The Stochastic-Quantum Theorem

Jacob A. Barandes[1,*]

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

(Dated: September 3, 2023)

This paper introduces several new classes of mathematical structures that have close connections with physics and with the theory of dynamical systems. The most general of these structures, called generalized stochastic systems, collectively encompass many important kinds of stochastic processes, including Markov chains and random dynamical systems. This paper then states and proves a new theorem that establishes a precise correspondence between any generalized stochastic system and a unitarily evolving quantum system. This theorem therefore leads to a new formulation of quantum theory, alongside the Hilbert-space, path-integral, and quasiprobability formulations. The theorem also provides a first-principles explanation for why quantum systems are based on the complex numbers, Hilbert spaces, linear-unitary time evolution, and the Born rule. In addition, the theorem suggests that by selecting a suitable Hilbert space, together with an appropriate choice of unitary evolution, one can simulate any generalized stochastic system on a quantum computer, thereby potentially opening up an extensive set of novel applications for quantum computing.

I. INTRODUCTION

In the development of physical theories, it sometimes turns out that existing definitions are too conceptually limiting, and that more flexible definitions are needed. Working with more flexible definitions at a higher level of abstraction or generality may make it easier to discover new connections or prove new theorems, which would then also apply down at the lower level of the original definitions.

This paper will argue that by appropriately generalizing standard definitions of dynamical systems, one can obtain novel classes of mathematical structures that encompass an extensive array of physically important models. As with a traditionally defined dynamical system, each such mathematical structure describes a physical system moving deterministically or stochastically along some trajectory in a configuration space, albeit with a more general set of laws than according to standard definitions. (For pedagogical treatments of the standard theory of dynamical systems, see, for instance, [1–3].)

This paper also states and proves a new theorem showing that despite being based on trajectories in configuration spaces, the newly introduced class of *generalized stochastic systems* actually includes all quantum systems. As a consequence, this *stochastic-quantum theorem* offers a more conceptually transparent way to understand quantum systems, with superpositions no longer regarded as literal blends of physical states. The theorem also provides a first-principles explanation

* jacob_barandes@harvard.edu

for features of quantum theory that are usually taken to be axiomatic, including Hilbert spaces over the complex numbers, linear-unitary evolution, and the Born rule.

Seen from another point of view, this *stochastic-quantum correspondence* yields an alternative way to formulate quantum theory, in the language of trajectories unfolding stochastically in configuration spaces. This alternative formulation is distinct from the traditional Hilbert-space formulation [4, 5], the path-integral formulation [6–8], and the quasi-probability formulation [9, 10].

From a more practical perspective, turning this stochastic-quantum correspondence around suggests that unitarily evolving quantum systems can be put to work simulating a very broad class of non-Markovian stochastic processes, thereby potentially opening up an extensive suite of new applications for quantum computers.

Section II begins by defining deterministic generalizations of dynamical systems, followed by the introduction of important distinctions between *indivisible*, *divisible*, and *Markovian* dynamics. Section III provides a generalized definition of a system with stochastic laws, shows how to represent such a generalized stochastic system in the formalism of linear algebra, describes connections between this work and the existing research literature, defines the relationship between a composite system and its subsystems, and introduces the crucial notion of a *unistochastic system*. Section IV states the stochastic-quantum theorem, whose proof is this paper's primary goal, and then discusses some important corollaries and provides a simple example of the theorem in practice. Section V lays out the theorem's proof, which entails explicitly constructing the claimed correspondence between stochastic systems and quantum systems along the way. Section VI concludes the paper with a brief discussion of future

II. DETERMINISTIC SYSTEMS

A. Generalized Dynamical Systems

Dynamical systems are abstract mathematical structures that usefully model many deterministic physical processes. According to the standard definition [1–3], a dynamical system consists of a map representing some kind of evolution law that can be repeatedly applied to the elements of some set of states. A dynamical system is usually assumed to be *divisible*, in the sense that one can 'divide up' its evolution law over any time duration into well-defined evolution laws that describe intermediate time durations. The more general case would be an *indivisible* dynamical system that might lack this feature.

The terms 'divisible' and 'indivisible' for dynamical laws are remarkably new. The terminology appears to be due to Wolf and Cirac, who introduced it in a 2008 paper [11] on quantum channels.[1]

To accommodate the eventual possibility of indivisible evolution, this paper will define a *generalized dynamical system* to mean a tuple of the form

$$(\mathcal{X}, \mathcal{T}, f) \tag{1}$$

[1] Note that this terminology is unrelated to the much older concept of *infinite divisibility*, which refers to a probability distribution that can be expressed as the probability distribution of a sum of any integer number of independent and identically distributed random variables.

that consists of the following data.

- The symbol $\mathcal{X}$ denotes a set that will be called the generalized dynamical system's *state space* (or *phase space*), and whose individual elements $i \in \mathcal{X}$ denote the system's *(allowed) states*.

- Note that $\mathcal{X}$ may or may not be a finite set, and it may or may not involve additional structure in its definition, such as measure-theoretic structure or a vector-space structure. For the purposes of this paper, no such additional structure will be specified or assumed. More broadly, for reasons of brevity and simplicity, this paper will entirely set aside measure-theoretic considerations that arise for the case of uncountable sets.

- The symbol $\mathcal{T}$ denotes a *set of times* $t \in \mathcal{T}$, where $\mathcal{T}$ may or may not be isomorphic to a subset of the real line $\mathbb{R}$ under addition.

- The symbol f denotes a map

$$f : \mathcal{X} \times \mathcal{T} \to \mathcal{X} \tag{2}$$

  that will be called the system's *dynamical map*. This dynamical map f takes as inputs any state i and any time t, and outputs a state $f(i,t) \in \mathcal{X}$:

$$i, t \mapsto f(i,t) \in \mathcal{X} \quad [\text{for all } i \in \mathcal{X},\, t \in \mathcal{T}]. \tag{3}$$

- Fixing the time t turns f into a *time-dependent dynamical map*

$$f_t : \mathcal{X} \to \mathcal{X} \tag{4}$$

  defined by

$$i \mapsto f_t(i) \equiv f(i,t) \quad [\text{for all } i \in \mathcal{X}]. \tag{5}$$

- Without any important loss of generality, the set of times $\mathcal{T}$ will be assumed to include an element denoted by 0 and called the *initial time*. It will be further assumed that at the initial time 0, the time-dependent dynamical map f_t trivializes to the identity map $\mathrm{id}_{\mathcal{X}}$ on $\mathcal{X}$:

$$f_0 = \mathrm{id}_{\mathcal{X}}, \quad \text{or} \quad f_0(i) = i \quad [\text{for all } i \in \mathcal{X}]. \tag{6}$$

- One can regard the argument i appearing in the expression $f_t(i)$ as an initial state of the system at the initial time 0, with the time-dependent dynamical map f_t then describing the evolution of that state i from the initial time 0 to the time t.

- Given a fixed state i, the set of states

$$\{f_t(i) \,|\, t \in \mathcal{T}\} \subset \mathcal{X} \tag{7}$$

describes the *orbit*, or *trajectory*, of the initial state i through the system's state space $\mathcal{X}$ according to the dynamical map f.

In many applications, one takes the set of times $\mathcal{T}$ to be a *semigroup*, meaning that the definition of $\mathcal{T}$ includes an associative binary operation $\star$ (which is often denoted instead by $+$ in the commutative case):

$$t, t' \mapsto t \star t' \in \mathcal{T} \quad \big[\text{for all } t, t' \in \mathcal{T}\big], \tag{8}$$

$$(t \star t') \star t'' = t \star (t' \star t'') \quad \big[\text{for all } t, t', t'' \in \mathcal{T}\big]. \tag{9}$$

One usually also takes the initial time 0 to be the identity element under this binary operation,

$$0 \star t = t \star 0 = t \quad [\text{for all } t \in \mathcal{T}], \tag{10}$$

in which case $\mathcal{T}$ becomes a *monoid*, meaning a semigroup with an identity element. If, furthermore, every time t has an inverse t' such that $t \star t' = t' \star t = 0$, then $\mathcal{T}$ becomes a *group*.

B. Divisible Dynamical Systems

In the most general case, a generalized dynamical system will not provide a way to evolve a system from a *non-initial* time $t' \neq 0$ to another time t. A generalized dynamical system will also generically lack any means of 'dividing up' the evolution from 0 to $t \neq 0$ into well-defined forms of evolution over intermediate time durations between 0 and t (even assuming that the set of times $\mathcal{T}$ has a notion of ordering). To make contact with the kinds of dynamical systems considered more widely in the research literature, it will therefore be necessary to introduce a somewhat less general class of mathematical structures.

This paper will define a *divisible dynamical system* to be a tuple of the form

$$(\mathcal{X}, \mathcal{T}, g). \tag{11}$$

- Here $\mathcal{X}$ is a state space and $\mathcal{T}$ is a set of times forming a monoid, whose identity element 0 plays the role of an initial time, as usual.

- The symbol g denotes a map

$$g : \mathcal{X} \times \mathcal{T}^2 \to \mathcal{X} \tag{12}$$

that will be called the divisible dynamical system's *transition map*. This transition map g

takes as inputs any state i and any pair of times t, t', and outputs a state $g(i,(t,t')) \in \mathcal{X}$:

$$i, t, t' \mapsto g\big(i,\big(t,t'\big)\big) \in \mathcal{X} \qquad \big[\text{for all } i \in \mathcal{X},\, t, t' \in \mathcal{T}\big]. \tag{13}$$

The transition map g here should be understood as describing the evolution or transition of the state i at the time t' to the state $g(i,(t,t'))$ at the time t.

