(1), \ldots, e(N)$.

For the dynamics, suppose for simplicity that the composite system evolves according to an

[11] The right-hand side of (77) is indeed a Cartesian product, not a tensor product, because this equation is solely a statement about the composite system's configuration space, not its dynamics or its Hilbert-space representation.

overall unistochastic transition matrix

$$\Gamma^{\mathcal{SE}}(t) = \overline{U^{\mathcal{SE}}(t)} \odot U^{\mathcal{SE}}(t), \tag{78}$$

or, in terms of individual entries,

$$\Gamma^{\mathcal{SE}}_{ie,i_0e_0}(t) = |U^{\mathcal{SE}}_{ie,i_0e_0}(t)|^2. \tag{79}$$

Furthermore, suppose that the subject system and the environment interact up to a time $t' > 0$ in such a way that they end up with joint probabilities of the form

$$p^{\mathcal{SE}}_{i'e'}(t') = p^{\mathcal{S}}_{i'}(t')\delta_{e'e(i')}, \tag{80}$$

which describe an idealized *statistical correlation* between the configuration i' of the subject system at t' and the corresponding configuration $e(i')$ of the environment.

If there is to be any possibility of the two subsystems evolving independently for times $t > t'$ after the interaction has concluded, then it should be possible to factorize the composite system's relative time-evolution operator $U^{\mathcal{SE}}(t \leftarrow t')$ between the two subsystems for $t > t'$ as the following tensor product:[12]

$$\left.\begin{aligned} U^{\mathcal{SE}}(t \leftarrow t') = U^{\mathcal{S}}(t \leftarrow t') \otimes U^{\mathcal{E}}(t \leftarrow t') \\ \text{for } t > t'. \end{aligned}\right\} \tag{81}$$

In terms of individual entries, one has

$$\left.\begin{aligned} U^{\mathcal{SE}}_{ie,i'e'}(t \leftarrow t') = U^{\mathcal{S}}_{ii'}(t \leftarrow t')U^{\mathcal{E}}_{ee'}(t \leftarrow t') \\ \text{for } t > t', \end{aligned}\right\} \tag{82}$$

meaning that each entry $U^{\mathcal{SE}}_{ie,i'e'}(t \leftarrow t')$ of the composite system's relative time-evolution operator is the product of corresponding entries $U^{\mathcal{S}}_{ii'}(t \leftarrow t')$ and $U^{\mathcal{E}}_{ee'}(t \leftarrow t')$ of the relative time-evolution operators for the two subsystems individually.

In light of the Born rule (40), the joint probabilities (80) correspond to a wave function[13]

$$\Psi^{\mathcal{SE}}_{i'e'}(t') = \Psi^{\mathcal{S}}_{i'}(t')\delta_{e'e(i')}. \tag{83}$$

The composite system's wave function at later times $t > t'$ after the interaction is therefore given in terms of the relative time-evolution operator (82) according to

---

[12] Note the natural appearance of a tensor product in (81), which is a statement about the composite system's dynamics in the system's Hilbert-space representation.

[13] If necessary, one can easily write down idealized examples of unitary time-evolution operators for the composite system that produce the wave function (83). For instance, one could use $U^{\mathcal{SE}}(t') \equiv \sum_{i'} P^{\mathcal{S}}_{i'} \otimes R^{\mathcal{E}}_{e(i')}$, where $P^{\mathcal{S}}_{i'}$ is the i'th configuration projector (21) for the subject system, and where $R^{\mathcal{E}}_{e(i')}$ is a unitary transformation that takes the environment's initial configuration to the configuration $e(i')$.

$$\left.\begin{aligned}\Psi^{\mathcal{SE}}_{ie}(t) &= \sum_{i',e'} U^{\mathcal{SE}}_{ie,i'e'}\big(t \leftarrow t'\big)\Psi^{\mathcal{SE}}_{i'e'}\big(t'\big) \\ &= \sum_{i'} U^{\mathcal{S}}_{ii'}\big(t \leftarrow t'\big)\Psi^{\mathcal{S}}_{i'}\big(t'\big)U^{\mathcal{E}}_{ee(i')}\big(t \leftarrow t'\big).\end{aligned}\right\} \tag{84}$$

From the Born rule (40), one sees that the joint probabilities for $t > t'$ are given by

$$p^{\mathcal{SE}}_{ie}(t) = \big|\Psi^{\mathcal{SE}}_{ie}(t)\big|^2. \tag{85}$$

Marginalizing over the configuration e of the environment and invoking the unitarity of the environment's relative time-evolution operator $U^{\mathcal{E}}(t \leftarrow t')$, one obtains the standalone probabilities $p^{\mathcal{S}}_i(t)$ for the subject system alone for $t > t'$:

$$\left.\begin{aligned}p^{\mathcal{S}}_i(t) &= \sum_e p^{\mathcal{SE}}_{ie}(t) \\ &= \sum_{i'_1,i'_2} \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')\Psi^{\mathcal{S}}_{i'_1}(t')}U^{\mathcal{S}}_{ii'_2}\big(t \leftarrow t'\big)\Psi^{\mathcal{S}}_{i'_2}\big(t'\big) \\ &\qquad \times \sum_e \overline{U^{\mathcal{E}}_{ee(i'_1)}(t \leftarrow t')}U^{\mathcal{E}}_{ee(i'_2)}\big(t \leftarrow t'\big) \\ &= \sum_{i'} |U^{\mathcal{S}}_{ii'}\big(t \leftarrow t'\big)|^2|\Psi^{\mathcal{S}}_{i'}\big(t'\big)|^2.\end{aligned}\right\} \tag{86}$$

Taking the limit $t \to t'$ in (86) and referring back to the Born rule (40) again, one sees that the subject system's standalone probabilities at the time t' are

$$p^{\mathcal{S}}_{i'}\big(t'\big) = |\Psi^{\mathcal{S}}_{i'}\big(t'\big)|^2. \tag{87}$$

One also sees from (86) that, as in (75), one can identify

$$\Gamma^{\mathcal{S}}_{ii'}\big(t \leftarrow t'\big) \equiv |U^{\mathcal{S}}_{ii'}\big(t \leftarrow t'\big)|^2. \tag{88}$$

Hence, (86) simplifies to a genuinely linear relationship that precisely mirrors the Bayesian marginalization formula (7), with t' now effectively serving as a new 'initial time':

$$p^{\mathcal{S}}_i(t) = \sum_{i'} \Gamma^{\mathcal{S}}_{ii'}\big(t \leftarrow t'\big)p^{\mathcal{S}}_{i'}\big(t'\big). \tag{89}$$

Applying the original Bayesian marginalization formula (7) from the actual initial time 0 to the time t', one also has the equation

$$p^{\mathcal{S}}_{i'}\big(t'\big) = \sum_j \Gamma^{\mathcal{S}}_{i'j}\big(t'\big)p^{\mathcal{S}}_j(0). \tag{90}$$

Combining (89) with (90) immediately yields

$$p_i^{\mathcal{S}}(t) = \sum_j \Gamma_{ij}^{\mathcal{S}}(t) p_j^{\mathcal{S}}(0), \tag{91}$$

where $\Gamma^{\mathcal{S}}(t)$ is a manifestly *divisible* transition matrix:

$$\Gamma^{\mathcal{S}}(t) \equiv \Gamma^{\mathcal{S}}\big(t \leftarrow t'\big)\Gamma^{\mathcal{S}}\big(t'\big). \tag{92}$$

Thus, the interaction between the subject system $\mathcal{S}$ and the environment $\mathcal{E}$ up to the time t' has led to a transition matrix $\Gamma^{\mathcal{S}}(t)$ for the subject system that is momentarily divisible at t'.

It is natural to refer to t' as a *division event*. An important corollary is that the initial time 0 is not a unique or special time, but is instead only one of many division events inevitably experienced by a system in sufficiently strong contact with a repeatedly eavesdropping environment—in the sense that the interactions with the environment lead to correlations that look approximately like those in (80). Division events will play a crucial role in the rest of this paper.

Suppose that these kinds of division events can be approximated as occurring regularly over a characteristic time scale δt. Suppose, moreover, that the unistochastic dynamics is homogeneous in time, in the sense that $U^{\mathcal{S}}(t + \delta t \leftarrow t) = U^{\mathcal{S}}(\delta t)$ for all times t. Then the subject system's transition matrix after any integer number $n \geq 1$ of time steps δt is given by

$$\Gamma^{\mathcal{S}}(n\,\delta t) = \big(\Gamma^{\mathcal{S}}\big)^n, \tag{93}$$

where

$$\Gamma_{ij}^{\mathcal{S}} \equiv |U_{ij}^{\mathcal{S}}(\delta t)|^2. \tag{94}$$

The stochastic dynamics therefore takes the form of a discrete-time Markov chain (13). This analysis therefore provides a theoretical explanation for the ubiquity of Markovian stochastic dynamics in so many real-world cases, and represents another new result.

G. Decoherence

Had the environment not interacted with the subject system, then the subject system's density matrix $\rho^{\mathcal{S}}(t')$ at the time t' would have generically been non-diagonal, in accordance with the general definition (32):

$$\left.\begin{aligned} \rho^{\mathcal{S}}\big(t'\big) &= U^{\mathcal{S}}\big(t'\big)\left[\sum_j p_j(0) P_j\right] U^{\mathcal{S}\dagger}\big(t'\big) \\ &= U^{\mathcal{S}}\big(t'\big)\mathrm{diag}(\ldots, p_j(0), \ldots) U^{\mathcal{S}\dagger}\big(t'\big). \end{aligned}\right\} \tag{95}$$

By contrast, suppose that the environment indeed interacts with the subject system to produce

a division event (92) at t'. In that case, the standalone probability $p_i^{\mathcal{S}}(t)$ for the subject system to occupy its ith configuration at $t > t'$ is given by the linear marginalization relationship (89), which can be written instead as

$$p_i^{\mathcal{S}}(t) = \mathrm{tr}(P_i \rho^{\mathcal{S}}(t)), \tag{96}$$

where

$$\rho^{\mathcal{S}}(t) \equiv U^{\mathcal{S}}\big(t \leftarrow t'\big)\rho^{\mathcal{S}}\big(t'\big)U^{\mathcal{S}\dagger}\big(t \leftarrow t'\big), \tag{97}$$

and where, in turn,

$$\rho^{\mathcal{S}}\big(t'\big) \equiv \sum_{i'} p_{i'}^{\mathcal{S}}\big(t'\big)P_{i'}^{\mathcal{S}} = \mathrm{diag}\big(\dots, p_{i'}^{\mathcal{S}}\big(t'\big), \dots\big), \tag{98}$$

which is diagonal.

On comparing the two expressions (95) and (98) for the subject system's density matrix $\rho(t')$ at t', one sees that the interaction with the environment has effectively eliminated the off-diagonal entries, or *coherences*, in the subject system's density matrix. This phenomenon is called *decoherence*, and the foregoing analysis makes clear that decoherence is nothing more than the mundane leakage of correlations into the environment when viewed through the lens of the Hilbert-space formulation.

This analysis also sheds new light on the meaning of coherences in density matrices, as well as on *superpositions* in state vectors, where superpositions are related to coherences in the case of a rank-one density matrix through the formula $\rho_{i_1 i_2}(t) = \Psi_{i_1}(t)\overline{\Psi_{i_2}(t)}$, in accordance with (39). From the standpoint of this analysis, superpositions and coherences are merely indications that one is catching a given system when it is in the midst of an indivisible stochastic process, between division events, rather than implying that the system is literally in 'multiple states at once.'

