The Stochastic-Quantum Correspondence

Jacob A. Barandes[1,*]

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

(Dated: February 19, 2023)

This paper introduces a precise correspondence between the theory of stochastic processes and quantum theory. This correspondence provides a new framework for using Hilbert-space methods to formulate highly generic types of stochastic dynamics, with potential applications throughout the sciences. This paper also uses this correspondence in the other direction to reconstruct quantum theory in general from physical models that consist of classical kinematics combined with stochastic dynamics. This reconstruction approach opens up new ways of understanding quantum-theoretic 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 in time according to probabilistic laws. 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 literature to identify a fundamental relationship connecting stochastic-processes theory and quantum theory [1–7]. This paper introduces a new and fully general correspondence between these two theories in the form of a simple 'dictionary' expressing any time-dependent stochastic matrix in terms of a suitable combination of Hilbert-space ingredients.

On the one hand, from a practical standpoint, this 'stochastic-quantum correspondence' provides a systematic framework for constructing highly generic forms of stochastic dynamics, much as the classical 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 examples. Importantly, this 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 any other frequently deployed approximations.

Taking a more foundational perspective, this paper also uses this stochastic-quantum correspondence to show that physical models based on classical kinematics 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 [8, 9]. In this alternative approach, systems have physical configurations in classical configuration spaces, and the mathematical objects of the Hilbert-space formulation serve a functional role akin to gauge-theoretic degrees of freedom.

In addition to establishing these new results, this paper identifies several forms of gauge invariance that have not previously been described in the literature, 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 hidden variables in hand, this paper also revisits and clarifies a number of important questions related to the status of nonlocality in quantum theory.

Given the mathematical simplicity of this stochastic-quantum correspondence, it is surprising that it has apparently not shown up in the 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 [6].[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 on classical configuration spaces.

* barandes@g.harvard.edu

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

II. STOCHASTIC PROCESSES

In the theory of stochastic processes [10], one starts with a configuration space $\mathcal{C}$ and a stochastic map $\Gamma(t)$ that acts linearly on probability distributions over $\mathcal{C}$ at an initial time $t = 0$ to yield corresponding probability distributions over $\mathcal{C}$ at other times $t \neq 0$. The formalism 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 probabilities at $t = 0$ can be denoted by

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

the probabilities at $t \neq 0$ can be denoted by

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

and the stochastic map consists of conditional probabilities

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

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 it is in its jth configuration at the time 0. (Note that no assumption is made here about whether $t > 0$ or $t < 0$.) Being probabilities, these quantities satisfy

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

and

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

Then from marginalization, one has the linear relationship

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

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

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

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

The conditions (6) on $\Gamma(t)$ identify it as a (left) stochastic matrix. On physical grounds, $\Gamma(t)$ will be assumed to satisfy the continuity condition that in the limit $t \to 0$, it approaches the $N \times N$ identity matrix $\mathbb{1}$:

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

Next, consider a random variable $A(t)$ with (not necessarily unique) 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 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 probability densities $p(y, 0)$ at $t = 0$ and $p(x, t)$ at $t \neq 0$, where x and y each symbolically denotes a set of real-valued coordinates. The linear 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 $t = 0$ as a special time. Section IX, however, will show that for systems in sufficiently strong contact with a noisy environment, $t = 0$ need not actually be a unique time, but will typically be only one of many times that play a similar role.

III. CONVENTIONAL APPROXIMATIONS

In textbook treatments of stochastic processes [10], one often introduces various approximations or simplifications of a system's time-dependent stochastic matrix

$\Gamma(t)$ to make it easier to construct and describe. A typical such approximation is to assume a discrete-time Markov chain, meaning that for some small but finite time scale Δt, one can express the time-dependent stochastic matrix $\Gamma(t = n\,\Delta t)$ at any integer number $n \geq 1$ of steps of duration Δt as n powers of a constant stochastic matrix Γ:

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

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

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

which is known as divisibility [11]. Here $\Gamma(t \leftarrow t')$ is likewise required to be 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 $\Gamma(t)$ to be a time-dependent permutation matrix, meaning a matrix whose rows and columns are permutations of the $N \times N$ identity matrix $\mathbb{1}$. In that case, $\Gamma(t)$ 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 a time-dependent, generically 'indivisible' $N \times N$ stochastic 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, indivisible examples, like the 2×2 stochastic 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.

It may not seem obvious how to construct smoothly time-dependent stochastic matrices $\Gamma(t)$ systematically, 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. Ideally, one immediate application would be making it possible to derive a self-contained theoretical justification for why the Markov and divisibility approximations work so well in many real-world cases.

IV. THE HILBERT-SPACE FORMULATION

This paper introduces a novel and highly general framework for formulating time-dependent stochastic matrices $\Gamma(t)$, conceptually akin to the Lagrangian or Hamiltonian frameworks for formulating deterministic dynamics for mechanical systems.

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

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

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, this paper will assume that $\Theta_{ij}(t)$ involves only the complex numbers.

On account of the general properties of $\Gamma(t)$ specified in (6), note that the matrix $\Theta(t)$ must satisfy

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

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, introduce the Schur-Hadamard product $\odot$, which is defined for arbitrary $N \times N$ matrices X and Y as entry-wise multiplication [12–14]:

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

One can then regard (17) as expressing the stochastic 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 not basis-independent. For the purposes of analyzing a given stochastic system, however, the system's configuration space $\mathcal{C}$ naturally singles out a specific basis, to be defined momentarily.

As an alternative approach that will turn out to have significant ramifications, start by defining an N-member collection of constant, diagonal $N \times N$ 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

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

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 these 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. The configuration projectors $P_1, \dots, P_N$ therefore constitute a projection-valued measure (PVM) [15, 16].

Letting tr( ) denote the usual matrix trace, one can then recast (17) instead 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' between the classical theory of 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.[2]

To understand what these mathematical tools are, introduce a set of $N \times 1$ column vectors $e_1, \dots, e_N$, where e_i has a 1 in its ith component and 0s in all its other components. That is, e_i has components

$$e_{i,j} = \delta_{ij}. \tag{27}$$

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

$$e_i^\dagger e_j = \delta_{ij}, \quad e_i e_i^\dagger = P_i, \tag{28}$$

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 $\mathcal{H}$, meaning a complete inner-product space over the complex numbers. The dictionary therefore provides a Hilbert-space formulation for constructing generic forms of stochastic dynamics.

Substituting the right-hand side of the dictionary (26) into the linear relationship (7) between the probabilities $p_j(0)$ at $t = 0$ and the probabilities $p_i(t)$ at $t \neq 0$, one finds that

$$p_i(t) = \mathrm{tr}(P_i \rho(t)), \tag{29}$$

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

$$\left.\begin{aligned} \rho(t) &\equiv \Theta(t) \left[\sum_{j=1}^{N} p_j(0) P_j \right] \Theta^\dagger(t) \\ &= \Theta(t) \mathrm{diag}(\dots, p_j(0), \dots) \Theta^\dagger(t) \\ &= \rho^\dagger(t), \\ &\qquad \mathrm{tr}(\rho(t)) = 1. \end{aligned}\right\} \tag{30}$$

Similarly, by substituting the formula (29) for $p_i(t)$ into the definition (10) of the expectation value of a random variable $A(t)$, one obtains

$$\langle A(t) \rangle = \mathrm{tr}(A(t)\rho(t)), \tag{31}$$

where $A(t)$ is now understood to be the diagonal $N \times N$ matrix defined as

$$A(t) \equiv \sum_{i=1}^{N} a_i(t) P_i = \mathrm{diag}(\dots, a_i(t), \dots). \tag{32}$$

In the special case in which the system's probability distribution at $t = 0$ is pure, meaning that one of the system's configurations j is occupied with probability 1, the system's probability vector at $t = 0$ is equal to the jth vector e_j in the configuration basis (27):

$$p(0) = e_j \quad [\text{pure}]. \tag{33}$$

Defining a unit-norm, $N \times 1$ column vector

$$\Psi(t) \equiv \Theta(t) e_j \quad \left[\Psi^\dagger(t)\Psi(t) = 1\right], \tag{34}$$

which is ultimately just the jth column of $\Theta(t)$, the ith component $\Psi_i(t)$ of $\Psi(t)$ is a purely law-like quantity equal to the specific matrix entry $\Theta_{ij}(t)$:

$$\Psi_i(t) = \Theta_{ij}(t). \tag{35}$$

It follows immediately that the self-adjoint matrix $\rho(t)$ defined in (30) is rank-one and has factorization

$$\rho(t) = \Psi(t)\Psi^\dagger(t) \quad [\text{pure}]. \tag{36}$$

[2] Similar-looking formulas appear in equations (3)–(6) of [17] as an intermediate step in proving a lemma used for conceptually different purposes.

The probability formula (29) then simplifies to

$$p_i(t) = |\Psi_i(t)|^2, \tag{37}$$

and the formula (31) for the expectation value of a random variable $A(t)$ becomes

$$\langle A(t)\rangle = \Psi^\dagger(t)A(t)\Psi(t). \tag{38}$$

Looking at all these results, one notices a striking resemblance to mathematical objects and formulas that are familiar from textbook quantum theory.[3] 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. The probability formulas (29) and (37) have the same form as the Born rule, and (31) and (38) have the same form as quantum-theoretic expectation values.