- Fixing two times t, t' turns g into a *time-dependent transition map*

$$g_{t\leftarrow t'} : \mathcal{X} \to \mathcal{X} \tag{14}$$

defined by

$$i \mapsto g_{t\leftarrow t'}(i) \equiv g\big(i,\big(t,t'\big)\big) \qquad [\text{for all } i \in \mathcal{X}]. \tag{15}$$

This time-dependent transition map will be required to trivialize to the identity map $\mathrm{id}_{\mathcal{X}}$ on $\mathcal{X}$ when $t' = t$,

$$g_{t\leftarrow t} = \mathrm{id}_{\mathcal{X}} \quad [\text{for all } t \in \mathcal{T}], \qquad \text{or} \quad g_{t\leftarrow t}(i) = i \quad [\text{for all } i \in \mathcal{X},\, t \in \mathcal{T}], \tag{16}$$

as well as satisfy the *divisibility condition*

$$g_{t\leftarrow t''} = g_{t\leftarrow t'} \circ g_{t'\leftarrow t''} \qquad \big[\text{for all } t, t', t'' \in \mathcal{T}\big], \tag{17}$$

where $\circ$ denotes function composition. The divisibility condition means that the time-dependent transition map $g_{t\leftarrow t''}$ factorizes or 'divides up' into a part $g_{t'\leftarrow t''}$ that carries out the evolution from t'' to t', followed by a part $g_{t\leftarrow t'}$ that carries out the evolution from t' to t.

- Choosing the time t' in the time-dependent transition map $g_{t\leftarrow t'}$ to be the initial time 0 naturally defines a time-dependent dynamical map (4),

$$f_t \equiv g_{t\leftarrow 0} \quad [\text{for all } t \in \mathcal{T}], \tag{18}$$

which then also defines an overall dynamical map $f : \mathcal{X} \times \mathcal{T} \to \mathcal{X}$ according to $f(i,t) \equiv f_t(i)$. The trivialization condition $g_{0\leftarrow 0} = \mathrm{id}_{\mathcal{X}}$ from (16) ensures that f_t satisfies the corresponding trivialization condition $f_0 = \mathrm{id}_{\mathcal{X}}$ in (6). It follows that every divisible dynamical system is, in particular, a generalized dynamical system, one that includes the additional structure that corresponds to having a transition map g.

- Meanwhile, setting $t'' = 0$ in the divisibility condition (17) on the transition map g yields the subsidiary divisibility condition

$$f_t = g_{t \leftarrow t'} \circ f_{t'} \quad \big[\text{for all } t, t' \in \mathcal{T}\big], \tag{19}$$

  which means that the time-dependent dynamical map f_t 'divides up' into a part $f_{t'}$ that carries out the evolution from the initial time 0 to t', followed by a part $g_{t \leftarrow t'}$ that carries out the evolution from t' to t.

- It also follows from the subsidiary divisibility condition (19), together with the trivialization condition $f_0 = \mathrm{id}_{\mathcal{X}}$ in (6), that the time-dependent dynamical map f_t has a well-defined inverse f_t^{-1} given by

$$f_t^{-1} = g_{0 \leftarrow t} \quad [\text{for all } t \in \mathcal{T}]. \tag{20}$$

  That is, a divisible dynamical system's time-dependent dynamical maps f_t are invertible.[2]

C. Markovian Dynamical Systems

A divisible dynamical system will be called *memoryless*, or *Markovian*, if it has the special property

$$g_{t \star t' \leftarrow t'} = g_{t \leftarrow 0} \quad \big[\text{for all } t, t' \in \mathcal{T}\big]. \tag{21}$$

Equivalently, using $f_t \equiv g_{t \leftarrow 0}$ from (18), one can write this special property as

$$g_{t \star t' \leftarrow t'} = f_t \quad \big[\text{for all } t, t' \in \mathcal{T}\big]. \tag{22}$$

It then follows immediately from the subsidiary divisibility condition (19) that the dynamical map $f : \mathcal{X} \times \mathcal{T} \to \mathcal{X}$ gives a semigroup action of the set of times $\mathcal{T}$ on the state space $\mathcal{X}$, in the sense that the semigroup operation $\star$ on $\mathcal{T}$ is mapped to function composition:

$$f_{t \star t'} = f_t \circ f_{t'} \quad \big[\text{for all } t, t' \in \mathcal{T}\big]. \tag{23}$$

This equation is an example of *Markovianity*, or a *Markov property*. Roughly speaking, a Markov property implies that time evolution depends only on duration, and not on history.

Notice that the Markov property (23) is phrased directly in terms of a dynamical map f, so it can be imposed on a generalized dynamical system $(\mathcal{X}, \mathcal{T}, f)$, without any need to invoke a transition map g. In that case, it also follows that there is no longer a meaningful distinction between a generalized dynamical system $(\mathcal{X}, \mathcal{T}, f)$ and a divisible dynamical system $(\mathcal{X}, \mathcal{T}, g)$. As

[2] If the set of times $\mathcal{T}$ has a well-defined ordering relation, and if one modifies the definition of a divisible dynamical system $(\mathcal{X}, \mathcal{T}, g)$ by restricting the transition map (12) so that $g_{t \leftarrow t'}$ is only defined when t' comes before t according to that ordering relation, then the arguments leading to (20) will break down. In that case, the time-dependent dynamical maps f_t will not necessarily have inverses.

such, if the Markov property (23) is imposed, then the resulting mathematical structure will simply be called a *Markovian dynamical system*.

What most references call a 'dynamical system' corresponds to what this paper would call a Markovian dynamical system. That is, a dynamical system, without any further qualifiers, is a tuple $(\mathcal{X}, \mathcal{T}, f)$ for which $\mathcal{X}$ is a state space, $\mathcal{T}$ is a set of times forming a monoid, and $f : \mathcal{X} \times \mathcal{T} \to \mathcal{X}$ is a dynamical map satisfying the Markov property (23). It follows that a generalized dynamical system (1), as befits its name, really is a generalization of the kinds of dynamical systems that are usually considered in textbooks, in the research literature, and in many applications.

In the most general cases, however, generalized dynamical systems will not satisfy the Markov property (23). Nor will it necessarily be possible to obtain a given generalized dynamical system $(\mathcal{X}, \mathcal{T}, f)$ by starting with a divisible dynamical system $(\mathcal{X}, \mathcal{T}, g)$ and then defining a dynamical map f according to $f_t \equiv g_{t \leftarrow 0}$ for all times t, as in (18). Indeed, in light of the inversion formula (20), the absence of any inverse time-dependent dynamical maps f_t^{-1} would present an immediate obstruction to deriving a dynamical map f from a transition map g.

A generalized dynamical system $(\mathcal{X}, \mathcal{T}, f)$ that cannot be derived from a divisible dynamical system $(\mathcal{X}, \mathcal{T}, g)$ will lack the structure needed to divide up its time evolution into intermediate durations, and will therefore be called an *indivisible dynamical system*.

III. STOCHASTIC SYSTEMS

A. Generalized Stochastic Systems

This paper will be primarily concerned with a mathematical structure that replaces the deterministic behavior of a generalized dynamical system (1) with probabilistic, or stochastic, behavior. A *generalized stochastic system* will be defined to mean a tuple of the form

$$(\mathcal{C}, \mathcal{T}, \Gamma, p, \mathcal{A}) \tag{24}$$

that consists of the following data.

- The symbol $\mathcal{C}$ will denote a set called the system's *configuration space*, and the elements of $\mathcal{C}$ will be called the system's *(allowed) configurations*. Configurations and configuration spaces will play an analogous role for generalized stochastic systems that states and state spaces play for generalized dynamical systems.

- The reason for switching the terminology from 'states' to 'configurations' is conceptual. In applications of the theory of dynamical systems to physical situations, such as in classical Hamiltonian mechanics, the notion of a 'state' is often taken to include rates of change or momenta in addition to configurations, because defining a 'state' in that way can make it possible to obtain deterministic laws in the form of first-order differential equations. For a generalized stochastic system, by contrast, the system's probabilistic laws may mean that rates of change and momenta are not well-defined in the absence of specifying a trajectory.

In that case, the only available notion of a 'state' is limited to the more rudimentary notion of a 'configuration.'

- For the purposes of this paper, $\mathcal{C}$ will be assumed to be a finite set, with some (possibly very large) number N of elements denoted by $1, 2, 3, \ldots, N$, and labeled abstractly by Latin letters $i, j, \ldots$. By an abuse of notation, the symbol $\mathcal{C}$ will sometimes be used to refer to the generalized stochastic system as a whole.

- The symbol $\mathcal{T}$ will again denote a *set of times*, including a time 0 that will be called the system's *initial time*.