These results may also help explain why the precise connection between quantum theory and stochastic processes has historically remained unclear for so long. If one assumes a Markov approximation, as is often the case in the research literature on stochastic processes, then coherences and superposition do not show up, meaning that density matrices remain diagonal, state vectors remain trivial, and nontrivial unistochastic dynamics cannot arise.[14]

H. Entanglement

Consider next a composite system $\mathcal{AB}$ consisting of a pair of subsystems $\mathcal{A}$ and $\mathcal{B}$. Suppose that the two subsystems do not interact from the initial time 0 up to some later time $t' > 0$, but then begin interacting at t'.

For times t between 0 and t', the absence of interactions means that the composite system's transition matrix $\Gamma^{\mathcal{AB}}(t)$ factorizes into the tensor product of a transition matrix $\Gamma^{\mathcal{A}}(t)$ for $\mathcal{A}$ and

[14] See [56] for an analysis of the connection between decoherence and Markovian dynamics within the standard Hilbert-space formulation of quantum theory.

a separate transition matrix $\Gamma^{\mathcal{B}}(t)$ for $\mathcal{B}$:

$$\Gamma^{\mathcal{AB}}(t) = \Gamma^{\mathcal{A}}(t) \otimes \Gamma^{\mathcal{B}}(t) \quad \text{for } 0 \leq t < t'. \tag{99}$$

Starting at the time t', however, the composite system's transition matrix $\Gamma^{\mathcal{AB}}(t)$, which encodes cumulative statistical information and therefore correlations, will fail to factorize between the two subsystems, in the sense that

$$\Gamma^{\mathcal{AB}}(t) \neq \Gamma^{\mathcal{A}}(t) \otimes \Gamma^{\mathcal{B}}(t) \quad \text{for } t > t', \tag{100}$$

for any possible transition matrices $\Gamma^{\mathcal{A}}(t)$ and $\Gamma^{\mathcal{B}}(t)$ that properly capture the respective dynamics of the two subsystems. (It is worth noting that this loss of factorization gives a highly general, model-independent way to define an interaction.) Even if the two subsystems have a notion of localizability in space, and are eventually placed at a large separation distance at some time $t > t'$, the composite system's transition matrix will still fail to factorize between the two subsystems, thereby leading to the appearance of what looks like nonlocal stochastic dynamics across that separation distance.[15]

However, if the composite system exhibits a division event of the form (92) at some later time $t'' > t'$, perhaps due to interactions between one of the subsystems and the larger environment, then the composite system's transition matrix $\Gamma^{\mathcal{AB}}(t)$ will divide at t'':

$$\Gamma^{\mathcal{AB}}(t) = \Gamma^{\mathcal{AB}}\big(t \leftarrow t''\big)\Gamma^{\mathcal{AB}}\big(t''\big) \quad \text{for } t > t'' > t'. \tag{101}$$

If the two subsystems $\mathcal{A}$ and $\mathcal{B}$ do not interact with each other after t', then the *relative* transition matrix $\Gamma^{\mathcal{AB}}(t \leftarrow t'')$ appearing in (101) will factorize between them,

$$\Gamma^{\mathcal{AB}}\big(t \leftarrow t''\big) = \Gamma^{\mathcal{A}}\big(t \leftarrow t''\big) \otimes \Gamma^{\mathcal{B}}\big(t \leftarrow t''\big), \tag{102}$$

so the two subsystems will cease exhibiting what had looked like nonlocal stochastic dynamics.

This analysis precisely captures the quantum-theoretic notion of *entanglement*. Systems that interact with each other start to exhibit what appears to be a nonlocal kind of stochastic dynamics, even if the systems are moved far apart in physical space, and decoherence by the environment effectively causes a breakdown in that apparent dynamical nonlocality.

This stochastic picture of entanglement and nonlocality provides a new way to understand why they occur in the first place. The indivisible nature of generic stochastic dynamics could be viewed as a form of nonlocality in time, which then leads to an apparent nonlocality across space. A division event leads to an instantaneous restoration of locality in time, which then leads to a momentary restoration of manifest locality across space.

---

[15] Questions about nonlocality will be addressed in detail in Subsection VI C.

V. MEASUREMENTS

A. Emergeables

The preceding sections have shown that a generalized stochastic system—that is, a physical model with kinematics based on a configuration space and dynamics based on a suitable stochastic law—is capable of accounting for signature features of quantum theory, like superposition, interference, decoherence, and entanglement. In addition, the Hilbert-space side of the dictionary (26) contains many expressions and equations that are identical to those found in quantum theory.

However, an actual quantum system also includes observables beyond those that are diagonal, as in (34), in a single basis. Indeed, the existence of noncommuting observables represented by non-diagonal, self-adjoint matrices is another hallmark feature of quantum theory.

Remarkably, a generalized stochastic system will generically contain such observables as well. Specifically, Subsection V B will establish that non-diagonal, self-adjoint matrices represent candidate observables that naturally satisfy the usual probabilistic rules of quantum theory, including the Born rule, all without the need to introduce any new fundamental axioms. In so doing, the analysis ahead will demonstrate that the dictionary (26) is not merely a tool for studying a broad class of stochastic processes, but truly defines a comprehensive stochastic-quantum correspondence.

As motivation, let A be a time-independent (diagonal) random variable (34), and consider the time derivative of its Heisenberg-picture counterpart $A^H(t)$, as defined for a generic time-evolution operator $\Theta(t)$ by (42):

$$\frac{dA^H(t)}{dt} = \frac{\partial\Theta^\dagger(t)}{\partial t}A\Theta(t) + \Theta^\dagger(t)A\frac{\partial\Theta(t)}{\partial t}. \tag{103}$$

Evaluating this matrix in the limit $t \to 0$ gives an $N \times N$ self-adjoint, generically non-diagonal matrix $\dot{A}$ at the initial time 0:

$$\dot{A} \equiv \lim_{t\to 0}\frac{dA^H(t)}{dt} = \dot{A}^\dagger. \tag{104}$$

This matrix will not generally commute with the original random variable A itself:

$$[A, \dot{A}] \neq 0. \tag{105}$$

Nonetheless, the matrix $\dot{A}$ has physical uses. For example, one has

$$\mathrm{tr}(\dot{A}\rho(0)) = \lim_{t\to 0}\frac{d\langle A(t)\rangle}{dt}, \tag{106}$$

which is a perfectly meaningful physical quantity, even though the time derivative of an expectation value is not necessarily the expectation value of something physical.

The matrix $\dot{A}$ therefore resembles a random variable in some ways, but incorporates stochastic dynamics directly into its definition (104) through the time-evolution operator $\Theta(t)$, and does not have a transparent interpretation at the level of the generalized stochastic system's underlying

configuration space $\mathcal{C}$ alone. Instead, $\dot{A}$ is an emergent amalgam of kinematical and dynamical ingredients, so it will be called an *emergeable*.[16] This terminology is intended to contrast $\dot{A}$ with the system's genuine random variables, which could be called *beables*—that is, 'be-ables'—to invoke a term coined by Bell in [57].

As a concrete example, consider a particle whose underlying *position* is regarded as a physical configuration, corresponding to some random variable A. If the particle's dynamics is stochastic, in the sense that the particle can be described as a generalized stochastic system, then the particle's *velocity* (or, equivalently, the particle's *momentum*) will not generally have a well-defined value at all times, and will naturally be representable as an emergeable $\dot{A}$ along the lines described here.

B. The Measurement Process

With all the requisite conceptual background now in place, one can model the measurement of a generic observable as a physical process. To start, consider a composite system $\mathcal{SDE}$ consisting of three subsystems that will be called a *subject system* $\mathcal{S}$, a *measuring device* $\mathcal{D}$, and an *environment* $\mathcal{E}$. Note that one of the additional goals ahead will be to identify the criteria for a subsystem like $\mathcal{D}$ to be regarded as a genuine measuring device.

Focusing momentarily on the subject system $\mathcal{S}$, consider an $N \times N$ self-adjoint matrix

$$\tilde{A}^{\mathcal{S}} = \tilde{A}^{\mathcal{S}\dagger}, \tag{107}$$

which may or may not be one of the subject system's diagonal random variables.[17] As a concrete example, $\tilde{A}^{\mathcal{S}}$ could be an emergeable like (104).

By the *spectral theorem*, $\tilde{A}^{\mathcal{S}}$ has a *spectral decomposition* of the form

$$\tilde{A}^{\mathcal{S}} = \sum_{\alpha} \tilde{a}_{\alpha} \tilde{P}^{\mathcal{S}}_{\alpha}, \tag{108}$$

where $\tilde{a}_{\alpha}$ are the eigenvalues of $\tilde{A}^{\mathcal{S}}$ and where $\tilde{P}^{\mathcal{S}}_{\alpha}$ are its eigenprojectors. These eigenprojectors $\tilde{P}^{\mathcal{S}}_{\alpha}$ are not generically diagonal, but they nonetheless satisfy the analogues of the mutual exclusivity condition (24),

$$\tilde{P}^{\mathcal{S}}_{\alpha} \tilde{P}^{\mathcal{S}}_{\alpha'} = \delta_{\alpha\alpha'} \tilde{P}^{\mathcal{S}}_{\alpha}, \tag{109}$$

and the completeness relation (25),

$$\sum_{\alpha} \tilde{P}^{\mathcal{S}}_{\alpha} = \mathbb{1}^{\mathcal{S}}, \tag{110}$$

where $\mathbb{1}^{\mathcal{S}}$ is the identity matrix for the subject system. These eigenprojectors therefore constitute a projection-valued measure (PVM) of their own. Letting $\tilde{e}^{\mathcal{S}}_{\alpha}$ be the corresponding orthonormal

[16] There is a sense in which emergeables are not an entirely new idea, but are closely related to emergent physical properties like temperatures or pressures.

[17] More generally, one could take $\tilde{A}^{\mathcal{S}}$ to be a *normal matrix*, meaning a matrix that commutes with its adjoint $\tilde{A}^{\mathcal{S}\dagger}$.

basis, one has, in analogy with (29), the following statements:

$$\tilde{e}_{\alpha}^{\mathcal{S}\dagger}\tilde{e}_{\alpha'}^{\mathcal{S}} = \delta_{\alpha\alpha'}, \quad \tilde{e}_{\alpha}^{\mathcal{S}}\tilde{e}_{\alpha}^{\mathcal{S}\dagger} = \tilde{P}_{\alpha}^{\mathcal{S}}. \tag{111}$$

If $\tilde{A}^{\mathcal{S}}$ happens to be one of the subject system's random variables (34), then the eigenvalues $\tilde{a}_{\alpha}$ are the random variable's magnitudes, and the eigenprojectors $\tilde{P}_{\alpha}^{\mathcal{S}}$ are the configuration projectors (21). More generally, however, the eigenvalues $\tilde{a}_{\alpha}$ and the eigenprojectors $\tilde{P}_{\alpha}^{\mathcal{S}}$ do not yet have an obvious physical meaning.

Suppose that the measuring device $\mathcal{D}$ has configurations $d(\alpha)$ that can be labeled by the same index α that appears in the spectral decomposition (108). Similarly, suppose that the environment $\mathcal{E}$ has configurations $e(\alpha)$ that can also be labeled by α.