These formulas are all expressed in what would conventionally be called the Schrödinger picture. One could instead work in the Heisenberg picture, with the definitions

$$\left.\begin{aligned} \rho^H \equiv \rho(0), \quad \Psi^H \equiv \Psi(0), \\ A^H(t) \equiv \Theta^\dagger(t)A(t)\Theta(t), \end{aligned}\right\} \tag{39}$$

where $A^H(t)$ now includes both a possible explicit dependence on time through its magnitudes $a_i(t)$ as well as implicit dependence on time through the time-evolution operator $\Theta(t)$. The probability formula (29) would then become[4]

$$p_i(t) = \mathrm{tr}(P^H(t)\rho^H), \tag{40}$$

and the formula (31) for expectation values would become

$$\langle A(t)\rangle = \mathrm{tr}(A^H(t)\rho^H). \tag{41}$$

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 stochastic matrix $\Gamma(t)$, whose entries are conditional probabilities $p(i, t|j, 0)$, in accordance with (4). The mathematical objects $\Theta(t)$, $\rho(t)$, $\Psi(t)$, $A(t)$, despite being extremely useful, are not uniquely defined by $\mathcal{C}$ or $\Gamma(t)$.

[3] For pedagogical treatments of quantum theory, see [18–22].

[4] 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.

V. GAUGE TRANSFORMATIONS

To make this non-uniqueness more manifest, it will be helpful to introduce an analogy with the Maxwell theory of classical electromagnetism.[5]

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, and any one such choice for the potentials is called a gauge choice. 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, 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, start by observing that the Schur-Hadamard product of the time-evolution operator $\Theta(t)$ with a matrix of time-dependent phases $\exp(i\theta_{ij})$ is a transformation of $\Theta(t)$ with no physical effects, and therefore corresponds to a genuine form of gauge invariance:

$$\Theta(t) \mapsto \Theta(t) \odot \begin{pmatrix} e^{i\theta_{11}(t)} & e^{i\theta_{12}(t)} & \\ e^{i\theta_{21}(t)} & \ddots & \\ & & e^{i\theta_{NN}(t)} \end{pmatrix}. \tag{42}$$

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

$$\Theta_{ij}(t) \mapsto \Theta_{ij}(t)e^{i\theta_{ij}(t)}. \tag{43}$$

To the author's knowledge, this kind of gauge invariance, which could be called a 'Schur-Hadamard gauge transformation,' has not yet been described in the literature. It will turn out to play a key role in the analysis of dynamical symmetries that will be presented in Section XVI, and will be extended in an interesting way in the context of Hilbert-space dilations in Section XVII.

The Hilbert-space formulation has another form of gauge invariance, which appears to have first been written down in [26] 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

[5] For pedagogical treatments of classical electromagnetism, see [23–25].

stochastic matrix $\Gamma(t)$ as a whole unchanged:[6]

$$\left.\begin{aligned} \rho(t) &\mapsto \rho_V(t) \equiv V(t)\rho(t)V^\dagger(t), \\ \Psi(t) &\mapsto \Psi_V(t) \equiv V(t)\Psi(t), \\ A(t) &\mapsto A_V(t) \equiv V(t)A(t)V^\dagger(t), \\ \Theta(t) &\mapsto \Theta_V(t) \equiv V(t)\Theta(t)V^\dagger(0). \end{aligned}\right\} \tag{44}$$

If the unitary matrix $V(t)$ is time-independent, then the gauge transformation (44) is merely a change of 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.

VI. KRAUS DECOMPOSITIONS

In the most general case, a time-evolution operator $\Theta(t)$ may not satisfy any nontrivial constraints apart from (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:

$$K_\beta(t) \equiv \begin{pmatrix} 0 & \cdots & 0 & \Theta_{1\beta}(t) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \Theta_{N\beta}(t) & 0 & \cdots & 0 \end{pmatrix} \quad [\beta = 1, \ldots, N]. \tag{45}$$

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

$$K_{\beta,ij}(t) = \delta_{\beta j}\Theta_{ij}(t). \tag{46}$$

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

$$\sum_{\beta=1}^{N} K_\beta^\dagger(t)K_\beta(t) = \mathbb{1}, \tag{47}$$

so these matrices are called Kraus operators [27]. One can then write the dictionary (26) in an alternative form called a Kraus decomposition:

$$\Gamma_{ij}(t) = \sum_{\beta=1}^{N} \mathrm{tr}(K_\beta^\dagger(t)P_iK_\beta(t)P_j). \tag{48}$$

Like all the other mathematical objects in the Hilbert-space formulation, the Kraus operators $K_1(t), \ldots, K_N(t)$ are not unique. Notice also that any number of $N \times N$ matrices satisfying the Kraus identity (47) are guaranteed to yield a valid stochastic matrix $\Gamma(t)$ via the Kraus decomposition (48).[7]

VII. UNISTOCHASTIC DYNAMICS

In the most minimal case in which the stochastic 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

$$\Gamma(t) = \overline{U(t)} \odot U(t). \tag{49}$$

That is,

$$\Gamma_{ij}(t) = |U_{ij}(t)|^2, \tag{50}$$

or, equivalently, in dictionary form (26),

$$\Gamma_{ij}(t) = \mathrm{tr}(U^\dagger(t)P_iU(t)P_j). \tag{51}$$

The Kraus identity (47), meanwhile, reduces to the statement that $U(t)$ is unitary,

$$U^\dagger(t) = U^{-1}(t), \tag{52}$$

and $\Gamma(t)$ is then said to be unistochastic [29, 30].[8] It follows immediately from the dictionary formula (51) that every unistochastic matrix is doubly stochastic, meaning that summing over its rows or its columns yields 1:

$$\sum_{i=1}^{N} \Gamma_{ij}(t) = \sum_{j=1}^{N} \Gamma_{ij}(t) = 1. \tag{53}$$

Note that $U(t)$ will not generically remain unitary under Schur-Hadamard gauge transformations (42), so writing a unistochastic matrix $\Gamma(t)$ in terms of a unitary time-evolution operator $U(t)$ corresponds to making a gauge

[6] Note the appearance of $t = 0$ in $V^\dagger(0)$ in the transformation rule for $\Theta(t)$.

[7] Kraus operators and Kraus decompositions play an important role in quantum information theory. They provide (non-unique) expressions for specific generalizations of unitary time evolution known as quantum channels, or completely positive trace-preserving (CPTP) maps. In particular, conditional probabilities similar in form to (48) were studied in [28].

[8] In Section XVII, it will be shown that all stochastic matrices can be expressed in terms of a unitary time-evolution operator on a suitably enlarged or dilated Hilbert space, so assuming unistochastic dynamics is not as special a condition as it might seem.

choice—or, somewhat more precisely, to partially fixing the gauge freedom (42). Notice also that every permutation matrix is unitary, so deterministic dynamics is a special case of unistochastic dynamics.

Assuming that $U(t)$ 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

$$H(t) \equiv i\hbar \frac{\partial U(t)}{\partial t} U^\dagger(t) = H^\dagger(t). \tag{54}$$

Here the factor of i ensures that $H(t)$ is self-adjoint, and, for present purposes, $\hbar$ is a fixed constant introduced for purposes of units.

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

$$i\hbar \frac{\partial \rho(t)}{\partial t} = [H(t), \rho(t)], \tag{55}$$

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

$$i\hbar \frac{\partial \Psi(t)}{\partial t} = H(t)\Psi(t), \tag{56}$$

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

$$\frac{dA^H(t)}{dt} = \frac{i}{\hbar}[H^H(t), A^H(t)] + \left(\frac{\partial A(t)}{\partial t}\right)^H, \tag{57}$$

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

$$\frac{d\langle A(t)\rangle}{dt} = \frac{i}{\hbar}\mathrm{tr}([H(t), A(t)]\rho(t)) + \left\langle \frac{\partial A(t)}{\partial t} \right\rangle. \tag{58}$$

The matrix $H^H(t)$ appearing in the Heisenberg equation of motion (57) is the Hamiltonian in the Heisenberg picture. 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 classical kinematics—with a classical configuration space $\mathcal{C}$—is a surprising new result.

If the system's time-evolution operator $\Theta(t) = U(t)$ is indeed unitary, then under the unitary gauge transformation defined by (44), the Hamiltonian transforms precisely as a non-Abelian gauge potential:[9]

$$\left.\begin{aligned} H(t) &\mapsto H_V(t) \\ &= V(t)H(t)V^\dagger(t) - i\hbar V(t)\frac{\partial V^\dagger(t)}{\partial t}. \end{aligned}\right\} \tag{59}$$

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.

Moreover, one can write the Schrödinger equation (56) as

$$\mathcal{D}(t)\Psi(t) = 0. \tag{60}$$

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

$$\mathcal{D}(t) \equiv \mathbb{1}\frac{\partial}{\partial t} + \frac{i}{\hbar}H(t). \tag{61}$$

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

Notice that if one picks

$$V(t) \equiv U^\dagger(t), \tag{62}$$

then the Hamiltonian precisely vanishes:

$$H_V(t) = 0. \tag{63}$$

This choice of gauge is nothing other than the definition (39) of the Heisenberg picture. Unitary gauge transformations (44) can therefore be viewed as generalized changes of time-evolution picture.[10]

VIII. INTERFERENCE

The appearance of the Schrödinger equation (56) 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 stochastic dynamics might have the resources to replicate the features of quantum theory more broadly.