- The symbol Γ will denote a map

$$\Gamma : \mathcal{C}^2 \times \mathcal{T} \to [0, 1] \subset \mathbb{R} \tag{25}$$

that will be called the system's *stochastic map*, where $[0, 1]$ denotes the closed unit interval $0 \leq x \leq 1$. Each value of this stochastic map will be labeled as

$$\begin{aligned} \Gamma_{ij}(t) &\equiv \Gamma((i, j), t) \\ &\quad [\text{for all } i, j \in \mathcal{C},\, t \in \mathcal{T}], \end{aligned} \tag{26}$$

and will be called the *conditional probability* or *transition probability* $p(i, t|j, 0)$ for the system to be in its ith configuration at the time t, given that the system is in its jth configuration at the initial time 0:

$$\begin{aligned} \Gamma_{ij}(t) &\equiv p(i, t|j, 0) \in [0, 1] \\ &\quad [\text{for all } i, j \in \mathcal{C},\, t \in \mathcal{T}]. \end{aligned} \tag{27}$$

- The stochastic map Γ will be required to satisfy the standard normalization condition

$$\begin{aligned} \sum_{i=1}^{N} \Gamma_{ij}(t) = \sum_{i=1}^{N} p(i, t|j, 0) = 1 \\ [\text{for all } j \in \mathcal{C},\, t \in \mathcal{T}], \end{aligned} \tag{28}$$

as well as the initial condition

$$\Gamma_{ij}(0) \equiv \delta_{ij} \equiv \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j, \end{cases} \tag{29}$$

where δ_{ij} is the usual Kronecker delta.

- The symbol p will denote a map

$$p : \mathcal{C} \times \mathcal{T} \to [0, 1] \subset \mathbb{R} \tag{30}$$

that will be called the system's *probability distribution*. Each value of this map will be labeled as

$$p_i(t) \equiv p(i,t) \in [0,1] \quad [\text{for all } i \in \mathcal{C},\, t \in \mathcal{T}], \tag{31}$$

and will be called the *standalone probability* for the system to be in its ith configuration at the time t.

- Only the standalone probabilities $p_1(0), \ldots, p_N(0)$ at the initial time 0 will be taken to be freely adjustable, subject to the standard normalization condition

$$\sum_{j=1}^{N} p_j(0) = 1. \tag{32}$$

The standalone probabilities at every other time t will be assumed to be defined by the following *Bayesian marginalization condition*:

$$\begin{aligned} p_i(t) &\equiv \sum_{j=1}^{N} \Gamma_{ij}(t) p_j(0) \\ &= \sum_{j=1}^{N} p(i,t|j,0) p_j(0) \\ &[\text{for all } i \in \mathcal{C},\, t \in \mathcal{T}]. \end{aligned} \tag{33}$$

The normalization condition (28) on the stochastic map Γ and the normalization condition (32) on the standalone probabilities at the initial time 0 then together ensure that the probability distribution p satisfies the standard normalization condition more generally:

$$\sum_{i=1}^{N} p_i(t) = 1 \quad [\text{for all } t \in \mathcal{T}]. \tag{34}$$

- Note that the definition of the stochastic map Γ is independent of the choice of standalone probabilities $p_1(0), \ldots, p_N(0)$ at the initial time 0. That is, Γ can be freely adjusted independently of those initial standalone probabilities. Importantly, the Bayesian marginalization condition (33) therefore defines a *linear relationship* between the standalone probabilities $p_1(0), \ldots, p_N(0)$ at the initial time 0 and the standalone probabilities $p_1(t), \ldots, p_N(t)$ at any other time t. In the work ahead, it will be argued that this linear relationship is ultimately responsible for the linearity of time evolution in quantum theory.

- The stochastic map Γ will *not* be assumed to have anything like a Markov or divisibility property. More precisely, no assumption will be made that given any pair of times t, t', there

will exist a time t'' such that the following Markov property holds:

$$\Gamma_{ij}(t) = \sum_{k=1}^{N} \Gamma_{ik}\big(t''\big)\Gamma_{kj}\big(t'\big) \qquad [\text{for all } i, j \in \mathcal{C}]. \tag{35}$$

Nor will it be assumed that Γ satisfies even the *weaker* divisibility condition that for any pair of times t, t', there exists a set of real-valued, non-negative quantities $X_{ik} \geq 0$ for which[3]

$$\sum_{i=1}^{N} X_{ij} = 1 \quad \text{and} \quad \Gamma_{ij}(t) = \sum_{k=1}^{N} X_{ik}\Gamma_{kj}\big(t'\big) \qquad [\text{for all } i, j \in \mathcal{C}]. \tag{36}$$

That is, a generalized stochastic system will generically be *non-Markovian* and *indivisible*, in a sense that parallels the notion of indivisibility for a generalized dynamical system, as introduced earlier in this paper.

- This notion of non-Markovianity represents a distinct way to generalize the Markovian case, as compared with most forms of non-Markovianity described in textbooks and in the research literature. According to those more traditional forms of non-Markovianity, one assumes that the set of times $\mathcal{T}$ has an ordering relation, and one further assumes the existence of higher-order conditional probabilities $p(i, t|j_1, t_1; j_2, t_2; \dots)$ that are conditioned on arbitrarily many times $t_1, t_2, \dots$. From that more traditional standpoint, the system is Markovian if and only if the latest conditioning time t_1 always screens off all the earlier conditioning times $t_2, \dots$, so that $p(i, t|j_1, t_1; j_2, t_2; \dots) = p(i, t|j_1, t_1)$. The definition of a generalized stochastic system presented in this paper does not assume the existence of such higher-order conditional probabilities in the first place.

- The symbol $\mathcal{A}$ will denote a commutative algebra of maps of the form

$$A : \mathcal{C} \times \mathcal{T} \to \mathbb{R} \tag{37}$$

under the usual rules of function arithmetic, and will be called the system's *algebra of random variables*. The individual values of each *random variable* $A \in \mathcal{A}$ will be labeled as

$$a_i(t) \equiv A(i, t) \in \mathbb{R} \quad [\text{for all } i \in \mathcal{C},\, t \in \mathcal{T}]. \tag{38}$$

Each such value $a_i(t)$ will be called the *magnitude* of the random variable A when the system is in its ith configuration at the time t. No assumption will be made here that these magnitudes are all distinct, even at any fixed time t.

[3] Note that if the quantities $\Gamma_{kj}(t')$ are regarded as forming an $N \times N$ stochastic matrix, a term to be defined shortly, then the inverse matrix, if it exists, will always have negative entries unless both matrices are permutation matrices. It follows that one cannot safely define the non-negative quantities X_{ij} by invoking inverse matrices.

- For the purposes of this paper, the algebra of random variables $\mathcal{A}$ will always be taken to be maximal, in the sense of containing every well-defined map of the form (37).

- Fixing the time t, the *expectation value* $\langle A(t)\rangle$ of a random variable A will denote its statistical average according to the probability distribution p at the time t:

$$\langle A(t)\rangle \equiv \sum_{i=1}^{N} a_i(t)p_i(t). \tag{39}$$

B. Ingredients from Linear Algebra

Given a generalized stochastic system $(\mathcal{C}, \mathcal{T}, \Gamma, p, \mathcal{A})$, with $\mathcal{C}$ a configuration space of finite integer size N, it will be convenient to introduce some formalism from linear algebra.

- Fixing the time t, let $p(t)$, called the system's *(time-dependent) probability vector*, denote the $N \times 1$ column vector whose ith entry is the ith standalone probability $p_i(t)$:

$$p(t) \equiv \begin{pmatrix} p_1(t) \\ \vdots \\ p_N(t) \end{pmatrix}. \tag{40}$$

- Again fixing the time t, let $\Gamma(t)$, called the system's *(time-dependent) transition matrix*, denote the $N \times N$ matrix for which the entry in the ith row, jth column is the transition probability $\Gamma_{ij}(t) \equiv p(i, t|j, 0)$:

$$\begin{aligned} \Gamma(t) &\equiv \begin{pmatrix} \Gamma_{11}(t) & \Gamma_{12}(t) & & \\ \Gamma_{21}(t) & \ddots & & \\ & & & \Gamma_{NN}(t) \end{pmatrix} \\ &= \begin{pmatrix} p(1,t|1,0) & p(1,t|2,0) & & \\ p(2,t|1,0) & \ddots & & \\ & & & p(N,t|N,0) \end{pmatrix}. \end{aligned} \tag{41}$$

It follows that for any time t, the transition matrix $\Gamma(t)$ is a *(column) stochastic matrix* in the mathematical sense, meaning that its entries are all non-negative real numbers,

$$\Gamma_{ij}(t) \geq 0 \quad [\text{for all } i, j \in \mathcal{C},\, t \in \mathcal{T}], \tag{42}$$

and that its columns each sum to 1, as required in (28). The initial condition (29) on the stochastic map Γ then becomes the statement that the transition matrix $\Gamma(0)$ at the initial time 0 is just the

$N \times N$ identity matrix $\mathbb{1}$:

$$\Gamma(0) = \mathbb{1} \equiv \begin{pmatrix} 1 & 0 & \\ 0 & \ddots & \\ & & 1 \end{pmatrix}. \tag{43}$$

Observe that the Bayesian marginalization condition (33) naturally takes the form of matrix multiplication:

$$p(t) = \Gamma(t)p(0) \quad [\text{for all } t \in \mathcal{T}]. \tag{44}$$

C. Connections with Other Constructions

There is a definite sense in which a generalized stochastic system $(\mathcal{C}, \mathcal{T}, \Gamma, p, \mathcal{A})$, as introduced in (24), provides a probabilistic extension of a generalized dynamical system $(\mathcal{X}, \mathcal{T}, f)$, as introduced in (1). Indeed, if the stochastic map $\Gamma : \mathcal{C}^2 \times \mathcal{T} \to [0, 1]$ from (25) outputs only the trivial probabilities 1 and 0, then Γ is effectively deterministic. In that case, one can naturally set $\mathcal{X} \equiv \mathcal{C}$ and define a dynamical map $f : \mathcal{X} \times \mathcal{T} \to \mathcal{X}$ according to

$$f(j, t) = i \quad \text{if and only if} \quad \Gamma((i, j), t) = 1 \quad [\text{for all } i, j \in \mathcal{X},\, t \in \mathcal{T}]. \tag{45}$$

A generalized stochastic system also naturally generalizes a mathematical structure called a *stochastic dynamical system*, or *random dynamical system* [12, 13].