Generalizing (79) from the earlier analysis of the decoherence process, suppose, moreover, that the composite system $\mathcal{SDE}$ evolves according to an overall unistochastic transition matrix:

$$\Gamma_{ide,i_0d_0e_0}^{\mathcal{SDE}}(t) = |U_{ide,i_0d_0e_0}^{\mathcal{SDE}}(t)|^2. \tag{112}$$

Generalizing (83) and letting $\tilde{e}_{\alpha',i'}^{\mathcal{S}}$ denote the i'th component of the basis vector $\tilde{e}_{\alpha'}^{\mathcal{S}}$ with respect to the subject system's configuration basis $e_{i'}^{\mathcal{S}}$, suppose that the three subsystems interact up to a time $t' > 0$ in such a way that they end up with the overall wave function[18]

$$\Psi_{i'd'e'}^{\mathcal{SDE}}\big(t'\big) = \sum_{\alpha'} \tilde{\Psi}_{\alpha'}^{\mathcal{S}}\big(t'\big)\tilde{e}_{\alpha',i'}^{\mathcal{S}}\delta_{d'd(\alpha')}\delta_{e'e(\alpha')}, \tag{113}$$

and that, mirroring (81), the composite system's relative time-evolution operator factorizes between the three subsystems for later times $t > t'$:

$$\left.\begin{aligned} U^{\mathcal{SDE}}\big(t \leftarrow t'\big) = U^{\mathcal{S}}\big(t \leftarrow t'\big) \otimes U^{\mathcal{D}}\big(t \leftarrow t'\big) \\ \otimes U^{\mathcal{E}}\big(t \leftarrow t'\big) \\ \text{for } t > t'. \end{aligned}\right\} \tag{114}$$

Then the composite system's wave function for times $t > t'$ after the interaction is

$$\left.\begin{aligned} \Psi_{ide}^{\mathcal{SDE}}(t) &= \sum_{i',e',d'} U_{ide,i'd'e'}^{\mathcal{SDE}}\big(t \leftarrow t'\big)\Psi_{i'd'e'}^{\mathcal{SDE}}\big(t'\big) \\ &= \sum_{i'}\sum_{\alpha'} U_{ii'}^{\mathcal{S}}\big(t \leftarrow t'\big)\tilde{\Psi}_{\alpha'}^{\mathcal{S}}\big(t'\big)\tilde{e}_{\alpha',i'}^{\mathcal{S}} \\ &\quad \times U_{dd(\alpha')}^{\mathcal{D}}\big(t \leftarrow t'\big)U_{ee(\alpha')}^{\mathcal{E}}\big(t \leftarrow t'\big). \end{aligned}\right\} \tag{115}$$

Invoking the Born rule (40), it follows from the explicit expression (115) for the composite

[18] It is straightforward to write down idealized examples of suitable unitary time-evolution operators for the composite system. One choice is $U^{\mathcal{SDE}}(t') \equiv \sum_{\alpha'} \tilde{P}_{\alpha'}^{\mathcal{S}} \otimes R_{d(\alpha')}^{\mathcal{D}} \otimes R_{e(\alpha')}^{\mathcal{E}}$, where $\tilde{P}_{\alpha'}^{\mathcal{S}}$ is the α'th eigenprojector appearing in the spectral decomposition (108), and where $R_{d(\alpha')}^{\mathcal{D}}$ and $R_{e(\alpha')}^{\mathcal{E}}$ are unitary transformations for the measuring device and the environment that respectively put them in the configurations $d(\alpha')$ and $e(\alpha')$.

system's wave function that the joint probabilities for $t > t'$ are given by

$$p_{ide}^{\mathcal{SDE}}(t) = \left|\Psi_{ide}^{\mathcal{SDE}}(t)\right|^2. \tag{116}$$

Marginalizing over the configuration i of the subject system $\mathcal{S}$ as well as the configuration e of the environment $\mathcal{E}$, and invoking the unitarity of both the subject system's relative time-evolution operator $U^{\mathcal{S}}(t \leftarrow t')$ and the environment's relative time-evolution operator $U^{\mathcal{E}}(t \leftarrow t')$, one obtains the standalone probabilities $p_d^{\mathcal{D}}(t)$ for the measuring device $\mathcal{D}$ alone for $t > t'$:

$$\left.\begin{aligned} p_d^{\mathcal{D}}(t) &= \sum_{i,e} p_{ide}^{\mathcal{SDE}}(t) \\ &= \sum_{i_1',i_2'} \sum_{\alpha_1',\alpha_2'} \overline{U_{dd(\alpha_1')}^{\mathcal{D}}(t \leftarrow t') \tilde{\Psi}_{\alpha_1'}^{\mathcal{S}}(t') \tilde{e}_{\alpha_1',i_1'}^{\mathcal{S}}} \\ &\qquad \times U_{dd(\alpha_2')}^{\mathcal{D}}\big(t \leftarrow t'\big) \tilde{\Psi}_{\alpha_2'}^{\mathcal{S}}\big(t'\big) \tilde{e}_{\alpha_2',i_2'}^{\mathcal{S}} \\ &\qquad \times \sum_i \overline{U_{ii_1'}^{\mathcal{S}}(t \leftarrow t')} U_{ii_2'}^{\mathcal{S}}\big(t \leftarrow t'\big) \\ &\qquad \times \sum_e \overline{U_{ee(\alpha_1')}^{\mathcal{E}}(t \leftarrow t')} U_{ee(\alpha_2')}^{\mathcal{E}}\big(t \leftarrow t'\big) \\ &= \sum_{\alpha'} |U_{dd(\alpha')}^{\mathcal{D}}\big(t \leftarrow t'\big)|^2 |\tilde{\Psi}_{\alpha'}^{\mathcal{S}}\big(t'\big)|^2. \end{aligned}\right\} \tag{117}$$

In the limit $t \to t'$, the last line of (117) implies that

$$p_{d(\alpha')}^{\mathcal{D}}\big(t'\big) = |\tilde{\Psi}_{\alpha'}^{\mathcal{S}}\big(t'\big)|^2. \tag{118}$$

Hence, the measuring device $\mathcal{D}$ has a standalone probability $|\tilde{\Psi}_{\alpha'}^{\mathcal{S}}(t')|^2$ of ending up in its configuration $d(\alpha')$, exactly as predicted by the textbook version of the Born rule. One can then naturally define an expectation value $\langle \tilde{A}^{\mathcal{S}}(t') \rangle$ for $\tilde{A}^{\mathcal{S}}$ at t' as the usual kind of statistical average:

$$\langle \tilde{A}^{\mathcal{S}}\big(t'\big) \rangle \equiv \sum_{\alpha} \tilde{a}_{\alpha} p_{d(\alpha')}^{\mathcal{D}}\big(t'\big). \tag{119}$$

This analysis establishes that as long as there exists a form of unistochastic time evolution (112) for the composite system $\mathcal{SDE}$ that arrives at the wave function (113), the matrix $\tilde{A}^{\mathcal{S}}$ represents a genuine observable, in the sense that the time evolution (112) leads to the measuring device ending up in the correct configuration with the correct Born-rule probability.

For times $t > t'$ after the interaction, the last line of (117) implies that the time t' is a division event for the measuring device:

$$\Gamma^{\mathcal{D}}(t) = \Gamma^{\mathcal{D}}\big(t \leftarrow t'\big)\Gamma^{\mathcal{D}}\big(t'\big) \quad \text{for } t > t'. \tag{120}$$

Here the measuring device's dynamics for times $t > t'$ is given by the relative unistochastic transi-

tion matrix

$$\Gamma^{\mathcal{D}}_{dd(\alpha')}\big(t \leftarrow t'\big) \equiv |U^{\mathcal{D}}_{dd(\alpha')}\big(t \leftarrow t'\big)|^2. \tag{121}$$

By contrast, if the observable $\tilde{A}^{\mathcal{S}}$ is an emergeable, as opposed to one of the subject system's (diagonal) random variables (34), then the subject system $\mathcal{S}$ does not experience a division event at t'. Instead, the subject system remains mired in indivisible time evolution at t', with some stochastically evolving underlying configuration. Moreover, if indeed $\tilde{A}^{\mathcal{S}}$ is an emergeable, then the measurement result obtained by the measuring device is an emergent effect of the interaction between the subject system and the measuring device, in a sense suggested by Bohr in 1935 [58, 59], rather than transparently revealing a physical aspect of the configuration of the subject system alone.

Despite t' not necessarily being a division event for the subject system $\mathcal{S}$, one can nevertheless compute the standalone probability $p^{\mathcal{S}}_i(t)$ for the subject system to be in its ith configuration for times $t > t'$ by marginalizing over the measuring device $\mathcal{D}$ and the environment $\mathcal{E}$:

$$\left.\begin{aligned}
p^{\mathcal{S}}_i(t) &= \sum_{d,e} p^{\mathcal{S}\mathcal{D}\mathcal{E}}_{ide}(t) \\
&= \sum_{i'_1,i'_2} \sum_{\alpha'_1,\alpha'_2} \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')\tilde{\Psi}^{\mathcal{S}}_{\alpha'_1}(t')\tilde{e}^{\mathcal{S}}_{\alpha'_1,i'_1}} \\
&\quad \times U^{\mathcal{S}}_{ii'_2}\big(t \leftarrow t'\big)\tilde{\Psi}^{\mathcal{S}}_{\alpha'_2}\big(t'\big)\tilde{e}^{\mathcal{S}}_{\alpha'_2,i'_2} \\
&\quad \times \sum_{d} \overline{U^{\mathcal{D}}_{dd(\alpha'_1)}(t \leftarrow t')} U^{\mathcal{D}}_{dd(\alpha'_2)}\big(t \leftarrow t'\big) \\
&\quad \times \sum_{e} \overline{U^{\mathcal{E}}_{ee(\alpha'_1)}(t \leftarrow t')} U^{\mathcal{E}}_{ee(\alpha'_2)}\big(t \leftarrow t'\big) \\
&= \sum_{\alpha'} \left[\sum_{i'_1,i'_2} \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')} U^{\mathcal{S}}_{ii'_2}\big(t \leftarrow t'\big)\tilde{e}^{\mathcal{S}}_{\alpha',i'_2}\overline{\tilde{e}^{\mathcal{S}}_{\alpha',i'_1}} \right] \\
&\quad \times |\tilde{\Psi}^{\mathcal{S}}_{\alpha'}\big(t'\big)|^2.
\end{aligned}\right\} \tag{122}$$

Recognizing $|\tilde{\Psi}^{\mathcal{S}}_{\alpha'}(t')|^2$ from (118) as the standalone probability $p^{\mathcal{D}}_{d(\alpha')}(t')$ for the measuring device $\mathcal{D}$ to end up in its configuration $d(\alpha')$ at the time t', and recalling both the configuration projectors $P^{\mathcal{S}}_i$ defined in (21) as well as the eigenprojectors $\tilde{P}^{\mathcal{S}}_{\alpha}$ appearing in the spectral decomposition (108) for $\tilde{A}^{\mathcal{S}}$, one can write (122) more succinctly as

$$p^{\mathcal{S}}_i(t) = \mathrm{tr}(P^{\mathcal{S}}_i \rho^{\mathcal{S}}(t)). \tag{123}$$

Here the subject system's density matrix $\rho^{\mathcal{S}}(t)$ for $t > t'$ is given by

$$\rho^{\mathcal{S}}(t) \equiv U^{\mathcal{S}}\big(t \leftarrow t'\big) \left[\sum_{\alpha'} p^{\mathcal{D}}_{d(\alpha')}\big(t'\big)\tilde{P}^{\mathcal{S}}_{\alpha'} \right] U^{\mathcal{S}\dagger}\big(t \leftarrow t'\big). \tag{124}$$

One can therefore recast the expectation value (119) for $\tilde{A}^{\mathcal{S}}$ as

$$\langle \tilde{A}^{\mathcal{S}}\big(t'\big)\rangle = \mathrm{tr}(\tilde{A}^{\mathcal{S}}\rho^{\mathcal{S}}\big(t'\big)), \tag{125}$$

which precisely mirrors the formula (33) for the expectation value of a (diagonal) random variable.