As another hint pointing in this direction, start by noting that an arbitrary time-dependent stochastic 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, suppose that at some time t', $\Gamma(t')$ has a matrix inverse $\Gamma^{-1}(t')$, and let

$$\tilde{\Gamma}(t \leftarrow t') \equiv \Gamma(t)\Gamma^{-1}(t'). \tag{64}$$

[9] For pedagogical treatments of non-Abelian gauge theories, see [31, 32].

[10] 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 over a one-dimensional base manifold parameterized by the time t, has vanishing curvature.

As an immediate consequence, one then has

$$\Gamma(t) = \tilde{\Gamma}(t \leftarrow t')\Gamma(t'), \tag{65}$$

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.[11] Hence, the matrix $\tilde{\Gamma}(t \leftarrow t')$ defined in (64) is not generically stochastic, so (65) 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 stochastic matrix $\Gamma(t)$. To this end, suppose that $\Gamma(t)$ happens to be unistochastic for simplicity, 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

$$U(t \leftarrow t') \equiv U(t)U^{\dagger}(t'), \tag{66}$$

which yields the composition law

$$U(t) = U(t \leftarrow t')U(t'). \tag{67}$$

At the level of the unistochastic matrix $\Gamma(t)$, one has from the Schur-Hadamard factorization (49) that

$$\left.\begin{aligned} \Gamma(t) &= \overline{U(t)} \odot U(t) \\ &= \overline{[U(t \leftarrow t')U(t')]} \odot [U(t \leftarrow t')U(t')], \end{aligned}\right\} \tag{68}$$

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

$$\left.\begin{aligned} \Gamma_{ij}(t) &= \sum_{k=1}^{N} |U_{ik}(t \leftarrow t')|^2 |U_{kj}(t')|^2 \\ &+ \sum_{k \neq l} \overline{U_{ik}(t \leftarrow t')U_{kj}(t')} U_{il}(t \leftarrow t')U_{lj}(t'). \end{aligned}\right\} \tag{69}$$

With $\Gamma_{kj}(t')$ defined according to (50) as usual,

$$\Gamma_{kj}(t') = |U_{kj}(t')|^2, \tag{70}$$

and defining

$$\Gamma_{ik}(t \leftarrow t') \equiv |U_{ik}(t \leftarrow t')|^2, \tag{71}$$

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

$$\left.\begin{aligned} &\Gamma_{ij}(t) - [\Gamma(t \leftarrow t')\Gamma(t')]_{ij} \\ &= \sum_{k \neq l} \overline{U_{ik}(t \leftarrow t')\Psi_k(t')} U_{il}(t \leftarrow t')\Psi_l(t'), \end{aligned}\right\} \tag{72}$$

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 (34). Remarkably, the right-hand side of (72) has precisely the mathematical form of quantum-theoretic 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 indivisible stochastic dynamics and divisible stochastic dynamics.

In particular, interference does not imply that matter is physically wavelike, contrary to frequent claims in textbook treatments like [33]. Indeed, from the perspective of the present discussion, the notion that interference ever suggested a wavelike quality for matter was merely an unfortunate accident of history, arising from the fact that many early empirical examples of interference in quantum-theoretic systems happened to resemble the behavior of interfering waves propagating in three-dimensional physical space. These historical examples were clearly special cases, as is evident from considering interference in multiparticle systems, whose purported waves would need to propagate through high-dimensional configuration spaces (as was noted by Schrödinger in his early work on wave mechanics [34]), or in more abstract systems, like qubits, that lack continuous configuration spaces altogether.

Nor does interference mean that, say, a particle in a double-slit experiment fails to go through one slit or the other.[12] According to the approach laid out in this paper, the particle does go through a specific slit in each

[11] 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. QED

[12] The exposition in [33] 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.

run of the experiment. 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, and 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.

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 (32) in the system's configuration basis. Indeed, Section XIII will show that non-diagonal self-adjoint matrices will turn out to be candidate observables as well.

These results also suggest that interference should arise in a much broader class of contexts than just for quantum systems. One could imagine experimentally measuring interference effects for essentially any system that can be modeled using indivisible stochastic dynamics.

IX. DIVISION EVENTS

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 environment, delicate correlations from one time to another 'wash out' over short time scales as those correlations leak out into the environment. Using the framework presented in this paper, it is possible to make this intuitive picture more precise.

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

$$\mathcal{C}_{\mathcal{SE}} = \mathcal{C}_{\mathcal{S}} \times \mathcal{C}_{\mathcal{E}}, \tag{73}$$

meaning that each element of $\mathcal{C}_{\mathcal{SE}}$ is a simple ordered pair of the form (i, e).[13] Single 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 overall unistochastic matrix

$$\Gamma^{\mathcal{SE}}(t) = \overline{U^{\mathcal{SE}}(t)} \odot U^{\mathcal{SE}}(t), \tag{74}$$

or, in terms of individual entries,

$$\Gamma^{\mathcal{SE}}_{ie,i_0e_0}(t) = |U^{\mathcal{SE}}_{ie,i_0e_0}(t)|^2. \tag{75}$$

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

$$p^{\mathcal{SE}}_{i'e'}(t') = p^{\mathcal{S}}_{i'}(t')\delta_{e'e(i')}, \tag{76}$$

which describe an idealized 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:

$$\left.\begin{aligned} U^{\mathcal{SE}}(t \leftarrow t') = U^{\mathcal{S}}(t \leftarrow t') \otimes U^{\mathcal{E}}(t \leftarrow t') \\ \text{for } t > t'. \end{aligned}\right\} \tag{77}$$

In terms of individual entries, one has

$$\left.\begin{aligned} U^{\mathcal{SE}}_{ie,i'e'}(t \leftarrow t') = U^{\mathcal{S}}_{ii'}(t \leftarrow t')U^{\mathcal{E}}_{ee'}(t \leftarrow t') \\ \text{for } t > t', \end{aligned}\right\} \tag{78}$$

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.[14]

In light of the Born rule (37), the joint probabilities (76) correspond to a wave function

$$\Psi^{\mathcal{SE}}_{i'e'}(t') = \Psi^{\mathcal{S}}_{i'}(t')\delta_{e'e(i')}, \tag{79}$$

so the composite system's wave function at later times $t > t'$ after the interaction is given in terms of the relative time-evolution operator (78) according to

$$\left.\begin{aligned} \Psi^{\mathcal{SE}}_{ie}(t) &= \sum_{i',e'} U^{\mathcal{SE}}_{ie,i'e'}(t \leftarrow t')\Psi^{\mathcal{SE}}_{i'e'}(t') \\ &= \sum_{i'} U^{\mathcal{S}}_{ii'}(t \leftarrow t')\Psi^{\mathcal{S}}_{i'}(t')U^{\mathcal{E}}_{ee(i')}(t \leftarrow t'). \end{aligned}\right\} \tag{80}$$

[13] Note that the right-hand side of (73) is indeed a Cartesian product, not a tensor product, because this equation is solely a statement about the composite system's kinematics, not its dynamics.

[14] Note the natural appearance of a tensor product in (77) and (78), which are statements about the composite system's dynamics.

From the Born rule (37), one sees that the joint probabilities for $t > t'$ are given by

$$p_{ie}^{\mathcal{SE}}(t) = \left|\Psi_{ie}^{\mathcal{SE}}(t)\right|^2. \tag{81}$$

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 $p_i^{\mathcal{S}}(t)$ for the subject system alone for $t > t'$:

$$\left.\begin{aligned} p_i^{\mathcal{S}}(t) &= \sum_e p_{ie}^{\mathcal{SE}}(t) \\ &= \sum_{i'_1, i'_2} \overline{U_{ii'_1}^{\mathcal{S}}(t \leftarrow t')\Psi_{i'_1}^{\mathcal{S}}(t')} U_{ii'_2}^{\mathcal{S}}(t \leftarrow t')\Psi_{i'_2}^{\mathcal{S}}(t') \\ &\qquad \times \sum_e \overline{U_{ee(i'_1)}^{\mathcal{E}}(t \leftarrow t')} U_{ee(i'_2)}^{\mathcal{E}}(t \leftarrow t') \\ &= \sum_{i'} |U_{ii'}^{\mathcal{S}}(t \leftarrow t')|^2 |\Psi_{i'}^{\mathcal{S}}(t')|^2. \end{aligned}\right\} \tag{82}$$

Taking the limit $t \to t'$ in (82) and referring back to the Born rule (37) again, one sees that the subject system's standalone probabilities at $t' > 0$ are

$$p_{i'}^{\mathcal{S}}(t') = |\Psi_{i'}^{\mathcal{S}}(t')|^2. \tag{83}$$

One also sees from (82) that, as in (71), one can identify

$$\Gamma_{ii'}^{\mathcal{S}}(t \leftarrow t') \equiv |U_{ii'}^{\mathcal{S}}(t \leftarrow t')|^2. \tag{84}$$

Hence, (82) simplifies to a genuinely linear relationship that precisely mirrors the basic marginalization formula (7) for a stochastic process, with $t' > 0$ now serving as the 'initial time':

$$p_i^{\mathcal{S}}(t) = \sum_{i'} \Gamma_{ii'}^{\mathcal{S}}(t \leftarrow t') p_{i'}^{\mathcal{S}}(t'). \tag{85}$$

Applying the basic marginalization formula (7) to the stochastic process from $t = 0$ to $t' > 0$, one also has the equation

$$p_{i'}^{\mathcal{S}}(t') = \sum_j \Gamma_{i'j}^{\mathcal{S}}(t') p_j^{\mathcal{S}}(0). \tag{86}$$

Combining (85) with (86) immediately yields

$$p_i^{\mathcal{S}}(t) = \sum_j \Gamma_{ij}^{\mathcal{S}}(t) p_j^{\mathcal{S}}(0), \tag{87}$$

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

$$\Gamma^{\mathcal{S}}(t) \equiv \Gamma^{\mathcal{S}}(t \leftarrow t')\Gamma^{\mathcal{S}}(t'). \tag{88}$$

Thus, the interaction between the subject system $\mathcal{S}$ and the environment $\mathcal{E}$ up to the time $t' > 0$ has led to a stochastic matrix $\Gamma^{\mathcal{S}}(t)$ for the subject system that is instantaneously divisible at t'.