- The definition of a random dynamical system starts with what this paper would call a generalized dynamical system $(\mathcal{X}, \mathcal{T}, f)$, with $\mathcal{T}$ a monoid, and replaces the single dynamical map f with a probabilistic family or ensemble of dynamical maps $\{f_\omega \,|\, \omega \in \Omega\}$ that are indexed by a point ω in some sample space Ω. That is, a randomly sampled point ω in Ω picks out an entire dynamical map f_ω, as a whole.

- For any randomly sampled point ω and any fixed time t, one can define a map

  $$f_{\omega,t} : \mathcal{X} \to \mathcal{X} \tag{46}$$

  by

  $$i \mapsto f_{\omega,t}(i) \equiv f_\omega(i, t) \quad [\text{for all } i \in \mathcal{X}]. \tag{47}$$

  A random dynamical system is then assumed to satisfy an initial condition that generalizes (6),

  $$f_{\omega,0} = \mathrm{id}_{\mathcal{X}} \quad [\text{for all } \omega \in \Omega], \tag{48}$$

as well as a Markov property that generalizes (23),

$$f_{\omega',t\star t'} = f_{\theta(\omega',t'),t} \circ f_{\omega',t'} \qquad \big[\text{for all } \omega' \in \Omega,\, t, t' \in \mathcal{T}\big]. \tag{49}$$

Here $\theta : \Omega \times \mathcal{T} \to \Omega$ is a map that is part of the definition of the random dynamical system, and specifies a deterministic rule for updating the sample point ω' and the dynamical map $f_{\omega'}$ to accommodate replacing the original initial time 0 with the effectively new initial time t'. That is, θ is necessary to account for any (purportedly deterministic) evolution in the underlying source of randomness itself.

- It follows from the foregoing definitions that the conditional probability $p(i, t|j, 0)$ for the system to be in its ith state at the time t, given that the system is in its jth state at the initial time 0, is obtained by adding up the probabilities for all the points ω in the sample space Ω whose corresponding dynamical maps f_ω take the state j at the initial time 0 and yield the state i at the final time t:

$$p(i, t|j, 0) = \text{probability}(\{\omega \in \Omega \,|\, f_{\omega,t}(j) = i\}) \qquad [\text{for all } i, j \in \mathcal{X},\, t \in \mathcal{T}]. \tag{50}$$

- Notice that a random dynamical system assigns a well-defined probability to any set of dynamical maps $\{f_\omega \,|\, \omega \in \Omega'\}$ that corresponds to a measurable subset Ω' of the sample space Ω. Thus, after fixing a state j at the initial time 0, a random dynamical system assigns a well-defined probability to any set of trajectories of the form $\{\{f_{\omega,t}(j) \,|\, t \in \mathcal{T}\} \,|\, \omega \in \Omega'\}$. (The nested subset notation here is just intended to make clear that this expression refers to an indexed set of indexed sets.) In particular, by generalizing the right-hand side of (50) to impose more conditions on the system's dynamical maps, one can define conditional probabilities $p(i_1, t_1; i_2, t_2; \ldots |j, 0)$ that involve arbitrarily many times. Random dynamical systems therefore contain far more information than the generalized stochastic systems defined in this paper, meaning that a generalized stochastic system indeed represents a less constrained, more general kind of mathematical structure.

In a sense, the absolutely *most general* kind of stochastic mathematical structure is a *stochastic process*, which essentially requires only the specification of a state space $\mathcal{X}$, a set of times $\mathcal{T}$, an initial probability distribution p, and a set of one or more time-dependent random variables $\mathcal{A}$. (See [14–17] for textbook treatments.) Importantly, the definition of stochastic process lacks the specification of a *dynamical law*. By requiring a dynamical law in the form of a stochastic map Γ, a generalized stochastic system $(\mathcal{C}, \mathcal{T}, \Gamma, p, \mathcal{A})$ is not quite as general as a stochastic process, but will still be general enough to encompass a large class of physical and mathematical models.

In particular, one can regard any *Markov chain* as a special case of a generalized stochastic system. The starting place is to assume that the set of times $\mathcal{T}$ is isomorphic to the integers, $\mathcal{T} \cong \mathbb{Z}$, with each time $t = n\,\delta t$ an integer number $n \in \mathbb{Z}$ of steps of some fixed, elementary time

scale δt. One then further assumes that for each integer n, the time-dependent transition matrix $\Gamma(n\,\delta t)$ originally defined in (41) can be expressed as the nth power of the transition matrix $\Gamma(\delta t)$ that implements the evolution for just the first time step δt:

$$\Gamma(n\,\delta t) = [\Gamma(\delta t)]^n \quad [\text{for all } n \in \mathbb{Z}]. \tag{51}$$

More broadly, a generalized stochastic system can therefore be understood as a kind of non-Markovian generalization of a Markov chain.

As this paper will also show, the class of generalized stochastic systems essentially includes all quantum systems as well.

D. Composite Systems and Subsystems

Introducing a notation of primes and tildes now to distinguish between specific generalized stochastic systems, a generalized stochastic system $\left(\tilde{\mathcal{C}}, \tilde{\mathcal{T}}, \tilde{\Gamma}, \tilde{p}, \tilde{\mathcal{A}}\right)$ will be called a *composite system* if its configuration space $\tilde{\mathcal{C}}$ is naturally expressible as a nontrivial Cartesian product of two sets $\mathcal{C}$ and $\mathcal{C}'$:

$$\tilde{\mathcal{C}} = \mathcal{C} \times \mathcal{C}'. \tag{52}$$

A composite system has the following salient features.

- Letting N denote the size of $\mathcal{C}$, and letting N' denote the size of $\mathcal{C}'$, the composite system's configuration space $\tilde{\mathcal{C}}$ has size $\tilde{N} \equiv NN'$.

- Labeling the elements of $\mathcal{C}$ by unprimed Latin letters $i, j, \ldots$, and labeling the elements of $\mathcal{C}'$ by primed Latin letters $i', j', \ldots$, the composite system's stochastic map $\tilde{\Gamma} : \tilde{\mathcal{C}}^2 \times \tilde{\mathcal{T}} \to [0, 1]$ has individual values that will be denoted by

$$\begin{aligned} \tilde{\Gamma}_{ii',jj'}(t) &\equiv \tilde{\Gamma}\big(\big((i,i'),(j,j')\big),t\big) \\ &\equiv \tilde{p}\big((i,i'),t|(j,j'),0\big) \\ &\quad [\text{for all } i,j \in \mathcal{C},\, i',j' \in \mathcal{C}',\, t \in \tilde{\mathcal{T}}]. \end{aligned} \tag{53}$$

  The normalization condition (28) on the stochastic map $\tilde{\Gamma}$ takes the form

$$\begin{aligned} &\sum_{i=1}^{N} \sum_{i'=1}^{N'} \tilde{\Gamma}_{ii',jj'}(t) = 1 \\ &\quad [\text{for all } j \in \mathcal{C},\, j' \in \mathcal{C}',\, t \in \tilde{\mathcal{T}}]. \end{aligned} \tag{54}$$

- The composite system's probability distribution $\tilde{p} : \tilde{\mathcal{C}} \times \tilde{\mathcal{T}} \to [0, 1]$ has individual values at

any fixed time t that will be denoted by

$$\tilde{p}_{ii'}(t) \equiv p\big((i,i'),t\big) \quad [\text{for all } i \in \mathcal{C},\, i' \in \mathcal{C}',\, t \in \tilde{\mathcal{T}}]. \tag{55}$$

The normalization condition (34) on the probability distribution $\tilde{p}$ takes the form

$$\sum_{i=1}^{N} \sum_{i'=1}^{N'} \tilde{p}_{ii'}(t) = 1 \quad [\text{for all } t \in \tilde{\mathcal{T}}], \tag{56}$$

and the Bayesian marginalization condition (33) becomes the definition

$$\tilde{p}_{ii'}(t) \equiv \sum_{j=1}^{N} \sum_{j'=1}^{N'} \tilde{\Gamma}_{ii',jj'}(t) \tilde{p}_{jj'}(0) \quad [\text{for all } i \in \mathcal{C},\, i' \in \mathcal{C}',\, t \in \tilde{\mathcal{T}}]. \tag{57}$$

Given a composite system $(\tilde{\mathcal{C}}, \tilde{\mathcal{T}}, \tilde{\Gamma}, \tilde{p}, \tilde{\mathcal{A}})$ with configuration space $\tilde{\mathcal{C}} = \mathcal{C} \times \mathcal{C}'$, as in (52), the composite system's set of times $\tilde{\mathcal{T}}$ trivially defines two sets of times $\mathcal{T}$ and $\mathcal{T}'$ according to

$$\mathcal{T} \equiv \mathcal{T}' \equiv \tilde{\mathcal{T}}. \tag{58}$$

Meanwhile, the composite system's probability distribution $\tilde{p}$ defines two probability distributions $p : \mathcal{C} \times \mathcal{T} \to [0,1]$ and $p' : \mathcal{C}' \times \mathcal{T}' \to [0,1]$ according to the respective marginalization formulas

$$p_i(t) \equiv \sum_{i'=1}^{N'} \tilde{p}_{ii'}(t) \quad [\text{for all } i \in \mathcal{C},\, t \in \mathcal{T}], \tag{59}$$

$$p'_{i'}(t) \equiv \sum_{i=1}^{N} \tilde{p}_{ii'}(t) \quad [\text{for all } i' \in \mathcal{C}',\, t \in \mathcal{T}']. \tag{60}$$