Furthermore, (122) yields a *linear* relationship between the standalone probabilities $p^{\mathcal{D}}_{d(\alpha')}(t')$ for the measuring device $\mathcal{D}$ at t' and the standalone probabilities $p^{\mathcal{S}}_{i}(t)$ for the subject system $\mathcal{S}$ at $t > t'$:

$$p^{\mathcal{S}}_{i}(t) = \sum_{\alpha'} \Gamma^{\mathcal{SD}}_{i,d(\alpha')}\big(t \leftarrow t'\big) p^{\mathcal{D}}_{d(\alpha')}\big(t'\big). \tag{126}$$

The entries $\Gamma^{\mathcal{SD}}_{i,d(\alpha')}(t \leftarrow t')$ of the *hybrid* relative transition matrix appearing here are given explicitly by

$$\left.\begin{aligned} &\Gamma^{\mathcal{SD}}_{i,d(\alpha')}\big(t \leftarrow t'\big) \\ &\equiv \sum_{i'_1,i'_2} \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')} U^{\mathcal{S}}_{ii'_2}\big(t \leftarrow t'\big) \tilde{e}^{\mathcal{S}}_{\alpha',i'_2} \overline{\tilde{e}^{\mathcal{S}}_{\alpha',i'_1}}. \end{aligned}\right\} \tag{127}$$

Because these matrix entries do not depend on the measuring device's standalone probabilities $p^{\mathcal{D}}_{d(\alpha')}(t')$, they naturally serve as *conditional probabilities* for the subject system $\mathcal{S}$ to be in its ith configuration at the time $t > t'$, given that the measuring device $\mathcal{D}$ is in its configuration $d(\alpha')$ at t':

$$p^{\mathcal{SD}}\big(i, t | d(\alpha'), t'\big) \equiv \Gamma^{\mathcal{SD}}_{i,d(\alpha')}\big(t \leftarrow t'\big). \tag{128}$$

C. Wave-Function Collapse

Importantly, notice that one can write the hybrid transition matrix (127) in an overall form that resembles the dictionary (26):

$$\left.\begin{aligned} &\Gamma^{\mathcal{SD}}_{i,d(\alpha')}\big(t \leftarrow t'\big) \\ &\qquad = \mathrm{tr}(U^{\mathcal{S}\dagger}\big(t \leftarrow t'\big) P^{\mathcal{S}}_{i} U^{\mathcal{S}}\big(t \leftarrow t'\big) \tilde{P}^{\mathcal{S}}_{\alpha}). \end{aligned}\right\} \tag{129}$$

Rearranging the right-hand side gives the equation

$$\Gamma^{\mathcal{SD}}_{i,d(\alpha')}\big(t \leftarrow t'\big) = \mathrm{tr}(P^{\mathcal{S}}_{i} \rho^{\mathcal{S}|\alpha',t'}(t)), \tag{130}$$

with a *conditional* density matrix $\rho^{\mathcal{S}|\alpha',t'}(t)$ for the subject system $\mathcal{S}$ at the time $t > t'$ naturally defined by time-evolving the eigenprojector $\tilde{P}^{\mathcal{S}}_{\alpha'}$ from t' to t:

$$\rho^{\mathcal{S}|\alpha',t'}(t) \equiv U^{\mathcal{S}}\big(t \leftarrow t'\big) \tilde{P}^{\mathcal{S}}_{\alpha'} U^{\mathcal{S}\dagger}\big(t \leftarrow t'\big). \tag{131}$$

Thus, the calculation (122) reduces to the statement that the standalone probabilities $p_i^{\mathcal{S}}(t)$ for the subject system at $t > t'$ are given by

$$p_i^{\mathcal{S}}(t) = \mathrm{tr}(P_i^{\mathcal{S}} \rho^{\mathcal{S}}(t)), \tag{132}$$

where the subject system's density matrix $\rho^{\mathcal{S}}(t)$, which was originally defined in (124), can equivalently be expressed as a probabilistic mixture of the conditional density matrices $\rho^{\mathcal{S}|\alpha',t'}(t)$ defined in (131), statistically weighted by the measurement probabilities $p^{\mathcal{D}}_{d(\alpha')}(t')$:

$$\rho^{\mathcal{S}}(t) \equiv \sum_{\alpha'} \rho^{\mathcal{S}|\alpha',t'}(t) p^{\mathcal{D}}_{d(\alpha')}\big(t'\big). \tag{133}$$

Taking stock of these results, one sees that to make future predictions for $t > t'$ about the subject system $\mathcal{S}$, conditioned on the measuring device's result $d(\alpha')$ at t', one uses the conditional probabilities (130), in which the subject system's density matrix has been effectively replaced with the conditional density matrix $\rho^{\mathcal{S}|\alpha',t'}(t)$. This conditional density matrix corresponds to a *collapsed* state vector or wave function defined as

$$\Psi^{\mathcal{S}|\alpha',t'}(t) \equiv U\big(t \leftarrow t'\big) \tilde{e}^{\mathcal{S}}_{\alpha}. \tag{134}$$

The phenomenon of *wave-function collapse* therefore reduces to a prosaic example of conditioning, a conclusion that represents another new result.

By contrast, for an observer who does not know the specific measurement result $d(\alpha')$, the correct density matrix $\rho^{\mathcal{S}}(t)$ to use is the one defined in (133). Again, this density matrix consists of an appropriate probabilistic mixture of conditional or collapsed density matrices that are statistically weighted over the measurement results.

D. The Measurement Problem

According to the foregoing treatment of the measurement process, a measuring device is an ordinary physical system that can carry out a measurement of an observable, and then ends up in a final configuration that reflects a definite measurement outcome. The probabilities for a measuring device's various possible measurement outcomes are given by the textbook Born rule (118), and conditioning on the specific measurement outcome leads to the textbook formula (134) for wave-function collapse. Hence, the picture of quantum theory presented in this paper arguably has the resources to solve the measurement problem [60].

The stochastic-quantum correspondence is also helpful for understanding the measurement process in another important way. Textbook treatments of quantum theory typically regard measuring devices as axiomatic primitives or posits, without providing clear principles for deciding which kinds of systems merit being called measuring devices and which do not. The approach taken toward the measurement process in this paper not only gives a candidate resolution of the measurement problem, but also yields a natural set of criteria for defining what counts as a good measuring

device in the first place, without the need to regard measuring devices as special among all other systems in any truly fundamental way.

Based on the approach in this paper, one sees that a good measuring device should be a physical system with at least as many configurations as possible outcomes for the observable to be measured (at least up to the desired level of experimental resolution), it should admit an overall form of dynamics that results in the correct final correlations, and it should be in sufficiently strong contact with a noisy and intrusive environment to generate a robust division event at the conclusion of the measurement interaction. It is worth noting that the first two of these three criteria would be standard requirements for a measuring device even without worrying about indivisible stochastic dynamics or quantum theory.[19]

E. The Uncertainty Principle

Again, the foregoing treatment of the measurement process leads to the textbook Born rule (118) and the textbook formula (134) for wave-function collapse. As a consequence, any pair of observables $\tilde{A}, \tilde{B}$ and their respective standard deviations $\Delta\tilde{A}, \Delta\tilde{B}$ will satisfy the *Heisenberg-Robertson uncertainty principle* [62, 63],

$$\Delta\tilde{A}\,\Delta\tilde{B} \geq \frac{1}{2}|\mathrm{tr}(i[\tilde{A}\tilde{B} - \tilde{B}\tilde{A}]\rho)|, \tag{135}$$

as follows from any of the standard proofs.[20]

The stochastic-quantum correspondence goes beyond replicating the uncertainty principle by painting a clearer picture of what the uncertainty principle physically means. Consider for simplicity the case in which $\tilde{A} = A$ is a random variable, or beable, and $\tilde{B}$ is an emergeable, in the language of Subsection V A. Then A has a direct interpretation solely in terms of the subject system's configuration space, whereas $\tilde{B}$ encodes an emergent pattern in the subject system's dynamics that can nonetheless show up in the measurement outcomes of a measuring device.

Suppose that A has a definite value or magnitude at some initial time 0. Then the subject system must be in a specific configuration with probability 1 at the initial time 0. The overall stochastic dynamics will then lead to uncertainty in the outcome of any measurement of $\tilde{B}$.

Suppose that one goes ahead and measures $\tilde{B}$, so that a definite measurement outcome emergently shows up in the configuration of a measuring device at some time $t' > 0$. Then the analysis in Subsection V C implies that there is an inevitable disturbance in the subject system that leads its density matrix to end up effectively as a non-diagonal matrix equal to an eigenprojector of $\tilde{B}$. A non-diagonal density matrix signifies that the system is in the midst of an indivisible stochastic process, as explained in Subsection IV G. In the present circumstances, that indivisible stochastic process is precisely one that would ensure that if $\tilde{B}$ were measured again shortly after t', then

[19] Without the third criterion—strong contact with an environment—one obtains a "latent measurement" in the language of [56, 61].

[20] Here is one proof, which is adapted from [64]: Because density matrices are positive semidefinite, it follows from the spectral theorem that for any $N \times N$ matrix X and any $N \times N$ density matrix ρ, one has the inequality $\mathrm{tr}(X^\dagger X\rho) \geq 0$. Let $\tilde{A}, \tilde{B}$ be any pair of observables, assumed without any real loss of generality to have vanishing expectation values $\langle\tilde{A}\rangle = \langle\tilde{B}\rangle = 0$. Let x be a variable real number and let $X \equiv x\tilde{A} + i\tilde{B}$. Then $\mathrm{tr}(X^\dagger X\rho) \geq 0$ becomes the inequality $f(x) \geq 0$, where $f(x) \equiv ax^2 + bx + c$ is a quadratic function with coefficients $a \equiv (\Delta\tilde{A})^2$, $b \equiv \mathrm{tr}(i[\tilde{A}\tilde{B} - \tilde{B}\tilde{A}]\rho)$, and $c \equiv (\Delta\tilde{B})^2$. The minimum value of $f(x)$ is $-b^2/4a + c$, so one arrives at the inequality $-b^2/4a + c \geq 0$, which one can rearrange to give the Heisenberg-Robertson uncertainty principle. QED

the measuring device would obtain the same outcome for $\tilde{B}$ as before. However, being in the s of an indivisible stochastic process also implies uncertainty in the subject system's underlying configuration, thereby rendering the value of A uncertain.

VI. FURTHER IMPLICATIONS

A. Symmetries

The stochastic-quantum correspondence developed in this paper provides new ways to think about *dynamical symmetries* in quantum theory, meaning transformations that leave the dynamics invariant. Going in the other direction, the stochastic-quantum correspondence also makes it more straightforward to impose dynamical symmetries systematically as constraints in the construction of the dynamics for a given generalized stochastic system.

Classically, any invertible transformation of a system's configurations $i = 1, \ldots, N$ is a permutation transformation of the configuration projectors (21):

$$\left.\begin{array}{l} P_i \mapsto P_{\sigma(i)}, \\ \quad \text{with } \{\sigma(1), \ldots, \sigma(N)\} = \{1, \ldots, N\}. \end{array}\right\} \tag{136}$$

More generally, a transformation between two PVMs $P_1, \ldots, P_N$ and $\tilde{P}_1, \ldots, \tilde{P}_N$ is always a similarity transformation of the form

$$P_i \mapsto \tilde{P}_i \equiv V^\dagger P_i V, \tag{137}$$

where V is some unitary operator.[21] This similarity transformation reduces to the configurational transformation (136) if and only if V is a permutation matrix.