It is natural to refer to t' as a 'division event.' An important corollary is that $t = 0$ 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 noisy environment.

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 stochastic matrix after any integer number $n \geq 1$ of time steps Δt is given by

$$\Gamma^{\mathcal{S}}(t = n\Delta t) = \left(\Gamma^{\mathcal{S}}\right)^n, \tag{89}$$

where

$$\Gamma_{ij}^{\mathcal{S}} \equiv |U_{ij}^{\mathcal{S}}(\Delta t)|^2. \tag{90}$$

The stochastic dynamics therefore takes the form of a discrete-time Markov chain (13). This analysis therefore provides an explanation for the ubiquity of Markovian stochastic dynamics in so many real-world cases.

X. DECOHERENCE

Had the environment not interacted with the subject system, then the subject system's density matrix $\rho^{\mathcal{S}}(t')$ at $t' > 0$ would have generically been non-diagonal, in accordance with the general definition (30):

$$\left.\begin{aligned} \rho^{\mathcal{S}}(t') &= U^{\mathcal{S}}(t')\left[\sum_j p_j(0) P_j\right] U^{\mathcal{S}\dagger}(t') \\ &= U^{\mathcal{S}}(t')\mathrm{diag}(\dots, p_j(0), \dots) U^{\mathcal{S}\dagger}(t'). \end{aligned}\right\} \tag{91}$$

By contrast, suppose that the environment indeed interacts with the subject system to produce a division event (88) 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 (82), which can be written instead as

$$p_i^{\mathcal{S}}(t) = \mathrm{tr}(P_i \rho^{\mathcal{S}}(t)), \tag{92}$$

where

$$\rho^{\mathcal{S}}(t) \equiv U^{\mathcal{S}}(t \leftarrow t')\rho^{\mathcal{S}}(t') U^{\mathcal{S}\dagger}(t \leftarrow t'), \tag{93}$$

and where, in turn,

$$\rho^{\mathcal{S}}(t') \equiv \sum_{i'} p_{i'}^{\mathcal{S}}(t') P_{i'}^{\mathcal{S}} = \mathrm{diag}\big(\dots, p_{i'}^{\mathcal{S}}(t'), \dots\big), \tag{94}$$

which is diagonal.

On comparing the two expressions (91) and (94) 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 unremarkable 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 (36). 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 in 'multiple states at once.'

These results may also help explain why the precise connection between quantum theory and stochastic processes remained unclear for so long. If one assumes a Markov approximation, as is often the case in the 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.

XI. 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 $t = 0$ up to some time $t' > 0$, but then begin interacting at t'.

For times t between 0 and t', the composite system's stochastic matrix $\Gamma^{\mathcal{AB}}(t)$ factorizes into the tensor product of a stochastic matrix $\Gamma^{\mathcal{A}}(t)$ for $\mathcal{A}$ and a stochastic matrix $\Gamma^{\mathcal{B}}(t)$ for $\mathcal{B}$:

$$\Gamma^{\mathcal{AB}}(t) = \Gamma^{\mathcal{A}}(t) \otimes \Gamma^{\mathcal{B}}(t) \quad \text{for } 0 \leq t < t'. \tag{95}$$

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

$$\Gamma^{\mathcal{AB}}(t) \neq \Gamma^{\mathcal{A}}(t) \otimes \Gamma^{\mathcal{B}}(t) \quad \text{for } t > t', \tag{96}$$

for any possible stochastic matrices $\Gamma^{\mathcal{A}}(t)$ and $\Gamma^{\mathcal{B}}(t)$ that properly capture the respective dynamics of the two subsystems. 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 stochastic 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 (88) at some later time $t'' > t'$, perhaps due to interactions with the larger environment, then the composite system's stochastic matrix $\Gamma^{\mathcal{AB}}(t)$ will divide at t'':

$$\Gamma^{\mathcal{AB}}(t) = \Gamma^{\mathcal{AB}}(t \leftarrow t'')\Gamma^{\mathcal{AB}}(t'') \quad \text{for } t > t'' > t'. \tag{97}$$

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

$$\Gamma^{\mathcal{AB}}(t \leftarrow t'') = \Gamma^{\mathcal{A}}(t \leftarrow t'') \otimes \Gamma^{\mathcal{B}}(t \leftarrow t''), \tag{98}$$

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, but 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.

XII. EMERGEABLES

The preceding sections have shown that a model with kinematics based on a classical configuration space and dynamics based on a suitable stochastic process is capable of accounting for signature features of quantum theory, like 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 like (32) that are diagonal in a single basis. Indeed, the existence of noncommuting observables is another hallmark feature of quantum theory.

Remarkably, a stochastic system will generically contain such observables as well. Specifically, Section XIII will establish that these mathematical objects represent candidate observables that naturally satisfy the usual

[15] Questions about nonlocality will be addressed in detail in Section XVIII.

probabilistic rules of quantum theory, all without introducing any new fundamental axioms. In so doing, the analysis ahead will demonstrate that the dictionary (26) is not merely a tool for studying stochastic processes, but defines a comprehensive stochastic-quantum correspondence.

As motivation, let A be a time-independent (diagonal) random variable (32), and consider the time derivative of its Heisenberg-picture counterpart $A^H(t)$, as defined for a generic time-evolution operator $\Theta(t)$ by (39):

$$\frac{dA^H(t)}{dt} = \frac{\partial\Theta^\dagger(t)}{\partial t}A\Theta(t) + \Theta^\dagger(t)A\frac{\partial\Theta(t)}{\partial t}. \tag{99}$$

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

$$\dot{A} \equiv \lim_{t\to 0}\frac{dA^H(t)}{dt} = \dot{A}^\dagger. \tag{100}$$

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

$$[A, \dot{A}] \neq 0. \tag{101}$$

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

$$\mathrm{tr}(\dot{A}\rho(0)) = \lim_{t\to 0}\frac{d\langle A(t)\rangle}{dt}, \tag{102}$$

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 (100), through the time-evolution operator $\Theta(t)$, and does not have a transparent interpretation at the level of the system's underlying configuration space $\mathcal{C}$. Instead, $\dot{A}$ is an emergent amalgam of kinematical and dynamical ingredients, so it will be called an 'emergeable.' This terminology is intended to contrast $\dot{A}$ with the system's genuine random variables, which could be called 'beables' or 'be-ables', as coined in [35].

XIII. MEASUREMENTS

Consider now 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 $\mathcal{E}$. One of the goals ahead will be to identify the necessary criteria for a subsystem like $\mathcal{D}$ to be regarded as a genuine measuring device.

Focusing momentarily on the subject system, consider a self-adjoint $N \times N$ matrix $\tilde{A}^{\mathcal{S}} = \tilde{A}^{\mathcal{S}\dagger}$, which may or may not be one of the subject system's diagonal random variables.[16] As a concrete example, $\tilde{A}^{\mathcal{S}}$ could be an emergeable like (100).