These two probability distributions p and p' do not necessarily turn $\mathcal{C}$ and $\mathcal{C}'$ into generalized stochastic systems of their own, in the sense of (24), due to the generic lack of well-defined stochastic maps $\Gamma : \mathcal{C}^2 \times \mathcal{T} \to [0,1]$ and $\Gamma' : \mathcal{C}'^2 \times \mathcal{T}' \to [0,1]$. Indeed, in place of the Bayesian marginalization condition (33), one instead has the relations

$$p_i(t) \equiv \sum_{i'=1}^{N'} \sum_{j=1}^{N} \sum_{j'=1}^{N'} \tilde{\Gamma}_{ii',jj'}(t) \tilde{p}_{jj'}(0) \quad [\text{for all } i \in \mathcal{C},\, t \in \mathcal{T}], \tag{61}$$

$$p'_{i'}(t) \equiv \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{j'=1}^{N'} \tilde{\Gamma}_{ii',jj'}(t) \tilde{p}_{jj'}(0) \quad [\text{for all } i' \in \mathcal{C}',\, t \in \mathcal{T}']. \tag{62}$$

The two sets $\mathcal{C}$ and $\mathcal{C}'$ will be called the configuration spaces of *subsystems* of the composite system $(\tilde{\mathcal{C}}, \tilde{\mathcal{T}}, \tilde{\Gamma}, \tilde{p}, \tilde{\mathcal{A}})$.

E. Unistochastic Systems

Recall that an $N \times N$ matrix U with complex-valued entries is called *unitary* if it satisfies

$$U^\dagger U = U U^\dagger = \mathbb{1}. \tag{63}$$

Here † denotes the adjoint operation, meaning complex conjugation combined with the transpose operation.

A stochastic matrix X is called *unistochastic* if each of its entries X_{ij} is the modulus-squared $|U_{ij}|^2$ of the corresponding entry U_{ij} of a unitary matrix U:

$$X = \begin{pmatrix} X_{11} & X_{12} & & \\ X_{21} & \ddots & & \\ & & & X_{NN} \end{pmatrix} = \begin{pmatrix} |U_{11}|^2 & |U_{12}|^2 & & \\ |U_{21}|^2 & \ddots & & \\ & & & |U_{NN}|^2 \end{pmatrix}. \tag{64}$$

The study of unistochastic matrices was initiated in a 1954 paper [18] by Horn, who originally called them 'ortho-stochastic matrices.' Unistochastic matrices were given their modern name in 1989 by Thompson [19, 20], and today *orthostochastic matrices* refer to the special case in which U can be taken to be a real orthogonal matrix, meaning a unitary matrix with entries that are all real-valued.

Every unistochastic matrix is *doubly stochastic*, or *bistochastic*, meaning that its columns *and* rows each sum to 1.[4] All 2×2 doubly stochastic matrices are unistochastic,[5] but this equivalence does not extend beyond the 2×2 case.[6] Again beyond the 2×2 case, not every unistochastic matrix is orthostochastic.[7]

A generalized stochastic system $(\mathcal{C}, \mathcal{T}, \Gamma, p, \mathcal{A})$ whose stochastic map Γ defines a unistochastic matrix $\Gamma(t)$ at every time t will be called a *unistochastic system*. One of the main goals of this paper will be to establish that the study of generalized stochastic systems essentially reduces to the study of unistochastic systems, which will also turn out to correspond to unitarily evolving quantum systems.

---

[4] Proof: $\sum_i |V_{ij}|^2 = \sum_i \overline{V_{ij}} V_{ij} = \left[V^\dagger V\right]_{jj} = 1$ and $\sum_j |V_{ij}|^2 = \sum_j V_{ij} \overline{V_{ij}} = \left[V V^\dagger\right]_{ii} = 1$, where bars denote complex conjugation. QED

[5] Proof: Every 2×2 doubly stochastic matrix is of the form $\left(\begin{smallmatrix} x & 1-x \\ 1-x & x \end{smallmatrix}\right)$, where $0 \leq x \leq 1$, and the matrix $\left(\begin{smallmatrix} \sqrt{x} & -\sqrt{1-x} \\ \sqrt{1-x} & \sqrt{x} \end{smallmatrix}\right)$ is easily seen to be unitary. QED

[6] For example, the 3×3 doubly stochastic matrix $\frac{1}{2}\left(\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{smallmatrix}\right)$ is provably not unistochastic.

[7] For example, the 3×3 matrix $\frac{1}{\sqrt{3}}\left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & z & z^2 \\ 1 & z^2 & z \end{smallmatrix}\right)$ is unitary for $z = \exp(2\pi i/3)$, so the 3×3 matrix $\frac{1}{3}\left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{smallmatrix}\right)$ is unistochastic. But this unistochastic matrix is not orthostochastic, because there does not exist a set of three mutually orthogonal 3×1 column vectors that feature only 1s and -1s as entries.

IV. THE STOCHASTIC-QUANTUM THEOREM

A. Statement of the Theorem

With all these preliminaries now in place, this paper's next major goal will be to prove the following theorem, which is a new result.

The Stochastic-Quantum Theorem
Every generalized stochastic system can be regarded as a subsystem of a unistochastic system. (65)

Remarkably, to study the class of generalized stochastic systems, this theorem implies that it suffices to restrict one's attention to the subclass of unistochastic systems.

B. Corollaries of the Theorem

The proof of the stochastic-quantum theorem (65) will involve the construction of a representation of the given generalized stochastic system in the formalism of Hilbert spaces, and will show that every generalized stochastic system corresponds to a unitarily evolving quantum system on a Hilbert space. This paper will therefore establish an important new correspondence between generalized stochastic systems and quantum systems, and thereby turn some of the puzzling axiomatic ingredients of quantum theory—the complex numbers,[8] Hilbert spaces, linear-unitary time evolution, and the Born rule in particular—into the output of a theorem.

One can also read this *stochastic-quantum correspondence* in the other direction, as the statement that all generalized stochastic systems can be modeled in terms of unitarily evolving quantum systems. From this perspective, unitarily evolving quantum systems actually represent the most general way to model a system with stochastic dynamical laws.

C. A Simple Example

The notion of embedding a generalized stochastic system into a unistochastic system does not always require an elaborate construction. As a simple example, consider a *discrete* Markovian dynamical system $(\mathcal{X}, \mathcal{T}, f)$ whose state space $\mathcal{X}$ has some finite size N, whose set of times $\mathcal{T}$ is isomorphic to the integers $\mathbb{Z}$ under addition, and whose dynamical map $f : \mathcal{X} \times \mathcal{T} \to \mathcal{X}$ satisfies the appropriate version of the Markov property (23):

$$f_{t+t'} = f_t \circ f_{t'} \quad \big[\text{for all } t, t' \in \mathcal{T} \cong \mathbb{Z}\big]. \tag{66}$$

[8] *Time-reversal transformations* for quantum systems are carried out by *anti-unitary operators*, which take the general form VK, where V is unitary and K is an abstract operator that carries out complex-conjugation. By definition, $K^2 = 1$ and $Ki = -iK$, where $i \equiv \sqrt{-1}$ is the imaginary unit. One can therefore show that i, K, and iK generate a Clifford algebra called the *pseudo-quaternions* [21]. As a consequence, there is a sense in which quantum systems are actually defined over the pseudo-quaternions, rather than merely over the complex numbers.

Based on these definitions and assumptions, it follows from the Markov property and the trivialization condition $f_0 = \mathrm{id}_{\mathcal{X}}$ from (6) that the time-dependent dynamical map f_t is invertible for any time $t \in \mathcal{T}$, with inverse given by

$$f_t^{-1} = f_{-t} \quad [\text{for all } t \in \mathcal{T} \cong \mathbb{Z}]. \tag{67}$$

Markovian dynamical systems of this kind provide a discretization of many of the kinds of deterministic, time-reversible physical systems that show up in classical physics.

Letting δt be the elementary duration of a single time step, and letting $n \in \mathbb{Z}$ denote an integer number of time steps, the system's specific state i at the time $t = n\,\delta t$ can be identified as an $N \times 1$ column vector with a 1 in its ith entry and 0s in all its other entries. Moreover, the time-dependent dynamical map $f_{n\,\delta t}$ can be expressed as the nth power of a fixed *permutation matrix* Σ, meaning a matrix consisting of only 1s and 0s, with a single 1 in each row and in each column.

Every permutation matrix is unitary, and is also unistochastic, because computing the modulus-squares of 1s and 0s gives back 1s and 0s. Hence, a discrete Markovian dynamical system defined in this way is already a unistochastic system, and therefore trivially satisfies the stochastic-quantum theorem (65).