The more general transformation (137) is a dynamical symmetry, meaning that it leaves the stochastic dynamics invariant, precisely if the right-hand side of the stochastic-quantum dictionary (26) remains unchanged:

$$\mathrm{tr}(\Theta^\dagger(t)\tilde{P}_i\Theta(t)\tilde{P}_j) = \mathrm{tr}(\Theta^\dagger(t)P_i\Theta(t)P_j). \tag{138}$$

This condition is equivalent to the statement that

$$\mathrm{tr}(\tilde{\Theta}^\dagger(t)P_i\tilde{\Theta}(t)P_j) = \mathrm{tr}(\Theta^\dagger(t)P_i\Theta(t)P_j), \tag{139}$$

where

$$\tilde{\Theta}(t) \equiv V\Theta(t)V^\dagger. \tag{140}$$

[21] Proof: Let $e_1, \ldots, e_N$ be the orthonormal configuration basis (27), with $e_i^\dagger e_j = \delta_{ij}$ and $e_i e_i^\dagger = P_i$ as in (29), and let $\tilde{e}_1, \ldots, \tilde{e}_N$ be an orthonormal basis related to the new projectors $\tilde{P}_i$ in the analogous way, with $\tilde{e}_i^\dagger \tilde{e}_j = \delta_{ij}$ and $\tilde{e}_i \tilde{e}_i^\dagger = \tilde{P}_i$. Then the $N \times N$ matrix defined by $V \equiv \sum_i e_i \tilde{e}_i^\dagger$ is unitary and satisfies $V^\dagger P_i V = \tilde{P}_i$. Going the other way, if V is an $N \times N$ unitary matrix, then the $N \times N$ matrices defined for $i = 1, \ldots, N$ by $\tilde{P}_i \equiv V^\dagger P_i V$ are guaranteed to constitute a PVM. QED

Re-expressing both sides of the equivalent condition (139) in terms of squared absolute values, as in (17), one sees that (140) is a dynamical symmetry precisely if

$$|\tilde{\Theta}_{ij}(t)|^2 = |\Theta_{ij}(t)|^2. \tag{141}$$

It follows immediately that $\tilde{\Theta}(t)$ can differ from $\Theta(t)$ by at most a Schur-Hadamard gauge transformation (45), so a necessary and sufficient condition for a unitary matrix V to give a dynamical symmetry is that

$$V\Theta(t)V^\dagger = \Theta(t) \odot \begin{pmatrix} e^{i\theta_{11}(t)} & e^{i\theta_{12}(t)} & & \\ e^{i\theta_{21}(t)} & \ddots & & \\ & & & e^{i\theta_{NN}(t)} \end{pmatrix}. \tag{142}$$

As special cases, this condition includes *unitary* dynamical symmetries,

$$V\Theta(t)V^\dagger = \Theta(t), \tag{143}$$

as well as *anti-unitary* dynamical symmetries,

$$V\Theta(t)V^\dagger = \overline{\Theta(t)}, \tag{144}$$

but (142) may also open up the possible existence of dynamical symmetries distinct from these two cases.

For the specific case of an anti-unitary dynamical symmetry, note that if one redefines $V \mapsto \overline{V}$, which is still unitary, then one can re-express (144) in the somewhat more conventional form

$$VK\Theta(t)KV^\dagger = \Theta(t). \tag{145}$$

Here K denotes the *complex-conjugation operator*, meaning that K is an *involution*,

$$K^2 = 1, \tag{146}$$

and, for any $N \times N$ matrix X, one has

$$KXK = \overline{X}. \tag{147}$$

The composite operator VK as a whole is then said to be an *anti-unitary operator*. Anti-unitary operators play an important role in describing *time-reversal symmetries*.[22]

If $\Theta(t) = U(t)$ is unitary, then $V\Theta(t)V^\dagger$ will likewise be unitary. In that case, suppose either that V is continuously connected to the identity matrix $\mathbb{1}$ by some smooth parameter, with a corresponding self-adjoint generator $G = G^\dagger$, or, alternatively, that V is an involution, meaning

[22] Intriguingly, because K anticommutes with i, meaning that $Ki = -iK$, the three mathematical objects i, K, and iK satisfy $-i^2 = K^2 = (iK)^2 = iK(iK) = 1$, and therefore generate a Clifford algebra isomorphic to the *pseudo-quaternions* [64]. In a sense, then, the Hilbert spaces of quantum systems are actually defined not over the complex numbers alone, but over the pseudo-quaternions.

that $V^2 = \mathbb{1}$, in which case $G \equiv V = V^\dagger$ is *itself* self-adjoint. Either way, the expectation value $\langle G(t)\rangle$ is a physically meaningful quantity, and *Noether's theorem* easily follows as the statement that this expectation value is constant in time, or *conserved*:

$$\langle G(t)\rangle = \mathrm{tr}(GU(t)\rho(0)U^\dagger(t)) = \langle G(0)\rangle. \tag{148}$$

B. Dilations

In most textbook treatments of quantum theory, a quantum system is axiomatically defined as a particular Hilbert space together with a preferred set of self-adjoint operators designated as observables with predetermined physical meanings, along with a particular Hamiltonian to define the system's time evolution.[23] From that point of view, modifying a system's Hilbert-space formulation in any nontrivial way would necessarily mean fundamentally modifying the system itself.

From the alternative point of view developed in this paper, by contrast, a Hilbert-space formulation is merely a collection of mathematical tools for constructing the dynamics of a given generalized stochastic system. The generalized stochastic system itself is ultimately defined by a configuration space and a dynamical law that stand apart from any arbitrary choice of Hilbert-space formulation. As a consequence, one is free to modify a generalized stochastic system's Hilbert-space formulation as needed, much like changing from one gauge choice to another in a gauge theory, or like adding physically meaningless variables to the Lagrangian formulation of a deterministic classical system.

With this motivation in place, recall again the basic stochastic-quantum dictionary (26):

$$\Gamma_{ij}(t) = \mathrm{tr}(\Theta^\dagger(t)P_i\Theta(t)P_j). \tag{149}$$

The Hilbert-space formulation expressed by the right-hand side can be manipulated for convenience, provided that the left-hand side of the dictionary remains unchanged.

In particular, for any integer $D \geq 2$, one can freely enlarge, or *dilate*, the Hilbert-space formulation to a larger dimension ND by the following *dilation transformation*:

$$\left.\begin{aligned} \Theta(t) &\mapsto \Theta(t)\otimes\mathbb{1}^{\mathcal{I}}, \\ P_i(t) &\mapsto P_i(t)\otimes\mathbb{1}^{\mathcal{I}}, \\ P_j(t) &\mapsto P_j(t)\otimes P^{\mathcal{I}}_\gamma. \end{aligned}\right\} \tag{150}$$

Here $\mathbb{1}^{\mathcal{I}}$ is the $D\times D$ identity matrix on a new *internal* Hilbert space $\mathcal{H}_{\mathcal{I}}$, and $P^{\mathcal{I}}_1, \ldots, P^{\mathcal{I}}_D$ collectively form any PVM on that internal Hilbert space satisfying the usual conditions of mutual exclusivity,

$$P^{\mathcal{I}}_\gamma P^{\mathcal{I}}_{\gamma'} = \delta_{\gamma\gamma'}P^{\mathcal{I}}_\gamma, \tag{151}$$

[23] In some circumstances, it may turn out to be more convenient to define a quantum system by a formal C*-algebra of observables alone, without picking a specific Hilbert-space representation [6, 65, 66].

and completeness,

$$\sum_{\gamma=1}^{D} P_\gamma^{\mathcal{I}} = \mathbb{1}^{\mathcal{I}}. \tag{152}$$

It is then a mathematical identity that one can rewrite the stochastic-quantum dictionary (26) as

$$\left.\begin{aligned}\Gamma_{ij}(t) = \mathrm{tr}\Big(\mathrm{tr}_{\mathcal{I}}\Big(\Big[\Theta^\dagger(t)\otimes\mathbb{1}^{\mathcal{I}}\Big]\big[P_i\otimes\mathbb{1}^{\mathcal{I}}\big] \\ \times\ \big[\Theta(t)\otimes\mathbb{1}^{\mathcal{I}}\big]\big[P_j\otimes P_\gamma^{\mathcal{I}}\big]\Big)\Big),\end{aligned}\right\} \tag{153}$$

with a second trace, or *partial trace*, over the internal Hilbert space $\mathcal{H}_{\mathcal{I}}$. The choice of value for the label γ here is immaterial, with different choices of γ related by gauge transformations.

One can equivalently write the dilated form (153) of the dictionary in *block-matrix form* as

$$\Gamma_{ij}(t) = \mathrm{tr}_{\mathcal{I}}\Big([\Theta_{ij}(t)]^{\mathcal{I}\dagger}[\Theta_{ij}(t)]^{\mathcal{I}}P_\gamma^{\mathcal{I}}\Big). \tag{154}$$

Here $[\Theta_{ij}(t)]^{\mathcal{I}}$ is a diagonal $D\times D$ matrix consisting of repeated copies of the specific entry $\Theta_{ij}(t)$ (for fixed i,j) along the diagonal:

$$[\Theta_{ij}(t)]^{\mathcal{I}} \equiv \Theta_{ij}(t)\,\mathbb{1}^{\mathcal{I}}. \tag{155}$$

Meanwhile, the adjoint operation $\dagger$ in (154) acts on this $D\times D$ block matrix $[\Theta_{ij}(t)]^{\mathcal{I}}$, so it does not transpose the indices i and j on the $N\times N$ matrix $\Theta_{ij}(t)$ itself:

$$[\Theta_{ij}(t)]^{\mathcal{I}\dagger} \equiv \overline{[\Theta_{ij}(t)]}^{\mathcal{I}}. \tag{156}$$

It follows that

$$[\Theta_{ij}(t)]^{\mathcal{I}\dagger}[\Theta_{ij}(t)]^{\mathcal{I}}P_\gamma^{\mathcal{I}} = |\Theta_{ij}(t)|^2 P_\gamma^{\mathcal{I}}, \tag{157}$$

so the trace over $\mathcal{H}_{\mathcal{I}}$ indeed yields $|\Theta_{ij}(t)|^2 = \Gamma_{ij}(t)$, as required.

In this dilated version of the Hilbert-space formulation, Schur-Hadamard gauge transformations (45) are enhanced to the following local-in-time gauge transformations, which have not yet been described in the research literature and therefore constitute another new result:

$$[\Theta_{ij}(t)]^{\mathcal{I}} \mapsto V_{(ij)}^{\mathcal{I}}(t)[\Theta_{ij}(t)]^{\mathcal{I}}. \tag{158}$$

Here $V_{(ij)}^{\mathcal{I}}(t)$ are a set of N^2 unitary, $D\times D$ matrices, where each such unitary matrix as a whole is labeled by a specific pair (ij) of configuration labels:

$$V_{(ij)}^{\mathcal{I}\dagger}(t) = (V_{(ij)}^{\mathcal{I}}(t))^{-1}. \tag{159}$$

The gauge transformations (158) will not generally preserve the factorization $\Theta(t)\otimes\mathbb{1}^{\mathcal{I}}$ appearing

in (153), so they motivate considering more general $ND \times ND$ time-evolution operators $\tilde{\Theta}(t)$, in terms of which the dilated dictionary (153) takes the form

$$\Gamma_{ij}(t) = \mathrm{tr}\Big(\mathrm{tr}_{\mathcal{I}}\Big(\tilde{\Theta}^{\dagger}(t)\big[P_i \otimes \mathbb{1}^{\mathcal{I}}\big]\tilde{\Theta}(t)\big[P_j \otimes P^{\mathcal{I}}_{\gamma}\big]\Big)\Big). \tag{160}$$

Any $ND \times ND$ matrix $\tilde{\Theta}(t)$ appearing on the right-hand side of this dictionary and satisfying the natural generalization of the summation condition (18) is guaranteed to lead to a valid transition matrix $\Gamma_{ij}(t)$ on the left-hand side, so working with a dilated Hilbert-space formulation essentially provides a larger 'canvas' for designing transition matrices.