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

$$\tilde{A}^{\mathcal{S}} = \sum_\alpha \tilde{a}_\alpha \tilde{P}^{\mathcal{S}}_\alpha, \tag{103}$$

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 satisfy the analogues of the mutual exclusivity condition (24),

$$\tilde{P}^{\mathcal{S}}_\alpha \tilde{P}^{\mathcal{S}}_{\alpha'} = \delta_{\alpha\alpha'}\tilde{P}^{\mathcal{S}}_\alpha, \tag{104}$$

and the completeness relation (25),

$$\sum_\alpha \tilde{P}^{\mathcal{S}}_\alpha = \mathbb{1}^{\mathcal{S}}, \tag{105}$$

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 $\tilde{e}_\alpha$ be the corresponding orthonormal basis, one has

$$\tilde{e}^\dagger_\alpha \tilde{e}_{\alpha'} = \delta_{\alpha\alpha'}, \quad \tilde{e}_\alpha \tilde{e}^\dagger_\alpha = \tilde{P}_\alpha. \tag{106}$$

If $\tilde{A}^{\mathcal{S}}$ happens to be one of the subject system's random variables (32), then the eigenvalues $\tilde{a}_\alpha$ are the random variable's magnitudes and the eigenprojectors $\tilde{P}^{\mathcal{S}}_\alpha$ are the configuration projectors (21). More generally, however, the eigenvalues $\tilde{a}_\alpha$ and eigenprojectors $\tilde{P}_\alpha$ 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 index α appearing in the spectral decomposition (103), and, similarly, that the environment $\mathcal{E}$ has configurations $e(\alpha)$ that can be labeled by α. Generalizing (75), suppose, moreover, that the composite system $\mathcal{SDE}$ evolves according to an overall unistochastic matrix

$$\Gamma^{\mathcal{SDE}}_{ide,i_0d_0e_0}(t) = |U^{\mathcal{SDE}}_{ide,i_0d_0e_0}(t)|^2. \tag{107}$$

Generalizing (79) and letting $\tilde{e}_{\alpha',i'}$ denote the i'th component of the basis vector $\tilde{e}_{\alpha'}$ with respect to the configuration basis $e_{i'}$, suppose that the three subsystems interact up to a time $t' > 0$ in such a way that they end up with overall wave function

$$\Psi^{\mathcal{SDE}}_{i'd'e'}(t') = \sum_{\alpha'}\tilde{\Psi}^{\mathcal{S}}_{\alpha'}(t')\tilde{e}_{\alpha',i'}\delta_{d'd(\alpha')}\delta_{e'e(\alpha')}, \tag{108}$$

and that, mirroring (77), the composite system's relative time-evolution operator factorizes between the three

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

subsystems for later times $t > t'$:

$$\left.\begin{aligned} U^{\mathcal{SDE}}(t \leftarrow t') = U^{\mathcal{S}}(t \leftarrow t') \otimes U^{\mathcal{D}}(t \leftarrow t') \\ \otimes U^{\mathcal{E}}(t \leftarrow t') \\ \text{for } t > t'. \end{aligned}\right\} \tag{109}$$

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

$$\left.\begin{aligned} \Psi^{\mathcal{SDE}}_{ide}(t) &= \sum_{i',e',d'} U^{\mathcal{SDE}}_{ide,i'd'e'}(t \leftarrow t')\Psi^{\mathcal{SDE}}_{i'd'e'}(t') \\ &= \sum_{i'}\sum_{\alpha'} U^{\mathcal{S}}_{ii'}(t \leftarrow t')\tilde{\Psi}^{\mathcal{S}}_{\alpha'}(t')\tilde{e}_{\alpha',i'} \\ &\quad \times U^{\mathcal{D}}_{dd(\alpha')}(t \leftarrow t')U^{\mathcal{E}}_{ee(\alpha')}(t \leftarrow t'). \end{aligned}\right\} \tag{110}$$

Invoking the Born rule (37), it follows that the joint probabilities for $t > t'$ are given by

$$p^{\mathcal{SDE}}_{ide}(t) = \left|\Psi^{\mathcal{SDE}}_{ide}(t)\right|^2. \tag{111}$$

Marginalizing over the configuration i of the subject system as well as the configuration e of the environment, 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 $p^{\mathcal{D}}_d(t)$ for the measuring device alone for $t > t'$:

$$\left.\begin{aligned} p^{\mathcal{D}}_d(t) &= \sum_{i,e} p^{\mathcal{SDE}}_{ide}(t) \\ &= \sum_{i'_1,i'_2}\sum_{\alpha'_1,\alpha'_2} \overline{U^{\mathcal{D}}_{dd(\alpha'_1)}(t \leftarrow t')\tilde{\Psi}^{\mathcal{S}}_{\alpha'_1}(t')\tilde{e}_{\alpha'_1,i'_1}} \\ &\quad \times U^{\mathcal{D}}_{dd(\alpha'_2)}(t \leftarrow t')\tilde{\Psi}^{\mathcal{S}}_{\alpha'_2}(t')\tilde{e}_{\alpha'_2,i'_2} \\ &\quad \times \sum_i \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')}U^{\mathcal{S}}_{ii'_2}(t \leftarrow t') \\ &\quad \times \sum_e \overline{U^{\mathcal{E}}_{ee(\alpha'_1)}(t \leftarrow t')}U^{\mathcal{E}}_{ee(\alpha'_2)}(t \leftarrow t') \\ &= \sum_{\alpha'} |U^{\mathcal{D}}_{dd(\alpha')}(t \leftarrow t')|^2 |\tilde{\Psi}^{\mathcal{S}}_{\alpha'}(t')|^2. \end{aligned}\right\} \tag{112}$$

In the limit $t \to t'$, the last line implies that

$$p^{\mathcal{D}}_{d(\alpha')}(t') = |\tilde{\Psi}^{\mathcal{S}}_{\alpha'}(t')|^2, \tag{113}$$

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

$$\langle \tilde{A}^{\mathcal{S}}(t)\rangle \equiv \sum_{\alpha} \tilde{a}_{\alpha} p^{\mathcal{D}}_{d(\alpha')}(t'). \tag{114}$$

This analysis establishes that as long as there exists a form of unistochastic time evolution (107) for the composite system $\mathcal{SDE}$ that arrives at the wave function (108), the matrix $\tilde{A}^{\mathcal{S}}$ represents a genuine observable, in the sense that the time evolution (107) 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 (112) implies that $t' > 0$ is a division event for the measuring device:

$$\Gamma^{\mathcal{D}}(t) = \Gamma^{\mathcal{D}}(t \leftarrow t')\Gamma^{\mathcal{D}}(t') \quad \text{for } t > t'. \tag{115}$$

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

$$\Gamma^{\mathcal{D}}_{dd(\alpha')}(t \leftarrow t') \equiv |U^{\mathcal{D}}_{dd(\alpha')}(t \leftarrow t')|^2. \tag{116}$$

By contrast, unless the observable $\tilde{A}^{\mathcal{S}}$ happens to be one of the subject system's (diagonal) random variables (32), the subject system 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. Nonetheless, for times $t > t'$, one can compute the standalone probability $p^{\mathcal{S}}_i(t)$ for the subject system to be in its ith configuration by marginalizing over the measuring device and the environment:

$$\left.\begin{aligned} p^{\mathcal{S}}_i(t) &= \sum_{d,e} p^{\mathcal{SDE}}_{ide}(t) \\ &= \sum_{i'_1,i'_2}\sum_{\alpha'_1,\alpha'_2} \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')\tilde{\Psi}^{\mathcal{S}}_{\alpha'_1}(t')\tilde{e}_{\alpha'_1,i'_1}} \\ &\quad \times U^{\mathcal{S}}_{ii'_2}(t \leftarrow t')\tilde{\Psi}^{\mathcal{S}}_{\alpha'_2}(t')\tilde{e}_{\alpha'_2,i'_2} \\ &\quad \times \sum_d \overline{U^{\mathcal{D}}_{dd(\alpha'_1)}(t \leftarrow t')}U^{\mathcal{D}}_{dd(\alpha'_2)}(t \leftarrow t') \\ &\quad \times \sum_e \overline{U^{\mathcal{E}}_{ee(\alpha'_1)}(t \leftarrow t')}U^{\mathcal{E}}_{ee(\alpha'_2)}(t \leftarrow t') \\ &= \sum_{\alpha'} \left[\sum_{i'_1,i'_2} \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')}U^{\mathcal{S}}_{ii'_2}(t \leftarrow t')\tilde{e}_{\alpha',i'_2}\overline{\tilde{e}_{\alpha',i'_1}}\right] \\ &\quad \times |\tilde{\Psi}^{\mathcal{S}}_{\alpha'}(t')|^2. \end{aligned}\right\} \tag{117}$$

Recognizing $|\tilde{\Psi}^{\mathcal{S}}_{\alpha'}(t')|^2$ from (113) as the probability $p^{\mathcal{D}}_{d(\alpha')}(t')$ for the measuring device to end up in its configuration $d(\alpha')$ at $t' > 0$, 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 (103) for $\tilde{A}^{\mathcal{S}}$, one can write (117) more succinctly as

$$p^{\mathcal{S}}_i(t) = \mathrm{tr}(P^{\mathcal{S}}_i \rho^{\mathcal{S}}(t)). \tag{118}$$

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

$$\rho^{\mathcal{S}}(t) \equiv U^{\mathcal{S}}(t \leftarrow t')\left[\sum_{\alpha'} p^{\mathcal{D}}_{d(\alpha')}P^{\mathcal{S}}_{\alpha'}\right]U^{\mathcal{S}\dagger}(t \leftarrow t'). \tag{119}$$

Hence, one can recast the expectation value (114) for $\tilde{A}^{\mathcal{S}}$ as

$$\langle \tilde{A}^{\mathcal{S}}(t) \rangle = \mathrm{tr}(\tilde{A}^{\mathcal{S}} \rho^{\mathcal{S}}(t)), \tag{120}$$

which precisely mirrors the formula (31) for the expectation value of a (diagonal) random variable.