Interestingly, there exists an analytic interpolation of this *discrete-time* unistochastic system that yields a corresponding *continuous-time* unistochastic system, at the cost of introducing non-trivial stochasticity into the intervals between the discrete time steps. This new unistochastic system is based on a time-dependent unistochastic matrix $\Gamma(t)$ whose entries are the modulus-squares of the corresponding entries of the $N \times N$ unitary time-evolution operator defined by $U(t) \equiv \Sigma^{t/\delta t}$, and therefore satisfies $\Gamma(n\,\delta t) = \Sigma^n$ for any integer n. One can even go on to define an $N \times N$ self-adjoint *Hamiltonian* H as the infinitesimal generator of $U(t)$, so there is ultimately a *Schrödinger equation* underlying this system.[9]

V. THE HILBERT-SPACE REPRESENTATION

A. The Time-Evolution Operator

To commence the proof of the stochastic-quantum theorem (65), one starts with a given generalized stochastic system $(\mathcal{C}, \mathcal{T}, \Gamma, p, \mathcal{A})$. The non-negativity (42) of the system's transition probabilities, $\Gamma_{ij}(t) \geq 0$, means that each transition probability can be written as the modulus-squared of a (non-unique) complex number $\Theta_{ij}(t)$:

$$\Gamma_{ij}(t) = |\Theta_{ij}(t)|^2 \quad [\text{for all } i, j \in \mathcal{C},\, t \in \mathcal{T}]. \tag{68}$$

[9] To define the unitary matrix $\Sigma^{t/\delta t}$, one starts by using the fact that every permutation matrix is also unitary to write the permutation matrix Σ as $V^\dagger D V$ for some unitary matrix V and some diagonal matrix D, where the entries of D are the eigenvalues of Σ. Because Σ is unitary, its eigenvalues are all phase factors $\exp(i\theta_m)$, where $m = 1, \ldots, N$, and where each phase θ_m is a real number. Setting all the eigenvalues to the power $t/\delta t$, so that they each take the new form $\exp(i\theta_m t/\delta t)$, one ends up with the matrix $\Sigma^{t/\delta t} \equiv V^\dagger D^{t/\delta t} V$, which is still unitary and now depends analytically on the time parameter t. The quantum-theoretic Hamiltonian matrix H for this system then has energy eigenvalues defined by $E_m \equiv -\hbar\theta_m/\delta t$ for $m = 1, \ldots, N$, and is diagonalized by the same unitary matrix V that diagonalizes Σ.

For each fixed time t, the complex numbers $\Theta_{ij}(t)$ collectively form their own $N \times N$ matrix, which will be called the system's *time-evolution operator* $\Theta(t)$:

$$\Theta(t) \equiv \begin{pmatrix} \Theta_{11}(t) & \Theta_{12}(t) & & \\ \Theta_{21}(t) & \ddots & & \\ & & & \Theta_{NN}(t) \end{pmatrix}. \tag{69}$$

(As a concession to terminological conventions, the terms 'matrix' and 'operator' will be used more-or-less interchangeably in what follows.)

The normalization condition (28) on the system's transition matrix $\Gamma(t)$ then becomes the summation condition

$$\sum_{i=1}^{N} |\Theta_{ij}(t)|^2 = 1 \quad [\text{for all } j \in \mathcal{C},\, t \in \mathcal{T}], \tag{70}$$

which can roughly be regarded as a generalization of a unitarity constraint. In keeping with the initial condition $\Gamma(0) = \mathbb{1}$ from (43), the time-evolution operator $\Theta(0)$ at the initial time 0 will be taken to be the $N \times N$ identity matrix $\mathbb{1}$:

$$\Theta(0) \equiv \mathbb{1} \equiv \begin{pmatrix} 1 & 0 & \\ 0 & \ddots & \\ & & 1 \end{pmatrix}. \tag{71}$$

B. The Hilbert Space

The time-evolution operator $\Theta(t)$ acts on an N-dimensional *Hilbert space* $\mathcal{H}$ defined as the space $\mathbb{C}^N$ of $N \times 1$ column vectors with complex-valued entries, together with the standard inner product $v^\dagger w$ for all $v, w \in \mathbb{C}^N$:

$$\mathcal{H} \equiv \mathbb{C}^N. \tag{72}$$

The standard orthonormal basis $e_1, \ldots, e_N$ is defined by

$$e_1 \equiv \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \ldots, \quad e_N \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{73}$$

These vectors are labeled as e_i, where each value of $i = 1, \ldots, N$ denotes a configuration in the system's configuration space $\mathcal{C}$, so this basis will be called the system's *configuration basis*.

There exists an associated set of rank-one projectors $P_1, \ldots, P_N$ defined by

$$P_i \equiv e_i e_i^\dagger = \operatorname{diag}(0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{th entry}}}{1}, 0, \ldots, 0) \quad [\text{for all } i \in \mathcal{C}]. \tag{74}$$

These *configuration projectors* $P_1, \ldots, P_N$ form a *projection-valued measure (PVM)* [22, 23], meaning that they satisfy the conditions of *mutual exclusivity*,

$$P_i P_j = \delta_{ij} P_i \quad [\text{for all } i, j \in \mathcal{C}], \tag{75}$$

with δ_{ij} again the usual Kronecker delta, and *completeness*,

$$\sum_{i=1}^{N} P_i = \mathbb{1}, \tag{76}$$

with $\mathbb{1}$ again the $N \times N$ identity matrix.

C. The Dictionary

It follows from a short calculation that the relationship $\Gamma_{ij}(t) = |\Theta_{ij}(t)|^2$ in (68) can be expressed in terms of a matrix trace as

$$\boxed{\Gamma_{ij}(t) = \mathrm{tr}\big(\Theta^\dagger(t) P_i \Theta(t) P_j\big)} \qquad [\text{for all } i, j \in \mathcal{C},\, t \in \mathcal{T}]. \tag{77}$$

This 'dictionary' essentially translates between the theory of generalized stochastic systems, as expressed by the left-hand side, and a corresponding Hilbert-space representation, as expressed by the right-hand side. This equation will turn out to form the core of a *stochastic-quantum correspondence* that will play a crucial role in the proof of the stochastic-quantum theorem.

D. The Density Matrix

Inserting the dictionary (77) into the Bayesian marginalization condition (33) yields the following equation:

$$p_i(t) = \mathrm{tr}(P_i \rho(t)) \quad [\text{for all } i \in \mathcal{C},\, t \in \mathcal{T}]. \tag{78}$$

Here $\rho(t)$, which will be called the system's *density matrix*, is an $N \times N$ time-dependent, self-adjoint, unit-trace, generically non-diagonal matrix defined for any fixed time t by

$$\rho(t) \equiv \Theta(t) \rho(0) \Theta^\dagger(t), \tag{79}$$

where its value at the initial time 0 is the following $N \times N$ diagonal matrix:

$$\rho(0) \equiv \sum_{j=1}^{N} p_j(0) P_j \equiv \begin{pmatrix} p_1(0) & 0 & \\ 0 & \ddots & \\ & & p_N(0) \end{pmatrix}. \tag{80}$$

Importantly, one sees from this analysis that the linearity of the Bayesian marginalization condition (33) underlies the linearity of the relationship (79) between the system's density matrix $\rho(0)$ at the initial time 0 and its density matrix $\rho(t)$ at other times t.

For any fixed time t, one can similarly express the expectation value (39) of a random variable A as

$$\langle A(t) \rangle = \mathrm{tr}(A(t)\rho(t)). \tag{81}$$

Here $A(t)$ denotes the $N \times N$ diagonal matrix

$$A(t) \equiv \sum_{i=1}^{N} a_i(t) P_i \equiv \begin{pmatrix} a_1(t) & 0 & \\ 0 & \ddots & \\ & & a_N(t) \end{pmatrix}. \tag{82}$$

Observe, in particular, that the magnitudes $a_1(t), \ldots, a_N(t)$ of the random variable A become the *eigenvalues* of the $N \times N$ matrix $A(t)$.

Notice also that the formula for the standalone probability $p_i(t)$ in (78) and the formula for the expectation value $\langle A(t) \rangle$ in (81) are both special cases of the *Born rule*.

E. An Aside on 'Classical Wave Functions'

Pausing for a moment, recall the smooth unistochastic interpolation of a discrete Markovian dynamical system described in Subsection IV C. For that system, the unitary time-evolution operator $U(t) \equiv \Sigma^{t/\delta t}$ trivializes to a permutation matrix Σ^n at every integer time step $n\,\delta t$. It follows that the system's density matrix $\rho(t)$, as defined in terms of its initial value $\rho(0)$ from (79), reduces to a diagonal matrix $\rho(n\,\delta t)$ at each integer time step. Taking the square root of each of its diagonal entries $p_1(n\,\delta t), \ldots, p_N(n\,\delta t)$, and allowing for arbitrary phase factors, one obtains a 'classical wave function' with components $\Psi_1(n\,\delta t), \ldots, \Psi_N(n\,\delta t)$ satisfying $|\Psi_i(n\,\delta t)|^2 = p_i(n\,\delta t)$ and capturing precisely the same information as the diagonal density matrix $\rho(n\,\delta t)$.

This classical wave function is the starting place for a representation of classical deterministic physics known popularly as the *'Koopman-von Neumann' formulation*, due to its superficial resemblance to work by Koopman [24] and von Neumann [25, 26] in the 1930s. However, as pointed out explicitly in [27], Koopman and von Neumann were actually using Hilbert spaces to represent observables, rather than to represent probability distributions. The formulation of classical physics in terms of 'classical wave functions' is more properly due to Schönberg [28], Loinger [29], Della Riccia and Wiener [30], and Sudarshan [31].