As a simple example of a dilation for the case $D = 2$, one can formally eliminate the complex numbers from a quantum system's Hilbert space [67]. Specifically, by increasing the system's Hilbert-space dimension from N to $2N$, one can replace the imaginary unit $i \equiv \sqrt{-1}$ with the real-valued 2×2 matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, with the enhanced version (158) of Schur-Hadamard gauge transformations now consisting of two-dimensional rotations of the internal Hilbert space $\mathcal{H}_{\mathcal{I}}$.[24] One can then represent the complex-conjugation operator K appearing in (145) as the real-valued 2×2 matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. The result is that all unitary and anti-unitary operators become $2N \times 2N$ real orthogonal matrices. One cost of using this 'real' representation, however, is that the Hilbert spaces of composite systems will not factorize as neatly into Hilbert spaces for their constituent subsystems.[25]

As a much more significant application of dilations, recall that any transition matrix $\Gamma_{ij}(t)$ has a Kraus decomposition (51):

$$\Gamma_{ij}(t) = \sum_{\beta=1}^{N} \mathrm{tr}(K^{\dagger}_{\beta}(t) P_i K_{\beta}(t) P_j). \tag{161}$$

The *Stinespring dilation theorem* [68, 69] then guarantees that by an appropriate dilation to a larger Hilbert space if necessary, one can express $\Gamma_{ij}(t)$ in terms of a unitary time-evolution operator $\tilde{U}(t)$:

$$\Gamma_{ij}(t) = \mathrm{tr}\Big(\mathrm{tr}_{\mathcal{I}}\Big(\tilde{U}^{\dagger}(t)\big[P_i \otimes \mathbb{1}^{\mathcal{I}}\big]\tilde{U}(t)\big[P_j \otimes P^{\mathcal{I}}_{\gamma}\big]\Big)\Big). \tag{162}$$

This fact makes clear the inevitability of unitary time evolution in quantum theory. As mentioned earlier, it also implies at the level of a theorem [50] that every generalized stochastic system can be regarded as a subsystem of a unistochastic system.[26]

As yet another key application of dilations, a dilated Hilbert-space formulation can make it possible to describe new kinds of emergeables. Some of these *dilation-emergeables* may be observables

[24] Importantly, notice that the proof of the uncertainty principle presented in Footnote 20 works just as well with the imaginary unit i represented by a 2×2 matrix in this way.

[25] Without increasing the dimension N of a system's Hilbert space, one could instead attempt to limit the appearance of the complex numbers in a system's Hilbert-space formulation by using the original Schur-Hadamard gauge transformation (45) to make all the entries of the system's time-evolution operator $\Theta(t)$ real-valued. In this alternative approach, however, a unistochastic transition matrix $\Gamma(t)$ may not be expressible in terms of a unitary or orthogonal time-evolution operator, and the complex numbers will generally still be needed anyway to define various observables.

[26] From the starting assumptions presented here, one can sketch the following proof: Given $N \times N$ Kraus operators $K_{\beta}(t)$, with $\beta = 1, \ldots, N$, define an $N^3 \times N^2$ matrix $\tilde{V}(t)$ according to $\tilde{V}_{(i\beta m)(jl)}(t) \equiv K_{\beta,ij}(t)\delta_{lm}$, treating $(i\beta m)$ as the first index of $\tilde{V}(t)$ and treating (jl) as its second index. One can show that this matrix satisfies $\tilde{V}^{\dagger}(t)\tilde{V}(t) = \mathbb{1}_{N^2 \times N^2}$, so it defines a partial isometry, which can always be extended to a unitary $N^3 \times N^3$ matrix $\tilde{U}_{(i\beta m)(ja)}(t)$ by adding $N^3 - N^2$ additional columns that are mutually orthogonal with each other and with the previous N^2 columns already in $\tilde{V}(t)$, where the new index a runs through N^2 possible values. These additional columns can always be chosen so that at the initial time 0, where $\tilde{V}(0) = \mathbb{1}$ is the $N \times N$ identity matrix, they

that can yield definite results in measurement processes, along the lines of Subsection V B, despite not having a direct meaning solely at the level of the system's underlying configuration space.

In this way, a generalized stochastic system based on a configuration space can easily accommodate emergent observables that describe empirically meaningful patterns in the dynamics and that model all kinds of quantum phenomena. Indeed, obtaining a unitary time-evolution operator for a given system may require dilating the Hilbert space in just this way, as in (162).

It is important to keep in mind that whether or not one actually carries out this formal dilation of the Hilbert-space formulation, the stochastic dynamics of the underlying generalized stochastic system will still be the same. Any emergent patterns in the system's stochastic dynamics that are made manifest by the dilation, as represented by any new dilation-emergeables that arise, were always there all along, albeit in a non-manifest way.

An important example of this last application is *intrinsic spin*. To introduce spin as a dilation-emergeable, one merely dilates the Hilbert space to ND dimensions, introduces a D-dimensional representation of $SO(3)$ for the internal Hilbert space, and then requires that the dilated time-evolution operator has the appropriate form of rotation symmetry. This approach to representing spin ensures that despite picking an arbitrary three-dimensional coordinate axis in the process of formally carrying out the dilation of the Hilbert space—such as by choosing the spin-z operator to be diagonal on the dilated Hilbert space—the underlying generalized stochastic system does not fundamentally involve any preferred direction or entail any basic violation of rotation invariance.

C. Nonlocality

This paper has shown that systems based on trajectories in configuration spaces and evolving according to generically indivisible stochastic dynamics have Hilbert-space representations and can replicate the usual mathematical formalism and empirical predictions of quantum theory.

Technically speaking, the configurations in this new picture for quantum theory play the role of *hidden variables*, meaning physical parameters that exist separately from wave functions and density matrices. Of course, one could argue that configurations should more properly be called *physical variables*, given that wave functions and density matrices arise from the stochastic-quantum correspondence merely as secondary, representational constructs, rather than as fundamental entities in their own right. Either way, any mention of 'hidden variables' immediately raises questions about the potential invocation of *nonlocality*, the study of which has motivated famous papers like that of Einstein, Podolsky, and Rosen [70], and has led to the development of a number of important *no-go theorems* [71–74].

Before assessing the implications of these no-go theorems for the picture described in this paper, it will be important to note that these theorems do not rule out the possibility of hidden variables altogether. Nor do these theorems imply that introducing hidden variables would necessarily make quantum theory any more dynamically nonlocal than it already is.

Being mindful of these caveats, there is ample reason to probe the question of nonlocal dynamics in the approach to quantum theory taken in this paper. After all, looking back at the discussion of entanglement in Subsection IV H, a pair of systems that interact at some time will

generically exhibit what look like nonlocal stochastic dynamics after that time, at least until the later occurrence of a division event due to decoherence by an external system.

In what follows, it will be important to be keep in mind the distinction between *deterministic* hidden-variables theories and *stochastic* hidden-variables theories.

Bell's original nonlocality theorem, as formulated and proved in 1964 [71], only addressed the case of a *deterministic* hidden-variables theory. Specifically, Bell showed that if one assumes that a theory's hidden variables uniquely determine measurement outcomes, and if one further assumes that the hidden variables are local in the sense that measurement results should not depend on the settings of faraway measuring devices, then one arrives at an inequality that is expressly violated by quantum theory. Bell's 1964 theorem therefore establishes that any purported formulation of quantum theory based on local deterministic hidden variables is ruled out empirically.

At first glance, there might have seemed to be just two available options in response to Bell's nonlocality theorem. Either one could accept a theory of *nonlocal* deterministic hidden variables, or one could *deny* the existence of nonlocal deterministic hidden variables and thereby try to avoid having to introduce any ostensibly new dynamical nonlocality into quantum theory.

However, for a hidden-variables theory based on *stochastic* dynamics rather than on *deterministic* dynamics, the question of dynamical nonlocality becomes murkier. The generalization to stochastic dynamics means that one needs to rely on more abstract, statistical conditions for establishing whether or not the theory's hidden variables behave in a dynamically local manner.

The most frequently cited statistical locality criterion for stochastic hidden-variables theories was formulated by Bell later on, in 1975 [73, 75, 76]. That statistical locality criterion is a statement about how rich a theory's hidden variables should be in order for the theory to qualify as dynamically local.

To formulate Bell's statistical locality criterion, one starts by considering the case of a measurement outcome x based on local measurement settings a, and a far-separated measurement outcome y based on local measurement settings b. Then one supposes that the joint probabilities $p(x, y|a, b)$ for the measurement results x and y, conditioned on the measurement settings a and b, show a statistical correlation. Bell argued that in order for the theory in question to be considered dynamically local, the theory should contain enough hidden variables to account for the statistical correlation in the following precise sense: if one conditions on all the hidden variables λ in the past light cone of the two measurements, then those hidden variables should screen off the correlation between the measurement results, meaning that the joint probabilities should factorize according to

$$p(x, y|a, b, \lambda) = p(x|a, \lambda)p(y|b, \lambda). \tag{163}$$

Bell's statistical locality criterion is precisely the condition that the theory in question should have enough hidden variables to ensure that the factorization (163) is always possible. Based on this statistical locality criterion, which should hold even in cases of 'one-shot' measurements in which certain measurement outcomes can be assigned a 100% probability [74], one can again derive predictions that are violated by quantum theory, just as in the case of a deterministic

hidden-variables theory.

However, Bell's statistical locality criterion is broader than the conditions he studied in his 1964 theorem on deterministic hidden-variables theories. Bell's statistical locality condition is so broad, in fact, that Bell used it to argue that textbook quantum theory is *itself* already dynamically nonlocal [73, 77].

To understand why, observe that textbook quantum theory is committed to the existence of measurement settings and definite measurement outcomes that end up behaving precisely as a (highly incomplete) set of stochastically evolving hidden variables. In other words, although textbook quantum theory is not a *deterministic* hidden-variables theory, it is, in fact, a *stochastic* hidden-variables theory.

The stochastic-quantum correspondence derived in this paper makes these commitments by textbook quantum theory manifest. Indeed, one can regard textbook quantum theory as the insistence that for any measurement set-up consisting of a subject system $\mathcal{S}$, a measuring device $\mathcal{D}$, and an environment $\mathcal{E}$, as laid out in Subsection V B, the configurations of $\mathcal{D}$ are to be treated as hidden variables (that is, as beables), whereas the configurations of $\mathcal{S}$ and $\mathcal{E}$ are to be regarded merely as emergeables.

The stochastic-quantum correspondence therefore brings newfound clarity to the outstanding foundational problems of textbook quantum theory. Specifically, one sees from this novel perspective that the seemingly arbitrary division of the world into measuring devices, which purportedly have underlying configurations, and all other systems, which purportedly do not, leads directly to all the usual mysteries about the measurement process according to textbook quantum theory. After all, what, in the end, determines whether a given system counts as a measuring device, and therefore merits having underlying configurations?

More relevant to the present discussion is that because textbook quantum theory includes stochastic hidden variables for measuring devices, and because those stochastic hidden variables are insufficient to ensure the factorization property (163), the nonlocality theorems that employ Bell's statistical locality criterion imply that textbook quantum theory is itself dynamically nonlocal. Hence, there is no real cost to upgrading the configurations of $\mathcal{S}$ and $\mathcal{E}$ to being hidden variables on an equal footing with the configurations of $\mathcal{D}$. These additional hidden variables do not lead to the factorization property (163) either, but they also do not lead to any trouble for the *no-communication theorem* [78, 79], which precludes using quantum theory to send controllable signals faster than light.

The main conclusion of this analysis is that if one takes Bell's statistical locality criterion seriously as the proper way to identify dynamical nonlocality, then textbook quantum theory is already dynamically nonlocal, so adding some additional hidden variables to the theory will not ultimately make that dynamical nonlocality any worse. Alternatively, one could instead dispute that Bell's statistical locality criterion is an acceptable criterion for locality in the first place, perhaps by arguing that local correlation-producing interactions are valid common-causes explanations, but are simply not the sorts of things that can be conditioned on. However, if one denies the validity of Bell's criterion, then it cannot be used to argue that the picture of quantum theory presented in this paper is dynamically nonlocal. Either way, the approach taken toward quantum theory in

this paper is no more or less dynamically nonlocal than textbook quantum theory already is.