Moreover, (117) yields a linear relationship between the standalone probabilities $p^{\mathcal{D}}_{d(\alpha')}(t')$ for the measuring device at $t' > 0$ and the standalone probabilities $p^{\mathcal{S}}_i(t)$ for the subject system at $t > t'$:

$$p^{\mathcal{S}}_i(t) = \sum_{\alpha'} \Gamma^{\mathcal{SD}}_{i,d(\alpha')}(t \leftarrow t') p^{\mathcal{D}}_{d(\alpha')}(t'). \tag{121}$$

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

$$\left.\begin{aligned} &\Gamma^{\mathcal{SD}}_{i,d(\alpha')}(t \leftarrow t') \\ &\equiv \sum_{i'_1, i'_2} \overline{U^{\mathcal{S}}_{ii'_1}(t \leftarrow t')} U^{\mathcal{S}}_{ii'_2}(t \leftarrow t') \tilde{e}_{\alpha', i'_2} \overline{\tilde{e}_{\alpha', i'_1}}. \end{aligned}\right\} \tag{122}$$

Because these matrix entries do not depend on the measuring device's standalone probabilities $p^{\mathcal{D}}_{d(\alpha')}(t')$, they naturally define conditional probabilities for the subject system to be in its ith configuration at the time $t > t'$ given that the measuring device is in its configuration $d(\alpha')$ at $t' > 0$:

$$p^{\mathcal{SD}}(i, t | d(\alpha'), t') \equiv \Gamma^{\mathcal{SD}}_{i,d(\alpha')}(t \leftarrow t'). \tag{123}$$

XIV. WAVE-FUNCTION COLLAPSE

Importantly, notice also that one can write the hybrid stochastic matrix (122) in an overall form that resembles the dictionary (26):

$$\left.\begin{aligned} &\Gamma^{\mathcal{SD}}_{i,d(\alpha')}(t \leftarrow t') \\ &\quad = \mathrm{tr}(U^{\mathcal{S}\dagger}(t \leftarrow t') P^{\mathcal{S}}_i U(t \leftarrow t') \tilde{P}^{\mathcal{S}}_{\alpha}). \end{aligned}\right\} \tag{124}$$

Rearranging the right-hand side gives the equation

$$\Gamma^{\mathcal{SD}}_{i,d(\alpha')}(t \leftarrow t') = \mathrm{tr}(P^{\mathcal{S}}_i \rho^{\mathcal{S}|\alpha',t'}(t)), \tag{125}$$

with a 'conditional' density matrix $\rho^{\mathcal{S}|\alpha',t'}(t)$ for the subject system at the time $t > t'$ naturally defined by time-evolving the eigenprojector $\tilde{P}^{\mathcal{S}}_{\alpha'}$ from $t' > 0$ to $t > t'$:

$$\rho^{\mathcal{S}|\alpha',t'}(t) \equiv U(t \leftarrow t') \tilde{P}^{\mathcal{S}}_{\alpha'} U^{\mathcal{S}\dagger}(t \leftarrow t'). \tag{126}$$

Thus, the calculation (117) reduces to the statement that the standalone probabilities $p^{\mathcal{S}}_i(t)$ for the subject system at $t > t'$ are given by

$$p^{\mathcal{S}}_i(t) = \mathrm{tr}(P^{\mathcal{S}}_i \rho^{\mathcal{S}}(t)), \tag{127}$$

where the subject system's density matrix $\rho^{\mathcal{S}}(t)$, which was originally defined in (119), can equivalently be expressed as a probabilistic mixture of the conditional density matrices $\rho^{\mathcal{S}|\alpha',t'}(t)$ defined in (126), statistically weighted by the measurement probabilities $p^{\mathcal{D}}_{d(\alpha')}(t')$:

$$\rho^{\mathcal{S}}(t) \equiv \sum_{\alpha'} \rho^{\mathcal{S}|\alpha',t'}(t) p^{\mathcal{D}}_{d(\alpha')}(t'). \tag{128}$$

Taking stock of these results, one sees that to make future predictions for $t > t'$ about the subject system, conditioned on the measuring device's result $d(\alpha')$ at $t' > 0$, one effectively replaces the subject system's density matrix with the conditional density matrix $\rho^{\mathcal{S}|\alpha',t'}(t)$, corresponding to a 'collapsed' state vector or wave function defined as

$$\Psi^{\mathcal{S}|\alpha',t'}(t) \equiv U(t \leftarrow t') \tilde{e}_{\alpha}. \tag{129}$$

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 (128), which, again, consists of an appropriate mixture of conditional or collapsed density matrices probabilistically weighted over the measurement results.

XV. THE MEASUREMENT PROBLEM

According to the foregoing analysis, measuring devices are ordinary physical systems that carry out measurements of observables, and then end up in final configurations that reflect definite measurement outcomes, with the probabilities for those various measurement outcomes given by the Born rule. Hence, the picture of quantum theory presented in this paper arguably has the resources to solve the measurement problem.

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 metaphysical 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, 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, 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

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.

XVI. SYMMETRIES

The stochastic-quantum correspondence developed in this paper provides new ways to think about dynamical symmetries in quantum theory. 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 stochastic model.

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

$$\left.\begin{array}{l} P_i \mapsto P_{\sigma(i)}, \\ \quad \text{with } \{\sigma(1), \ldots, \sigma(N)\} = \{1, \ldots, N\}. \end{array}\right\} \tag{130}$$

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

$$P_i \mapsto \tilde{P}_i \equiv V^\dagger P_i V, \tag{131}$$

where V is some unitary operator.[17] This similarity transformation reduces to the configurational transformation (130) if and only if V is a permutation matrix.

The more general transformation (131) leaves the stochastic dynamics invariant precisely if the right-hand side of the stochastic-quantum dictionary (26) remains unchanged:

$$\mathrm{tr}(\Theta^\dagger(t)\tilde{P}_i\Theta(t)\tilde{P}_j) = \mathrm{tr}(\Theta^\dagger(t)P_i\Theta(t)P_j). \tag{132}$$

This condition is equivalent to the statement that

$$\mathrm{tr}(\tilde{\Theta}^\dagger(t)P_i\tilde{\Theta}(t)P_j) = \mathrm{tr}(\Theta^\dagger(t)P_i\Theta(t)P_j), \tag{133}$$

where

$$\tilde{\Theta}(t) \equiv V\Theta(t)V^\dagger. \tag{134}$$

Re-expressing both sides of the equivalent condition (133) in terms of squared absolute values, as in (17), one sees that (134) is a dynamical symmetry precisely if

$$|\tilde{\Theta}_{ij}(t)|^2 = |\Theta_{ij}(t)|^2. \tag{135}$$

It follows immediately that $\tilde{\Theta}(t)$ can differ from $\Theta(t)$ by at most a Schur-Hadamard gauge transformation (42), meaning that a necessary and sufficient condition for a unitary matrix V to give a dynamical symmetry is that

$$V\Theta(t)V^\dagger = \Theta(t) \odot \begin{pmatrix} e^{i\theta_{11}(t)} & e^{i\theta_{12}(t)} & \\ e^{i\theta_{21}(t)} & \ddots & \\ & & e^{i\theta_{NN}(t)} \end{pmatrix}. \tag{136}$$

As special cases, this condition includes unitary dynamical symmetries,

$$V\Theta(t)V^\dagger = \Theta(t), \tag{137}$$

as well as anti-unitary dynamical symmetries,

$$V\Theta(t)V^\dagger = \overline{\Theta(t)}. \tag{138}$$

Note that if one redefines $V \mapsto \overline{V}$, which is still unitary, then one can re-express (138) in the somewhat more conventional form

$$VK\Theta(t)KV^\dagger = \Theta(t). \tag{139}$$

Here K denotes the operation of complex conjugation, so that $K^2 = 1$ and $KXK = \overline{X}$ for any $N \times N$ matrix X. 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.[18]

If $\Theta(t) = U(t)$ is unitary, then $V\Theta(t)V^\dagger$ will likewise be unitary. In that case, if V is continuously connected to the identity $\mathbb{1}$ by some smooth parameter, with a corresponding self-adjoint generator $G = G^\dagger$, then Noether's theorem then easily follows as the statement that the expectation value $\langle G(t)\rangle$ of that generator is constant in time:

$$\langle G(t)\rangle = \mathrm{tr}(GU(t)\rho(0)U^\dagger(t)) = \langle G(0)\rangle. \tag{140}$$

XVII. DILATIONS

In most textbook treatments of quantum theory, a quantum system is axiomatically defined as a particular

[17] Proof: Let $e_1, \ldots, e_N$ be the orthonormal configuration basis (28), with $e_i^\dagger e_j = \delta_{ij}$ and $e_i e_i^\dagger = P_i$, 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 a unitary $N \times N$ 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

[18] Intriguingly, because K anticommutes with i, meaning that $Ki = -iK$, the four mathematical objects 1, i, K, and iK satisfy $-i^2 = K^2 = (iK)^2 = iK(iK) = 1$, and therefore define a Clifford algebra isomorphic to the pseudo-quaternions [36]. In a sense, then, the Hilbert spaces of quantum systems are actually defined not over the complex numbers, but over the pseudo-quaternions.