F. The Kraus Decomposition

Returning to the proof, and fixing the time t, and letting β denote an integer from 1 to N that will play a conceptually different role from a configuration index, let $K_\beta(t)$ denote the $N \times N$

matrix whose βth column agrees with the βth column of the time-evolution operator $\Theta(t)$, with 0s in all its other entries. That is, the entry in the ith row, jth column of $K_\beta(t)$ is given by

$$K_{\beta,ij}(t) \equiv \delta_{\beta j}\Theta_{ij}(t) \quad [\text{for all } \beta, i, j \in \mathcal{C},\, t \in \mathcal{T}]. \tag{83}$$

It follows from the summation condition (70) on $\Theta(t)$ that these new matrices satisfy the *Kraus condition*:

$$\sum_{\beta=1}^{N} K_\beta^\dagger(t) K_\beta(t) = \mathbb{1} \quad [\text{for all } t \in \mathcal{T}]. \tag{84}$$

Moreover, one can write the dictionary (77) as

$$\Gamma_{ij}(t) = \sum_{\beta=1}^{N} \mathrm{tr}\big(K_\beta^\dagger(t) P_i K_\beta(t) P_j\big) \tag{85}$$
$$[\text{for all } i, j \in \mathcal{C},\, t \in \mathcal{T}],$$

and one can express the time-evolution rule (79) for the system's density matrix as

$$\rho(t) = \sum_{\beta=1}^{N} K_\beta(t)\rho(0)K_\beta^\dagger(t). \tag{86}$$

The matrices $K_1(t), \ldots, K_N(t)$ are therefore *Kraus operators* [32], and (86) gives a *Kraus decomposition* of $\rho(t)$.

Abstracting these results, one obtains a *completely positive trace-preserving map*, or *quantum channel*,

$$\mathcal{E}_t : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}, \tag{87}$$

defined on the algebra $\mathbb{C}^{N\times N}$ of $N \times N$ matrices over the complex numbers according to

$$X \mapsto \mathcal{E}_t(X) \equiv \sum_{\beta=1}^{N} K_\beta(t) X K_\beta^\dagger(t). \tag{88}$$

G. Dilation

By the *Stinespring dilation theorem* [33, 34], the quantum channel (87) can be *purified*, meaning made into a form of *unitary* time-evolution. Specifically, for some integer $\tilde{N}$ in the *bounded* interval

$$N \leq \tilde{N} \leq N^3, \tag{89}$$

there exists an $\tilde{N} \times \tilde{N}$ unitary matrix $\tilde{U}(t)$ on a potentially enlarged or *dilated* Hilbert space

$$\tilde{\mathcal{H}} \equiv \mathbb{C}^{\tilde{N}} \tag{90}$$

such that the dictionary (77) can be written as[10]

$$\Gamma_{ij}(t) = \mathrm{tr}\Big(\mathrm{tr}'\Big(\tilde{U}^\dagger(t)[P_i \otimes \mathbb{1}']\tilde{U}(t)[P_j \otimes P'_{j'}]\Big)\Big) \tag{91}$$
$$[\text{for all } i, j \in \mathcal{C},\, t \in \mathcal{T}].$$

The formula (91) involves a number of ingredients.

- The dilated Hilbert space $\tilde{\mathcal{H}}$ is defined as the tensor product

$$\tilde{\mathcal{H}} \equiv \mathcal{H} \otimes \mathcal{H}', \tag{92}$$

  where $\mathcal{H} \equiv \mathbb{C}^N$ is the system's original Hilbert space (72), and where

$$\mathcal{H}' \equiv \mathbb{C}^{N'} \tag{93}$$

  is an *ancillary* Hilbert space whose dimension N' satisfies $\tilde{N} = NN' \leq N^3$. That is, N' is an integer lying in the bounded interval

$$1 \leq N' \leq N^2. \tag{94}$$

- The first trace $\mathrm{tr}(\cdots)$ in (91) denotes the *partial trace* over just the original Hilbert space $\mathcal{H}$, and the second trace $\mathrm{tr}'(\cdots)$ similarly denotes the partial trace over the ancillary Hilbert space $\mathcal{H}'$.

- Importantly, for each fixed time t, $\tilde{U}(t)$ is an $\tilde{N} \times \tilde{N}$ *unitary* matrix that reduces to the $\tilde{N} \times \tilde{N}$ identity matrix $\tilde{U}(0) = \tilde{\mathbb{1}}$ at the initial time 0.

- The symbol $\mathbb{1}'$ denotes the $N' \times N'$ identity matrix on the ancillary Hilbert space $\mathcal{H}'$.

- Letting $e'_1, \ldots, e'_{N'}$ denote the standard orthonormal basis for the ancillary Hilbert space $\mathcal{H}'$, in analogy with the configuration basis (73) for the system's original Hilbert space $\mathcal{H}$, and letting the primed Latin letters $i', j', \ldots$ each denote an element of an ancillary configuration

[10] 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 $\tilde{U}_{(i\beta m)(ja)}(t)$ by adding $N^3 - N^2$ additional columns that are mutually orthogonal with each other and with the previous N^2 columns already in $\tilde{V}(t)$, where the new index a runs through N^2 possible values. These additional columns can always be chosen so that at the initial time 0, where $V(0) = \mathbb{1}$ is the $N \times N$ identity matrix, they make $\tilde{U}(0)$ coincide with the $N^3 \times N^3$ identity matrix. The last step is to show that $\tilde{U}(t)$ satisfies (91), whose right-hand side reduces to $\sum_{\beta,m} |\tilde{U}_{(i\beta m)(jj')}(t)|^2 = \sum_\beta |K_{\beta,ij}(t)|^2$. QED

space $\mathcal{C}'$ consisting of the integers from 1 to N', the symbol $P'_{i'}$ denotes the rank-one projector

$$P'_{i'} \equiv e'_{i'} e'^{\dagger}_{i'} = \mathrm{diag}(0, \ldots, 0, \underset{\substack{\uparrow \\ i'\text{th entry}}}{1}, 0, \ldots, 0) \qquad \big[\text{for all } i' \in \mathcal{C}'\big], \tag{95}$$

which is an $N' \times N'$ diagonal matrix with a 1 in its i'th diagonal entry and 0s in all its other entries. These projectors form a projection-valued measure (PVM) on $\mathcal{H}'$ satisfying the conditions of mutual exclusivity,

$$P'_{i'} P'_{j'} = \delta_{i'j'} P'_{i'} \quad \big[\text{for all } i', j' \in \mathcal{C}'\big], \tag{96}$$

and completeness,

$$\sum_{i'=1}^{N'} P'_{i'} = \mathbb{1}'. \tag{97}$$

- Note that the left-hand side of (91) is insensitive to the specific choice of $j' \in \mathcal{C}'$ on the right-hand side.

Extending the foregoing construction, and fixing the time t, one obtains an $\tilde{N} \times \tilde{N}$ transition matrix given by the new dictionary

$$\tilde{\Gamma}_{ii',jj'}(t) \equiv \tilde{\mathrm{tr}}\big(\tilde{U}^{\dagger}(t) \tilde{P}_{ii'} \tilde{U}(t) \tilde{P}_{jj'}\big) \qquad \big[\text{for all } i, j \in \mathcal{C},\, i', j' \in \mathcal{C}',\, t \in \mathcal{T}\big]. \tag{98}$$

Here the trace is now over the dilated Hilbert space $\tilde{\mathcal{H}}$, and

$$\tilde{P}_{ii'} \equiv P_i \otimes \tilde{P}_{i'} \quad \big[\text{for all } i \in \mathcal{C},\, i' \in \mathcal{C}'\big] \tag{99}$$

defines a rank-one projector on $\tilde{\mathcal{H}}$.

Unfolding the notation, the dilated dictionary (98) reduces to the statement that

$$\tilde{\Gamma}_{ii',jj'}(t) = |\tilde{U}_{ii',jj'}(t)|^2 \qquad \big[\text{for all } i, j \in \mathcal{C},\, i', j' \in \mathcal{C}',\, t \in \mathcal{T}\big]. \tag{100}$$

That is, each entry $\tilde{\Gamma}_{ii',jj'}(t)$ of the $\tilde{N} \times \tilde{N}$ transition matrix $\tilde{\Gamma}(t)$ is the modulus-squared of the corresponding entry $\tilde{U}_{ii',jj'}(t)$ of an $\tilde{N} \times \tilde{N}$ *unitary* matrix $\tilde{U}(t)$. Again, a stochastic matrix with this special feature is called a *unistochastic matrix*.

H. The Dilated Generalized Stochastic System

One can now define a *dilated* generalized stochastic system $(\tilde{\mathcal{C}}, \tilde{\mathcal{T}}, \tilde{\Gamma}, \tilde{p}, \tilde{\mathcal{A}})$ in the following way.

- Let the dilated system's configuration space $\tilde{\mathcal{C}}$ be defined as the Cartesian product

$$\tilde{\mathcal{C}} \equiv \mathcal{C} \times \mathcal{C}', \tag{101}$$

where $\mathcal{C}$ is the original system's configuration space, with N elements labeled by unprimed Latin letters $i, j, \ldots$, and where $\mathcal{C}'$ is the configuration space of an ancillary subsystem, with N' elements labeled by primed Latin letters $i', j', \ldots$.