It may be that a more empirically meaningful notion of dynamical locality is the condition that new statistical correlations can only arise between systems that are in local contact, either directly or through the mediation of other systems. As the no-communication theorem shows, this empirical locality criterion is satisfied by textbook quantum theory, as well as by any physical theory or interpretative framework that reproduces the predictions of textbook quantum theory, and ultimately ensures the impossibility of sending controllable messages faster than light. One could argue that any stronger locality requirements, such as Bell's statistical locality condition, go beyond what is empirically verifiable, and are therefore matters of metaphysics.

A number of other important no-go theorems have been proved over the years, including von Neumann's early no-go theorem [17, 80], the Kochen-Specker theorem [81], the Pusey-Barrett-Rudolph theorem [82], and Myrvold's no-go theorem [83]. These theorems either assume that *all* observables are true random variables (that is, beables) that exist at the level of the given system's configuration space, or they assume that measurements are *passive operations* that merely reveal pre-existing values of observables without altering the behavior of measured systems in the process, or they assume that measurements are fundamentally *irreversible*, or they assume the existence of *additional* probability formulas. Because the picture of quantum theory introduced in this paper refrains from making any of these assumptions, it is consistent with these theorems.

VII. DISCUSSION AND FUTURE WORK

This paper has shown that one can reconstruct the mathematical formalism and all the empirical predictions of quantum theory using simpler, more physically transparent axioms than the standard Dirac-von Neumann axioms. Rather than postulating Hilbert spaces and their ingredients from the beginning, one instead posits a physical model, called a *generalized stochastic system*, based on trajectories in configuration spaces following generically indivisible stochastic dynamics. The *stochastic-quantum correspondence* then connects generalized stochastic systems with quantum systems in a fundamental way, showing that every quantum system can be viewed as the Hilbert-space representation of an underlying generalized stochastic system.

This perspective deflates some of the most mysterious features of quantum theory. In particular, one sees that density matrices, wave functions, and all the other appurtenances of Hilbert spaces, while highly useful, are merely gauge variables. These appurtenances should therefore not be assigned direct physical meanings or treated as though they directly represent physical objects, any more than Lagrangians or Hamilton's principal functions directly represent physical objects. Superposition is then not a literal smearing of physical objects, but is merely a mathematical artifact of catching a system in the middle of an indivisible stochastic process, as represented using a Hilbert-space formulation and wave functions.

Moreover, from this standpoint, *canonical quantization* need not be regarded as the promotion of classical observables to noncommutative operators by fiat, but can be implemented (when mathematically feasible) simply by generalizing a classical system's dynamics from being deterministic to being stochastic, with all the exotic features of quantum theory then emerging automatically. As

a consequence, this formulation of canonical quantization potentially offers more straightforward techniques for coupling classical systems to quantum systems in real-world applications.

In an important sense, the stochastic-quantum correspondence also legitimizes many standard practices followed in physics and in other scientific areas, like astronomy, chemistry, biology, and paleontology. To see why, notice that according to the thoroughly instrumentalist and operationalist Dirac-von Neumann axioms, the *only* predictions provided by textbook quantum theory are predictions of measurement outcomes, probabilities of measurement outcomes, and expectation values that are averages of measurement outcomes statistically weighted by measurement-outcome probabilities [36–40]. However, scientists in all areas of research talk about other kinds of phenomena that presumably just *happen* in some way, according to *happening* probabilities, in the past, present, or future. Strictly speaking, however, the *happening* of phenomena lies outside the axiomatic ambit of textbook quantum theory. The inability of textbook quantum theory to account for the happening of phenomena either means that scientists are not speaking honestly or coherently about their research, or that textbook quantum theory is inadequate as a physical theory.

Decoherence alone cannot bridge the gap between measurement-outcome probabilities and happening probabilities, because decoherence can only temporarily change whatever orthonormal basis momentarily diagonalizes a system's density matrix (and, after all, every density matrix is always diagonal in *some* orthonormal basis). After a system undergoes decoherence, textbook quantum theory then still requires one to make a direct appeal to the measurement axioms to translate the final density matrix into a statement about probabilities, which will then axiomatically end up being measurement-outcome probabilities, rather than happening probabilities.

Nor can appealing to some sort of *thermodynamic limit* resolve the discrepancy either. In order for a limit in a physical context to make sense, there should be clearly physical elementary ingredients or components involved. Furthermore, the end result of the limit should gradually emerge as a better and better physical approximation at *finite* stages of the limiting process, simply because a rigorous limit consists of inequalities between finite (if arbitrarily large or small) parameters. For example, in the *hydrodynamic limit* of a system of classical interacting particles, the particles are the physical elementary ingredients, and one sees fluid-like behavior gradually emerge as a better and better physical approximation as the number of particles progressively increases. In the case of textbook quantum theory, by contrast, every finite stage of any purported thermodynamic limit features only measurement outcomes and measurement-outcome probabilities, so there are no clearly physical elementary ingredients, and the gap between measurement outcomes and the happening of phenomena never closes.

The stochastic-quantum correspondence yields a much richer version of quantum theory in which physical phenomena really happen, with probabilities that are really happening probabilities, and therefore vindicates the ways that scientists talk about the world. Measurement-outcome probabilities are then merely a special case, arising when what is actually happening is a change to the configuration of a measuring device.

Because this paper's approach invokes hidden variables in the form of underlying physical configurations, this framework for quantum theory shares some aspects with the *de Broglie-Bohm for-*

mulation, or *Bohmian mechanics* [84–86]. However, in contrast to this paper's approach, Bohmian mechanics employs deterministic dynamics, and features a fundamental guiding equation that explicitly breaks Lorentz invariance by singling out a preferred foliation of spacetime into spacelike hypersurfaces. This paper instead takes seriously what experiments strongly suggest—that the dynamics of quantum theory is indeterministic, and that there is no fundamentally preferred foliation of spacetime. The formulation of quantum theory in this paper is also more flexible and model-independent than Bohmian mechanics, and works for all kinds of quantum systems, from particles to fields and beyond.

In contrast with the *Everett interpretation* [87, 88], also known as the *'many worlds' interpretation*, the framework presented in this paper assumes that quantum systems, like classical systems, have definite configurations in configuration spaces, and does not attempt to derive probability from non-probabilistic assumptions or grapple with fundamental aspects of personal identity in a universe continuously branching into large (and somewhat undefined) numbers of parallel worlds. The approach in this paper is therefore more modest, metaphysically speaking, than the Everett interpretation.

Neither this paper's approach nor the Everett interpretation satisfies the statistical locality criterion formulated by Bell and described in Subsection VI C, but the Everett interpretation arguably exhibits a different notion of dynamical locality at a level of description that *transcends* its individual world-branches [89]. However, because each individual world-branch looks no more or less nonlocal than the world according to textbook quantum theory, it is not clear whether the Everett interpretation's dynamical locality is more than a metaphysical statement, nor is it easy to discern what concrete benefits it truly provides.

Unlike *stochastic-collapse theories* [14, 15], this paper does not invoke any fundamental violations of unitarity, nor does it require introducing any new constants of nature to specify dynamical-collapse rates.

Looking forward, it would be interesting to see what implications the stochastic-quantum correspondence could have for both phenomenological stochastic processes, like those in biology or finance, as well as for future work in fundamental physics, like quantum gravity.

More broadly, by recasting the Hilbert-space formulation of quantum theory as merely a convenient way to represent a large class of stochastic processes, one opens the door to searching for totally different representations that might look nothing at all like Hilbert spaces and that could allow for the construction of more general kinds of stochastic processes. Perhaps one could even find a way to generalize beyond stochastic processes altogether. work.

ACKNOWLEDGMENTS

The author would especially like to acknowledge Emily Adlam, David Albert, Howard Georgi, David Kagan, and Logan McCarty for extensive discussions during the writing of this paper. The author would also like to thank Scott Aaronson, Guido Bacciagaluppi, David Baker, Craig Callender, Ignacio Cirac, Iris Cong, Jordan Cotler, Erik Curiel, Louis Deslauriers, Melissa Franklin, Philip Goyal, David Griffiths, Mina Himwich, Jenann Ismael, Daniel Jafferis, David Kaiser, Efthimios

Kaxiras, Gregory Kestin, Matthew Kleban, Serhii Kryhin, Barry Loewer, Curtis McMullen, Simon Milz, Arjun Mirani, Wayne Myrvold, Filip Niewinski, Jill North, John Norton, Matthew Reece, Vivishek Sudhir, Noel Swanson, Xi Yin, and Nicole Yunger Halpern for helpful conversations.

---

[1] E. C. G. Sudarshan, P. M. Mathews, and J. Rau. "Stochastic Dynamics of Quantum-Mechanical Systems". *Physical Review*, 121:920–924, February 1961. URL: http://link.aps.org/doi/10.1103/PhysRev.121.920, doi:10.1103/PhysRev.121.920.

[2] G. A. Skorobogatov and S. I. Svertilov. "Quantum Mechanics can be Formulated as a Non-Markovian Stochastic Process". *Physical Review A*, 58(5):3426–3432, November 1998. doi:10.1103/PhysRevA.58.3426.

[3] D. T. Gillespie. "Non-Markovian Stochastic Processes". In *AIP Conference Proceedings*, Volume 511, pages 49–56. American Institute of Physics, July 2000. doi:10.1063/1.60002.

[4] D. T. Gillespie, W. O. Alltop, and J. M. Martin. "Modeling Quantum Measurement Probability as a Classical Stochastic Process". *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(3):548–562, August 2001. doi:10.1063/1.1378791.

[5] S. Aaronson. "Quantum Computing and Hidden Variables". *Physical Review A*, 71(3):032325, March 2005. doi:10.1103/PhysRevA.71.032325.

[6] F. Strocchi. *An Introduction to the Mathematical Structure of Quantum Mechanics*. World Scientific, 2nd edition, 2008.

[7] M. Frasca. "Quantum Mechanics is the Square Root of a Stochastic Process". *arXiv*, 2012. URL: http://arxiv.org/abs/1201.5091, arXiv:1201.5091.

[8] F. A. Bopp. "Quantenmechanische Statistik und Korrelationsrechnung". *Zeitschrift für Naturforschung A*, 2(4):202–216, 1947. doi:10.1515/zna-1947-0402.

[9] F. A. Bopp. "Ein für die Quantenmechanik bemerkenswerter Satz der Korrelationsrechnung". *Zeitschrift für Naturforschung A*, 7(1):82–87, 1952. doi:10.1515/zna-1952-0117.

[10] F. A. Bopp. "Statistische Untersuchung des Grundprozesses der Quantentheorie der Elementarteilchen". *Zeitschrift für Naturforschung A*, 8(1):6–13, 1953. doi:10.1515/zna-1953-0103.

[11] I. Fényes. "Eine wahrscheinlichkeitstheoretische Begründung und Interpretation der Quantenmechanik". *Zeitschrift für Physik*, 132(1):81–106, February 1952. doi:10.1007/BF01338578.

[12] E. Nelson. "Dynamical Theories of Brownian Motion". 1967.

[13] E. Nelson. "Quantum Fluctuations". 1985.

[14] G. C. Ghirardi, A. Rimini, and T. Weber. "Unified Dynamics for Microscopic and Macroscopic Systems". *Physical Review D*, 34(2):470–491, 1986. doi:10.1103/PhysRevD.34.470.

[15] A. Bassi and G. C. Ghirardi. "Dynamical reduction models". *Physical Reports*, 379(5):257–426, 2003. URL: http://www.sciencedirect.com/science/article/pii/S0370157303001030, arXiv:quant-ph/0302164, doi:10.1016/S0370-1573(03)00103-0.

[16] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 1st edition, 1930.

[17] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer, 1932.

[18] P. A. M. Dirac. "The Lagrangian in Quantum Mechanics". *Physikalische Zeitschrift der Sowjetunion*, 3(1):64–72, 1933.

[19] R. P. Feynman. *The Principle of Least Action in Quantum Mechanics*. PhD thesis, Princeton University, Princeton, NJ, 1942.