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.[19] 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 stochastic system. The 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 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):

$$\Gamma_{ij}(t) = \mathrm{tr}(\Theta^\dagger(t) P_i \Theta(t) P_j). \tag{141}$$

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

$$\left.\begin{aligned} \Theta(t) &\mapsto \Theta(t) \otimes \mathbb{1}^{\mathcal{I}}, \\ P_i(t) &\mapsto P_i(t) \otimes \mathbb{1}^{\mathcal{I}}, \\ P_j(t) &\mapsto P_j(t) \otimes P_\gamma^{\mathcal{I}}. \end{aligned}\right\} \tag{142}$$

Here $\mathbb{1}^{\mathcal{I}}$ is the $D \times D$ identity matrix on a new 'internal' Hilbert space $\mathcal{H}_{\mathcal{I}}$, and $P_1^{\mathcal{I}}, \ldots, P_D^{\mathcal{I}}$ collectively form any PVM on that internal Hilbert space satisfying the usual conditions of mutual exclusivity,

$$P_\gamma^{\mathcal{I}} P_{\gamma'}^{\mathcal{I}} = \delta_{\gamma\gamma'} P_\gamma^{\mathcal{I}}, \tag{143}$$

and completeness,

$$\sum_{\gamma=1}^{D} P_\gamma^{\mathcal{I}} = \mathbb{1}^{\mathcal{I}}. \tag{144}$$

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

$$\left.\begin{aligned} \Gamma_{ij}(t) = \mathrm{tr}\big(\mathrm{tr}_{\mathcal{I}}\big(&\big[\Theta^\dagger(t) \otimes \mathbb{1}^{\mathcal{I}}\big]\big[P_i \otimes \mathbb{1}^{\mathcal{I}}\big] \\ &\times \big[\Theta(t) \otimes \mathbb{1}^{\mathcal{I}}\big]\big[P_j \otimes P_\gamma^{\mathcal{I}}\big]\big)\big), \end{aligned}\right\} \tag{145}$$

with a second 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 of the dictionary in block-matrix form as

$$\Gamma_{ij}(t) = \mathrm{tr}_{\mathcal{I}}\Big([\Theta_{ij}(t)]^{\mathcal{I}\dagger}[\Theta_{ij}(t)]^{\mathcal{I}} P_\gamma^{\mathcal{I}}\Big). \tag{146}$$

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,

$$[\Theta_{ij}(t)]^{\mathcal{I}} \equiv \Theta_{ij}(t) \mathbb{1}^{\mathcal{I}}, \tag{147}$$

and the adjoint operation $\dagger$ in (146) 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,

$$[\Theta_{ij}(t)]^{\mathcal{I}\dagger} \equiv \overline{[\Theta_{ij}(t)]}^{\mathcal{I}}. \tag{148}$$

In this dilated version of the Hilbert-space formulation, the Schur-Hadamard gauge transformation (42) is enhanced to the following local-in-time gauge transformation, which has not yet been described in the literature:

$$[\Theta_{ij}(t)]^{\mathcal{I}} \mapsto V_{(ij)}^{\mathcal{I}}(t) [\Theta_{ij}(t)]^{\mathcal{I}}. \tag{149}$$

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:

$$V_{(ij)}^{\mathcal{I}\dagger}(t) = (V_{(ij)}^{\mathcal{I}}(t))^{-1}. \tag{150}$$

The gauge transformations (149) will not generally preserve the factorization $\Theta(t) \otimes \mathbb{1}^{\mathcal{I}}$ appearing in (145), so they motivate considering more general $ND \times ND$ time-evolution operators $\tilde{\Theta}(t)$, in terms of which the dilated dictionary (145) takes the form

$$\Gamma_{ij}(t) = \mathrm{tr}\Big(\mathrm{tr}_{\mathcal{I}}\Big(\tilde{\Theta}^\dagger(t)\big[P_i \otimes \mathbb{1}^{\mathcal{I}}\big]\tilde{\Theta}(t)\big[P_j \otimes P_\gamma^{\mathcal{I}}\big]\Big)\Big). \tag{151}$$

Any $ND \times ND$ matrix $\tilde{\Theta}(t)$ appearing on the right-hand side of this dictionary is guaranteed to lead to a valid stochastic matrix $\Gamma_{ij}(t)$ on the left-hand side, so working with a dilated Hilbert-space formulation essentially provides a larger 'canvas' for designing stochastic 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 [37]. 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 (149) of the Schur-Hadamard gauge transformation now consisting of two-dimensional rotations of the internal Hilbert space $\mathcal{H}_{\mathcal{I}}$. One can then represent the

[19] 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.

complex-conjugation operator K appearing in (139) as the 2×2 matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. The result is that all unitary and anti-unitary operators become $2N \times 2N$ orthogonal matrices. One cost in using this 'real' representation, however, is that the Hilbert spaces of composite systems do not factorize as neatly into Hilbert spaces for their constituent subsystems.[20]

As a somewhat more significant application of dilations, recall that any stochastic matrix $\Gamma_{ij}(t)$ has a Kraus decomposition (48):

$$\Gamma_{ij}(t) = \sum_{\beta=1}^{N} \mathrm{tr}(K_\beta^\dagger(t) P_i K_\beta(t) P_j). \tag{152}$$

The Stinespring dilation theorem [38] 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 $\tilde{U}(t)$:

$$\Gamma_{ij}(t) = \mathrm{tr}\Big(\mathrm{tr}_{\mathcal{I}}\Big(\tilde{U}^\dagger(t)\big[P_i \otimes \mathbb{1}^{\mathcal{I}}\big]\tilde{U}(t)\big[P_j \otimes P_\gamma^{\mathcal{I}}\big]\Big)\Big). \tag{153}$$

This fact makes clear the inevitability of unitary time evolution in quantum theory.[21]

As another application, a dilated Hilbert-space formulation can make it possible to introduce new kinds of emergeables. Some of these emergeables may be observables that can yield definite results in measurement processes, along the lines of Section XIII, despite not having a direct meaning at the level of the system's underlying configuration space. In this way, a stochastically evolving system based on a classical configuration space can easily accommodate emergent observables that model all kinds of quantum-theoretic phenomena. Indeed, obtaining a unitary time-evolution operator for a given system may require dilating the Hilbert space in just this way, as in (153).

An important example of this last application is intrinsic spin. To introduce spin as an 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.

XVIII. NONLOCALITY

This paper has shown that systems with classical configuration spaces and indivisible stochastic dynamics have Hilbert-space representations and can replicate the usual mathematical formalism and empirical predictions of quantum theory.

The classical configurations in this new picture for quantum theory essentially play the role of hidden variables. The term 'hidden variables' immediately raises questions about the potential invocation of nonlocal dynamics, the study of which has motivated famous papers like [39] and has led to the development of a number of important theorems [35, 40–42].

Before assessing the implications of these 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 those 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 Section XI, 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 [40], 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 also assumes that local 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.

There seemed to be just two available options in response to this nonlocality theorem. One could either accept a theory of nonlocal deterministic hidden variables, or deny the existence of deterministic hidden variables and thereby avoid having to introduce any dynamical nonlocality into quantum theory.

---

[20] 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 (42) to make all the entries of the system's time-evolution operator $\Theta(t)$ real-valued. In this alternative approach, however, a unistochastic 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.

[21] From the starting assumptions presented here, one can sketch the following proof: Given $N \times N$ Kraus matrices $K_\beta(t)$, 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 $\tilde{U}_{(i\beta m)(j\gamma)}(t)$ 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 γ runs through N^2 possible values. The last step is to show that $\tilde{U}(t)$ satisfies (153), whose right-hand side reduces to $\sum_{\beta,m} |\tilde{U}_{(i\beta m)(j\gamma)}(t)|^2$. QED

However, for a hidden-variables theory based on stochastic dynamics rather than 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 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 [35, 43, 44]. 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 this statistical locality criterion, start 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 suppose 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 the joint probabilities should factorize according to

$$p(x, y|a, b, \lambda) = p(x, a|\lambda)p(y, b|\lambda). \tag{154}$$

Bell's statistical locality criterion is precisely the condition that the theory in question should have enough hidden variables to ensure that the factorization (154) 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 probability of 1 [42], one can again derive predictions that are violated by quantum theory, just as in the case of a deterministic hidden-variables theory.

However, this statistical locality criterion is broader than the conditions Bell studied in his earlier work on deterministic hidden-variables theories in [40]. Bell's statistical locality condition is so broad, in fact, that Bell used it to argue that textbook quantum theory is itself dynamically nonlocal [35, 45].

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 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 Section XIII, 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.

This seemingly arbitrary division of the world into measuring devices, which truly have underlying configurations, and all other systems, which do not, leads directly to all the usual mysteries about the measurement process according to textbook quantum theory. 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 (154), 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 (154) either, but they also do not lead to any trouble for the no-communication theorem [46, 47], 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, 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. If one instead disputes Bell's statistical locality 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.

A number of other important no-go theorems have been proved over the years, like von Neumann's early no-go theorem [9], the Kochen-Specker theorem [48], the Pusey-Barrett-Rudolph theorem [49], and Myrvold's no-go theorem [50]. 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 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.

XIX. DISCUSSION

This paper has shown that one can replicate the mathematical formalism and predictions of quantum theory using a physical model based on classical kinematics and stochastic dynamics, and therefore naturally suggests a new interpretation of quantum theory grounded in the theory of stochastic processes. According to this highly adaptable 'stochastic-process interpretation,' systems have underlying physical configurations in configuration spaces at all times, and their dynamics is no more or less nonlocal than the dynamics of textbook quantum theory.