- Let the dilated system's set of times $\tilde{\mathcal{T}}$ be the original system's set of times $\mathcal{T}$:

$$\tilde{\mathcal{T}} \equiv \mathcal{T}. \tag{102}$$

- Let $\tilde{\Gamma} : \tilde{\mathcal{C}}^2 \times \tilde{\mathcal{T}} \to [0, 1]$ be the stochastic map defined according to the dilated dictionary (100). Then for each fixed time t, the $\tilde{N} \times \tilde{N}$ matrix $\tilde{\Gamma}(t)$ is unistochastic. Moreover, by construction, $\tilde{\Gamma}(t)$ satisfies the marginalization condition

$$\sum_{i'=1}^{N'} \tilde{\Gamma}_{ii',jj'}(t) = \Gamma_{ij}(t) \qquad \big[\text{for all } i, j \in \mathcal{C},\, j' \in \mathcal{C}',\, t \in \mathcal{T}\big], \tag{103}$$

where, as in (91), the value of $j' \in \mathcal{C}'$ is irrelevant.

- Let the map $\tilde{p} : \tilde{\mathcal{C}} \to [0, 1]$ be the probability distribution defined by

$$\tilde{p}_{ii'}(t) \equiv \sum_{j=1}^{N} \tilde{\Gamma}_{ii',jj'}(t) p_j(0) \qquad \big[\text{for all } i \in \mathcal{C},\, i', j' \in \mathcal{C}',\, t \in \mathcal{T}\big]. \tag{104}$$

It follows from the marginalization condition (103), together with the Bayesian marginalization condition (33), that

$$\sum_{i'=1}^{N'} \tilde{p}_{ii'}(t) = p_i(t) \quad [\text{for all } i \in \mathcal{C},\, t \in \mathcal{T}], \tag{105}$$

as in (59).

- Let the algebra $\tilde{\mathcal{A}}$ of random variables be the set of all maps of the form $\tilde{A} : \tilde{\mathcal{C}} \times \tilde{\mathcal{T}} \to \mathbb{R}$.

These results establish that $(\tilde{\mathcal{C}}, \tilde{\mathcal{T}}, \tilde{\Gamma}, \tilde{p}, \tilde{\mathcal{A}})$ is a composite unistochastic system, and that the original generalized stochastic system $(\mathcal{C}, \mathcal{T}, \Gamma, p, \mathcal{A})$ can be regarded as one of its subsystems. This conclusion completes the proof of the stochastic-quantum theorem (65). QED

I. The Corresponding Quantum System

Corresponding to the dilated unistochastic system $(\tilde{\mathcal{C}}, \tilde{\mathcal{T}}, \tilde{\Gamma}, \tilde{p}, \tilde{\mathcal{A}})$ is a quantum system based on a Hilbert space $\tilde{\mathcal{H}}$ of dimension $\tilde{N} \leq N^3$, with linear-unitary time evolution encoded in the unitary time-evolution operator $\tilde{U}(t)$.

Unistochastic matrices are not generally orthostochastic, meaning that they are not guaranteed to be expressible in terms of real orthogonal matrices. As a consequence, if one is given a generalized stochastic system whose $N \times N$ transition matrix $\Gamma(t)$ is *already* unistochastic, then there is no guarantee that the corresponding $N \times N$ unitary time-evolution operator $U(t)$ can be assumed to be a real orthogonal matrix. The stochastic-quantum correspondence therefore implies that in order to provide Hilbert-space representations for the most general kinds of generalized stochastic systems, the complex numbers $\mathbb{C}$ will be an important feature of quantum theory.

Whether in that case or more generally, of course, one is always free to start with an $N \times N$ time-evolution operator $\Theta(t)$ in (68) whose individual entries are all real. With that choice, the $\tilde{N} \times \tilde{N}$ unitary time-evolution operator $\tilde{U}(t)$ obtained from the Stinespring dilation theorem will likewise be real, and will therefore be an orthogonal matrix.

However, it is important to keep in mind that from the point of view of the stochastic-quantum correspondence, a generalized stochastic system's Hilbert-space representations are convenient fictions, and so one is entirely free to assume that they involve the complex numbers anyway. Assuming that a given choice of Hilbert-space representation is defined over the complex numbers, rather than merely over the real numbers, allows one to take advantage of the spectral theorem, eigenvectors of the time-evolution operator $\tilde{U}(t)$, and anti-unitary operators. Assuming appropriate smoothness conditions in time, one can further make use of a self-adjoint Hamiltonian $\tilde{H}(t)$ with real-valued energy eigenvalues, as well as the Schrödinger equation. To the extent that these mathematical constructs are often taken by textbooks to be indisputable features of quantum systems, the complex numbers become an avoidable part of quantum theory.[11]

VI. DISCUSSION AND FUTURE WORK

The unitarily evolving quantum system that lies on the other side of the stochastic-quantum correspondence is not limited to a *commutative* algebra of observables represented by operators that are diagonal in the configuration basis.

Indeed, as explained in other work [35], one can model the quantum measurement process of an observable represented by an *arbitrary* self-adjoint operator in terms of a unistochastic system that contains a subject system, a measuring device, and an environment, all as explicitly defined subsystems. By taking the measuring device's allowed configurations to correspond to definite readings of outcomes, and by taking the overall unistochastic system's transition matrix to be based

[11] Even if one assumes that a given choice of Hilbert-space representation is defined over the complex numbers, one can always double the dimension of the Hilbert space from N to $2N$ and represent the imaginary unit $i \equiv \sqrt{-1}$ by the 2×2 real matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, in which case all $N \times N$ unitary matrices become $2N \times 2N$ real orthogonal matrices. Interestingly, one can then also represent the complex-conjugation operation K needed for anti-unitary operators as a 2×2 real matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. However, this approach still ultimately preserves the algebraic structure of the complex numbers in the Hilbert space, in the sense that they will correspond to a subalgebra of matrices on the overall Hilbert space that commute with the Hamiltonian and all the system's observables. Hence, this approach does not truly eliminate the complex numbers from the Hilbert-space formalism of quantum theory.

on precisely the type of unitary time-evolution operator employed in standard textbook treatments of the measurement process, one inevitably finds that the measuring device ends up in one of its measurement-reading configurations with a stochastic probability given by the general form of the *Born rule*. Hence, in principle, one has access to a quantum system's entire *noncommutative* algebra of observables.

With the stochastic-quantum correspondence in hand, one can refer other exotic features of quantum systems back to their associated generalized stochastic systems to give those features a more physically transparent interpretation. For example, as shown in [35], *interference* and *entanglement* can be understood as artifacts of the generic indivisibility of a generalized stochastic system's dynamics. (It is also shown in [35] how this interpretative framework navigates the various *no-go theorems* that have appeared in the research literature over the past century.)

This overall approach to quantum foundations therefore sheds new light on some of the strangest features of quantum theory, in addition to suggesting novel applications of quantum computers. This approach might even provide a helpful stepping stone for the development of self-consistent generalizations of quantum theory itself.

ACKNOWLEDGMENTS

The author would like to thank Scott Aaronson, David Baker, Craig Callender, Iris Cong, Jordan Cotler, Jenann Ismael, David Kagan, Logan McCarty, Arjun Mirani, and Jill North for helpful conversations.

[1] R. Devaney. "An Introduction to Chaotic Dynamical Systems". 1989.

[2] S. H. Strogatz. *Nonlinear Dynamics and Chaos*. Westview Press, 1994.

[3] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, 1995.

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

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

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

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

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

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

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

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

[12] J. Honerkamp. *Stochastic Dynamical Systems: Concepts, Numerical Methods, Data Analysis*. Wiley-VCH, 1996.

[13] L. Arnold. *Random Dynamical Systems*. Springer, 1998.

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

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

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

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

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

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

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

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

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

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

[24] B. O. Koopman. "Hamiltonian Systems and Transformations in Hilbert Space". *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931. doi:10.1073/pnas.17.5.315.

[25] J. von Neumann. "Zur Operatorenmethode In Der Klassischen Mechanik". *Annals of Mathematics*, 33(3):587–642, 1932. doi:10.2307/1968537.

[26] J. von Neumann. "Zusatze Zur Arbeit 'Zur Operatorenmethode...' ". *Annals of Mathematics*, 33(4):789–791, 1932. doi:10.2307/1968225.

[27] T. F. Jordan and E. C. G. Sudarshan. "Dynamical Mappings of Density Operators in Quantum Mechanics". *Journal of Mathematical Physics*, 2(6):772–775, 1961. URL: http://scitation.aip.org/content/aip/journal/jmp/2/6/10.1063/1.1724221, doi:10.1063/1.1724221.

[28] M. Schönberg. "Application of Second Quantization Methods to the Classical Statistical Mechanics (II)". *Il Nuovo Cimento*, 10(4):419–472, April 1953. doi:10.1007/BF02781980.

[29] A. Loinger. "Galilei Group and Liouville Equation". *Annals of Physics*, 20(1):132–144, 1962. doi:10.1016/0003-4916(62)90119-7.

[30] G. Della Riccia and N. Wiener. "Wave Mechanics in Classical Phase Space, Brownian Motion, and Quantum Theory". *Journal of Mathematical Physics*, 7(8):1372–1383, August 1966. doi:10.1063/1.1705047.

[31] E. C. G. Sudarshan. "Interaction between classical and quantum systems and the measurement of quantum observables". *Pramana*, 6(3):117–126, March 1976. doi:10.1007/BF02847120.

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

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

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

[35] J. A. Barandes. "The Stochastic-Quantum Correspondence", 2023. URL: https://arxiv.org/abs/2302.10778, arXiv:2302.10778.