[20] R. P. Feynman. "Space-Time Approach to Non-Relativistic Quantum Mechanics". *Reviews of modern physics*, 20(2):367, April 1948. doi:10.1103/RevModPhys.20.367.

[21] E. P. Wigner. "On the Quantum Correction For Thermodynamic Equilibrium". *Physical Review*, 40(5):749–759, June 1932. doi:10.1103/PhysRev.40.749.

[22] J. E. Moyal. "Quantum Mechanics as a Statistical Theory". *Mathematical Proceedings of the Cambridge Philosophical Society*, 45(1):99–124, January 1949. doi:10.1017/S0305004100000487.

[23] C. Wetterich. "Quantum Representation of Classical Statistical Ensembles". Unpublished draft, 2009. URL: https://www.thphys.uni-heidelberg.de/~wetterich/Material/quantum_representation.pdf.

[24] M. Rosenblatt. *Random Processes*. Oxford University Press, 1962.

[25] E. Parzen. *Stochastic Processes*. Holden-Day, 1962.

[26] J. L. Doob. *Stochastic Processes*. Wiley-Interscience, 1990.

[27] S. Ross. *Stochastic Processes*. John Wiley and Sons, 2nd edition, 1995.

[28] S. Milz and K. Modi. "Quantum Stochastic Processes and Quantum Non-Markovian Phenomena". *PRX Quantum*, 2:030201, May 2021. URL: https://dx.doi.org/10.1103/PRXQuantum.2.030201, arXiv:2012.01894v2, doi:10.1103/PRXQuantum.2.030201.

[29] M. M. Wolf and J. I. Cirac. "Dividing Quantum Channels". *Communications in Mathematical Physics*, 279:147–168, 2008. arXiv:math-ph/0611057, doi:10.1007/s00220-008-0411-y.

[30] I. Schur. "Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen". *Journal für die reine und angewandte Mathematik*, 140:1–28, 1911. URL: https://eudml.org/doc/149352.

[31] P. R. Halmos. *Finite-Dimensional Vector Spaces*. Van Nostrand, 2nd edition, 1958.

[32] R. A. Horn. "The Hadamard Product". In *Matrix Theory and Applications*, Volume 40, pages 87–169, 1990. doi:10.1090/psapm/040/1059485.

[33] G. W. Mackey. "Induced Representations of Locally Compact Groups I". *Annals of Mathematics*, 55(1):101–139, 1952. URL: http://www.jstor.org/stable/1969423, doi:10.2307/1969423.

[34] G. W. Mackey. "Quantum Mechanics and Hilbert Space". *The American Mathematical Monthly*, 64(8):45–57, 1957. URL: http://www.jstor.org/stable/2308516, doi:10.2307/2308516.

[35] A. Auffeves and P. Grangier. "Recovering the Quantum Formalism from Physically Realist Axioms". *Scientific Reports*, 7:43365, Dec. 2017. URL: http://doi.org/10.1038/srep43365, arXiv:1610.06164v2, doi:10.1038/srep43365.

[36] D. J. Griffiths and D. F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, 3rd edition, 2018.

[37] J. S. Townsend. *A Modern Approach to Quantum Mechanics*. Cambridge University Press, 2nd edition, 2012.

[38] R. Shankar. *Principles of Quantum Mechanics*. Plenum Press, 2nd edition, 1994.

[39] J. J. Sakurai and J. J. Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, 2nd edition, 2010.

[40] B. Schumacher and M. Westmoreland. *Quantum Processes, Systems, and Information*. Cambridge University Press, 1st edition, 2010.

[41] D. J. Griffiths. *Introduction to Electrodynamics*. Cambridge University Press, 4th edition, 2017.

[42] A. Zangwill. *Modern Electrodynamics*. Cambridge University Press, 2012.

[43] J. D. Jackson. *Classical Electrodynamics*. Wiley, 3rd edition, 1998.

[44] H. Brown. "Aspects of Objectivity in Quantum Mechanics". In *From Physics to Philosophy*, pages 45–70. Cambridge University Press, 1999. URL: https://philpapers.org/rec/BROAOO-2.

[45] K. Kraus. "General State Changes in Quantum Theory". *Annals of Physics*, 64(2):311–335, June 1971. URL: http://www.sciencedirect.com/science/article/pii/0003491671901084, doi:10.1016/0003-4916(71)90108-4.

[46] J. A. Barandes and D. Kagan. "Quantum Conditional Probabilities and New Measures of Quantum Information". *Annals of Physics*, 448:169192, January 2023. arXiv:2109.07447, doi:10.1016/j.aop.2022.169192.

[47] A. Horn. "Doubly Stochastic Matrices and the Diagonal of a Rotation Matrix". *American Journal of Mathematics*, 76(3):620–630, 1954. doi:10.2307/2372705.

[48] R. C. Thompson. "Lecture notes from a Johns Hopkins University lecture series". Unpublished lecture notes, 1989.

[49] P. Nylen, T.-Y. Tam, and F. Uhlig. "On the Eigenvalues of Principal Submatrices of Normal, Hermitian and Symmetric Matrices". *Linear and Multilinear Algebra*, 36(1):69–78, 1993. doi:10.1080/03081089308818276.

[50] J. A. Barandes. "The Stochastic-Quantum Theorem". In preparation.

[51] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[52] S. Weinberg. *The Quantum Theory of Fields*, Volume 2. Cambridge University Press, 1996.

[53] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics, Volume 3*. Addison-Wesley, 1965. URL: https://www.feynmanlectures.caltech.edu/III_toc.html.

[54] E. Schrödinger. "An Undulatory Theory of the Mechanics of Atoms and Molecules". *Physical Review*, 28(6):1049–1070, December 1926. doi:10.1103/PhysRev.28.1049.

[55] S. Weinberg. *The Quantum Theory of Fields*, Volume 1. Cambridge University Press, 1996.

[56] J. R. Glick and C. Adami. "Markovian and Non-Markovian Quantum Measurements". *Foundations of Physics*, July 2020. arXiv:1701.05636, doi:10.1007/s10701-020-00362-4.

[57] J. S. Bell. "Subject and Object". 1972. URL: https://cds.cern.ch/record/610096/files/CM-P00058496.pdf.

[58] N. Bohr. "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?". *Physical review*, 48(8):696, 1935. doi:10.1103/PhysRev.48.696.

[59] J. S. Bell. "Introduction to the Hidden-Variable Question". 1971. URL: http://cds.cern.ch/record/400330.

[60] W. Myrvold. "Philosophical Issues in Quantum Theory". In E. N. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2022 edition, 2022. URL: https://plato.stanford.edu/archives/fall2022/entries/qt-issues/.

[61] R. H. Dicke. "Quantum Measurements, Sequential and Latent". *Foundations of Physics*, 19(4):385–395, April 1989. doi:10.1007/BF00731832.

[62] W. Heisenberg. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik ('About the Descriptive Content of Quantum Theoretical Kinematics and Mechanics')". *Zeitschrift für Physik*, 43(3):172–198, 1927. doi:10.1007/BF01397280.

[63] H. P. Robertson. "The Uncertainty Principle". *Physical Review*, 34(1):163–164, July 1929. doi:10.1103/PhysRev.34.163.

[64] E. C. G. Stueckelberg. "Quantum Theory in Real Hilbert Space". *Helvetica Physica Acta*, 33(4):727–752, 1960. URL: https://www.e-periodica.ch/digbib/view?pid=hpa-001:1960:33::715#735.

[65] R. Haag. *Local Quantum Physics: Fields, Particles, Algebras*. Springer-Verlag, 1993.

[66] B. H. Feintzeig. "On Algebraic Methods in Quantum Theories". *PhD Thesis, University of California, Irvine*, 2016. URL: https://escholarship.org/uc/item/42t27123.

[67] J. Myrheim. "Quantum Mechanics on a Real Hilbert Space". 1999. URL: https://arxiv.org/abs/quant-ph/9905037, arXiv:quant-ph/9905037.

[68] W. F. Stinespring. "Positive functions on C*-algebras". *Proceedings of the American Mathematical Society*, 6(2):211–216, April 1955. doi:10.2307/2032342.

[69] M. Keyl. "Fundamentals of Quantum Information Theory". *Physics reports*, 369(5):431–548, October 2002. arXiv:0202122, doi:10.1016/S0370-1573(02)00266-1.

[70] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?". *Physical Review*, 47(10):777–780, May 1935. doi:10.1103/PhysRev.47.777.

[71] J. S. Bell. "On the Einstein-Podolsky-Rosen Paradox". *Physics*, 1(3):195–200, 1964.

[72] J. F. Clauser, M. A. Horne, A. E. Shimony, and R. A. Holt. "Proposed Experiment to Test Local Hidden-Variable Theories". *Physical Review Letters*, 23(15):880–884, October 1969. doi:10.1103/PhysRevLett.23.880.

[73] J. S. Bell. "The Theory of Local Beables". *CERN*, 1975. URL: https://cds.cern.ch/record/980036/files/197508125.pdf.

[74] D. M. Greenberger, M. A. Horne, and A. Zeilinger. "Going Beyond Bell's Theorem". In *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, Fundamental Theories of Physics, pages 69–72. Springer, 1989. arXiv:0712.0921, doi:10.1007/978-94-017-0849-4_10.

[75] J. Butterfield. "Bell's Theorem: What it Takes". *The British Journal for the Philosophy of Science*, 43(1):41–83, March 1992. URL: https://www.jstor.org/stable/687884, doi:10.1093/bjps/43.1.41.

[76] W. Myrvold, M. Genovese, and A. Shimony. "Bell's Theorem". In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2021 edition, 2021. URL: https://plato.stanford.edu/archives/fall2021/entries/bell-theorem/.

[77] J. S. Bell. "La Nouvelle Cuisine". In A. Sarlemijn and P. Kroes, editors, *Between Science and Technology*, pages 97–115. Elsevier, 1990.

[78] G. C. Ghirardi, A. Rimini, and T. Weber. "A General Argument Against Superluminal Transmission Through the Quantum Mechanical Measurement Process". *Lettere al Nuovo Cimento*, 27(10):293–298, 1980. doi:10.1007/BF02817189.

[79] T. F. Jordan. "Quantum Correlations do not Transmit Signals". *Physics Letters A*, 94(6):264, 1983. doi:10.1016/0375-9601(83)90713-2.

[80] J. von Neumann. "Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik". *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1927:245–272, 1927. URL: https://eudml.org/doc/59230.

[81] S. B. Kochen and E. P. Specker. "The Problem of Hidden Variables in Quantum Mechanics". *Indiana University Mathematics Journal*, 17(1):59–87, 1967.

[82] M. F. Pusey, J. Barrett, and T. Rudolph. "On the Reality of the Quantum State". *Nature Physics*, 8(6):475–478, 2012. arXiv:1111.3328, doi:10.1038/nphys2309.

[83] W. C. Myrvold. "Modal Interpretations and Relativity". *Foundations of Physics*, 32(11):1773–1784, 2002. URL: https://doi.org/10.1023/A%3A1021406924313, arXiv:quant-ph/0209109, doi:10.1023/A:1021406924313.

[84] L. de Broglie. *An Introduction to the Study of Wave Mechanics*. E. P. Dutton and Company, Inc., 1930.

[85] D. J. Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. I". *Physical Review*, 85(2):166–179, January 1952. doi:10.1103/PhysRev.85.166.

[86] D. J. Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. II". *Physical Review*, 85(2):180–193, January 1952. doi:10.1103/PhysRev.85.180.

[87] H. Everett III. " 'Relative State' Formulation of Quantum Mechanics". *Reviews of Modern Physics*, 29(3):454–462, July 1957. doi:10.1103/RevModPhys.29.454.

[88] H. Everett III. "The Theory of the Universal Wave Function". In *The Many-Worlds Interpretation of Quantum Mechanics*, Volume 1, page 3, 1973.

[89] D. Wallace. *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. Oxford University Press, 2012.