From this perspective, density matrices, wave functions, and other appurtenances of Hilbert spaces, while highly useful, are merely gauge variables and should not be assigned direct physical meanings or treated as though they directly represent physical objects, any more than a Lagrangian or a Hamilton's principal function directly describe 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 non-commuting 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.

This approach to quantum theory therefore shares some features with the de Broglie-Bohm formulation, or Bohmian mechanics [51–53]. However, in contrast to the stochastic-process interpretation, Bohmian mechanics employs deterministic dynamics, and features a guiding equation that fundamentally breaks Lorentz invariance by singling out a preferred foliation of spacetime into spacelike hypersurfaces. The stochastic-process interpretation instead takes seriously what experiments strongly suggest—that the dynamics of quantum theory is indeterministic.

In contrast with the Everett interpretation [54, 55], the stochastic-process interpretation 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 numbers of parallel worlds. The approach in this paper is therefore more modest, metaphysically speaking, than the Everett interpretation. The Everett interpretation arguably exhibits a manifest notion of dynamical locality at a level of description that transcends its individual world-branches [56], but because each individual world-branch looks no more or less nonlocal than the world according to textbook quantum theory or the stochastic-process interpretation, it is not clear what concrete benefits the Everett interpretation's dynamical locality truly provide.

Unlike stochastic-collapse theories [57, 58], the stochastic-process interpretation does not entail any fundamental violations of unitarity, nor does it require introducing any new constants of nature to specify dynamical-collapse rates.

The stochastic-process interpretation shares some features with the modal interpretations [59–61], including an insistence that systems always have definite configurations of some kind at every moment in time, while treating at least some forms of probability in a law-like, objective way. In particular, like the minimal modal interpretation [62], the stochastic-process interpretation uses conditional probabilities in a central way. One difference between the stochastic-process interpretation and most of the modal interpretations, however, is the insistence by the stochastic-process interpretation that the definite configuration of a given system is an element of a classical-looking configuration space, rather than corresponding more abstractly to an element of a Hilbert space. The stochastic-process interpretation also avoids some of the ontological instabilities that are a serious challenge for most of the modal interpretations [63].

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

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 David Baker, Louis Deslauriers, Melissa Franklin, Mina Himwich, David Kaiser, Efthimios Kaxiras, Gregory Kestin, Curtis McMullen, and Matthew Reece for helpful discussions.

[1] E. C. G. Sudarshan, P. M. Mathews, and J. Rau, Physical Review **121**, 920 (1961).
[2] G. A. Skorobogatov and S. I. Svertilov, Physical Review A **58**, 3426 (1998).
[3] D. T. Gillespie, in *AIP Conference Proceedings*, Vol. 511 (American Institute of Physics, 2000) pp. 49–56.
[4] D. T. Gillespie, W. O. Alltop, and J. M. Martin, Chaos: An Interdisciplinary Journal of Nonlinear Science **11**, 548 (2001).

[5] S. Aaronson, Physical Review A **71**, 032325 (2005).
[6] C. Wetterich, ""Quantum Representation of Classical Statistical Ensembles"," (2009), unpublished draft.
[7] M. Frasca, arXiv (2012), arXiv:1201.5091 [math-ph].
[8] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 1st ed. (Oxford University Press, 1930).
[9] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer, 1932).
[10] S. Ross, *Stochastic Processes*, 2nd ed. (John Wiley and Sons, 1995).
[11] S. Milz and K. Modi, PRX Quantum **2**, 030201 (2021), arXiv:2012.01894v2 [quant-ph].
[12] I. Schur, Journal für die reine und angewandte Mathematik **140**, 1 (1911).
[13] P. R. Halmos, *Finite-Dimensional Vector Spaces*, 2nd ed. (Van Nostrand, 1958).
[14] R. A. Horn, in *Matrix Theory and Applications*, Vol. 40 (1990) pp. 87–169.
[15] G. W. Mackey, Annals of Mathematics **55**, 101 (1952).
[16] G. W. Mackey, The American Mathematical Monthly **64**, 45 (1957).
[17] A. Auffeves and P. Grangier, Scientific Reports **7**, 43365 (2017), arXiv:1610.06164v2 [quant-ph].
[18] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*, 3rd ed. (Cambridge University Press, 2018).
[19] J. S. Townsend, *A Modern Approach to Quantum Mechanics*, 2nd ed. (Cambridge University Press, 2012).
[20] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. (Plenum Press, 1994).
[21] J. J. Sakurai and J. J. Napolitano, *Modern Quantum Mechanics*, 2nd ed. (Addison-Wesley, 2010).
[22] B. Schumacher and M. Westmoreland, *Quantum Processes, Systems, and Information*, 1st ed. (Cambridge University Press, 2010).
[23] D. J. Griffiths, *Introduction to Electrodynamics*, 4th ed. (Cambridge University Press, 2017).
[24] A. Zangwill, *Modern Electrodynamics* (Cambridge University Press, 2012).
[25] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, 1998).
[26] H. Brown, in *From Physics to Philosophy* (Cambridge University Press, 1999) pp. 45–70.
[27] K. Kraus, Annals of Physics **64**, 311 (1971).
[28] J. A. Barandes and D. Kagan, Annals of Physics **448**, 169192 (2023), arXiv:2109.07447.
[29] R. C. Thompson, "Unpublished Lecture Notes," (1989), unpublished lecture notes.
[30] P. Nylen, T.-Y. Tam, and F. Uhlig, Linear and Multilinear Algebra **36**, 69 (1993).
[31] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Westview Press, 1995).
[32] S. Weinberg, *The Quantum Theory of Fields*, Vol. 2 (Cambridge University Press, 1996).
[33] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics, Volume 3* (Addison-Wesley, 1965).
[34] E. Schrödinger, Physical Review **28**, 1049 (1926).
[35] J. S. Bell, CERN (1975).
[36] E. C. G. Stueckelberg, Helvetica Physica Acta **33**, 727 (1960).
[37] J. Myrheim, (1999), arXiv:quant-ph/9905037 [quant-ph].
[38] W. F. Stinespring, Proceedings of the American Mathematical Society **6**, 211 (1955).
[39] A. Einstein, B. Podolsky, and N. Rosen, Physical Review **47**, 777 (1935).
[40] J. S. Bell, Physics **1**, 195 (1964).
[41] J. F. Clauser, M. A. Horne, A. E. Shimony, and R. A. Holt, Physical Review Letters **23**, 880 (1969).
[42] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, Fundamental Theories of Physics (Springer, 1989) pp. 69–72, arXiv:0712.0921 [quant-ph].
[43] J. Butterfield, The British Journal for the Philosophy of Science **43**, 41 (1992).
[44] W. Myrvold, M. Genovese, and A. Shimony, in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta (Metaphysics Research Lab, Stanford University, 2021) Fall 2021 ed.
[45] J. S. Bell, in *Between Science and Technology*, edited by A. Sarlemijn and P. Kroes (Elsevier, 1990) pp. 97–115.
[46] G. C. Ghirardi, A. Rimini, and T. Weber, Lettere al Nuovo Cimento **27**, 293 (1980).
[47] T. F. Jordan, Physics Letters A **94**, 264 (1983).
[48] S. B. Kochen and E. P. Specker, Indiana University Mathematics Journal **17**, 59 (1967).
[49] M. F. Pusey, J. Barrett, and T. Rudolph, Nature Physics **8**, 475 (2012), arXiv:1111.3328 [quant-ph].
[50] W. C. Myrvold, Foundations of Physics **32**, 1773 (2002), arXiv:quant-ph/0209109 [quant-ph].
[51] L. de Broglie, *An Introduction to the Study of Wave Mechanics* (E. P. Dutton and Company, Inc., 1930).
[52] D. J. Bohm, Physical Review **85**, 166 (1952).
[53] D. J. Bohm, Physical Review **85**, 180 (1952).
[54] H. Everett III, Reviews of Modern Physics **29**, 454 (1957).
[55] H. Everett III, in *The Many-Worlds Interpretation of Quantum Mechanics*, Vol. 1 (1973) p. 3.
[56] D. Wallace, *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation* (Oxford University Press, 2012).
[57] G. C. Ghirardi, A. Rimini, and T. Weber, Physical Review D **34**, 470 (1986).
[58] A. Bassi and G. C. Ghirardi, Physical Reports **379**, 257 (2003), arXiv:quant-ph/0302164 [quant-ph].
[59] H. P. Krips, Philosophy of Science **36**, 145 (1969).
[60] B. C. Van Fraassen, in *"Paradigms and Paradoxes: The Philosophical Challenge of the Quantum Domain"*, edited by R. Colodny (University of Pittsburgh Press, 1972) pp. 303–366.
[61] P. E. Vermaas and D. G. B. J. Dieks, Foundations of Physics **25**, 145 (1995).
[62] J. A. Barandes and D. Kagan, Foundations of Physics **50**, 1189 (2020), arXiv:1807.07136 [quant-ph].
[63] P. E. Vermaas, *A Philosopher's Understanding of Quantum Mechanics: Possibilities and Impossibilities of a Modal Interpretation* (Cambridge University Press, 1999).