Standard Quantum Theory
Derived from First Physical Principles

Mehran Shaghaghi

University of Illinois-Chicago, Chicago 60612, IL, USA

The mathematical formalism of quantum theory has been established for nearly a century, yet its physical foundations remain elusive. In recent decades, connections between quantum theory and information theory have garnered increasing attention. This study presents a physical derivation of the mathematical formalism quantum theory based on information-theoretic considerations in physical systems. We postulate that quantum systems are characterized by single independent adjustable variables. Utilizing this physical postulate along with the conservation of total probability, we derive the standard Hilbert space formalism of quantum theory, including the Born probability rule. This comprehensive derivation of quantum theory offers a clear and concise physical foundation for the mathematical formalism of quantum mechanics.

1 Introduction

Quantum theory has been a successful mathematical framework for describing the behavior of quantum systems. Developed nearly a century ago, primarily by Dirac and von Neumann, it is based on Hermitian operators and their eigenvectors and eigenvalues [1, 2]. Despite its success, a fundamental physical foundation for quantum theory remains elusive. Since its inception, numerous efforts have been made to derive the formalism of quantum theory from physical axioms or first principles [3-20]. However, these attempts have either been incomplete or based on abstract mathematical assumptions that lack a clear physical basis.

In recent decades, there has been a growing interest in utilizing an information-theoretic approach to quantum theory. This interest is partly due to the advocacy of John Wheeler for the relevance of information theory for understanding quantum physics [21-23], as well as the advancements in the field of quantum information. Some efforts involve analyzing the internal structure and logic of the theory to identify its foundational principles. Others seek to gain new insights into the characteristics that shape quantum theory by examining the relationships between the various features of quantum systems. For example, Clifton, Bub, and Halvorson [17] used a C*-algebraic formalism to demonstrate how certain information-theoretic constraints on physical systems –such as no superluminal information transfer, no perfect broadcasting, and no bit commitment– lead to other features of quantum theory, including kinematic independence, noncommutativity, and nonlocality. Another approach, initiated by Hardy [16], focuses on the

general properties of probability theories and discusses the criteria that distinguish quantum theory from classical probability theories.

In this report, we present a derivation of the standard formalism of quantum theory based on a clear physical foundation. The fundamental concept we employ is the limited information capacity of quantum systems. Specifically, quantum systems are those physical systems that can hold only a single physical message (a piece of information) at a time, as formalized in the complementarity principle. This contrasts with classical systems, which can hold multiple messages simultaneously. The limits on the information capacity of a physical system can be understood as each independently adjustable variable of the system can physically represent one independent piece of information at a time. For physical systems with only one adjustable variable, this constraint prevents them from carrying more than one piece of information (one message) at a time.

A range of factors can restrict the adjustable variables of physical systems. For example, the adjustable variables can be confined by the constraints of the experimental setup, such as in the double-slit experiment where the use of a coherent beam of photons determines their energy and direction, leaving their polarization as the only adjustable variable. Alternatively, the adjustable variables of physical systems can be constrained under extreme conditions such as high pressures, low temperatures, or intense electromagnetic fields, as in laser-cooled trapped ions or superconductivity. In some cases, other independent adjustable variables of a system may be abstracted away, with specified properties (e.g., electron spin) considered isolated and separate from the rest. In this report, we demonstrate that the standard formalism of quantum mechanics, including the Hilbert space and the Born probability rule, can be explicitly and systematically derived as the theory of physical systems with a single adjustable variable.

2 Single Variable Systems: Information-Theoretic Considerations

Several information-theoretic definitions of a quantum system have been proposed in the literature as potential foundations for the derivation of quantum theory. For instance, Rovelli proposed the axiom that "there is a maximum amount of relevant information that can be extracted from a [quantum] system" [12]. Similarly, Zeilinger suggested "a foundational principle for quantum mechanics" that states a quantum system is "an elementary system [that] carries 1 bit of information" [14]. While these axioms offer some interesting explanations for certain quantum phenomena, such as randomness and entanglement [24-28], or a framework for reconstructing the formalism [12], they were not successful in deriving the full formalism of quantum theory.

Here, we present a mathematical theory that describes the behavior of systems with a single adjustable variable under measurements. Several information-theoretic considerations influence the properties of such systems. As these systems have no more than one adjustable variable, they

can contain only one message (i.e., one piece of information) at a time. We refer to these systems as Single Message (SM) systems. Given that an SM system contains only a single piece of information, its state can be defined by the outcome of the last measurement performed on it. It is important to note that performing a measurement always yields an outcome, even if the outcome is a zero reading.

The unique property of SM systems is that when the information content of an SM system is determined via a measurement, performing any subsequent independent measurement (which measures an independent proposition) must yield a result with zero informational gain, i.e, a random outcome. The equivalence of randomness and zero information gain is well established through Shannon's information measure [29]. This implies that in SM systems, performing independent measurements leads to random transformations of the system into new, unpredictable states. In other words, performing measurements on SM systems generally involves an element of randomness, as the system cannot hold more than one piece of information at any given time.

In summary, due to the single messaging capacity of the SM systems, performing independent measurements does not yield deterministic results but rather involves elements of unpredictability and randomness, similar to measuring the X component of an electron spin that is in the Z^+ state. This constraint implies that performing independent measurements on an SM system results in completely random changes in the system's state, making it impossible to predict measurement outcomes with certainty. However, if measurements are not entirely independent, the future states of the SM system can be predicted probabilistically based on the outcome of the last measurement performed on the system. Therefore, the dynamics of SM systems can be described probabilistically.

3 Construction of the Formalism

In analyzing the possible outcomes of an SM system in a given measurement scenario, it is crucial to examine the relationship between two measurements: the last measurement performed on the system (which has defined the state of the system) and the measurement for which we want to calculate the probabilities of its outcomes. These two measurements can either be dependent or independent. In the case of independent measurements, the outcome of the first measurement does not influence the outcome of the second, while for dependent measurements, certain outcomes of the second measurement can happen less or more likely based on the outcome of the first. A dependent measurement, for example, would be measuring the spin of an electron in the direction that is tilted 20 degrees from the z-axis in the zx-plane when it is in S_+^z state.

Consider an SM system in a certain state. The aim is to determine the probabilities of the system to end up in each of the outcomes of a measurement, M, to be performed on it. Without loss of generality, we consider measurements with N distinguishable outcomes (generalization to the

infinite case is straightforward). In what follows this notation is used: a measurement of type K is represented as M^K, with N possible outcomes $m_1^K, m_2^K, .., m_N^K$ that are independent members of the set $S(M^K)$ defined as:

$$S(M^K) = \{m_1^K, .., m_N^K\} \tag{1}$$

The probabilities of the SM system for the outcomes of this measurement can be represented as:

$$P(M^K) = \{P_1^K, .., P_N^K\} \tag{2}$$

Since performing the measurement eventually results in an outcome, it follows that:

$$\sum_{j=1}^{N} P_j^K = 1 \tag{3}$$

In certain cases, the probabilities, P_j^K, are easily determinable. For example, if the SM system has just undergone the measurement M^Q and the outcome m_i^Q is resulted, then repeating the same measurement M^Q will not change the state of the system. Therefore, in the case that the next measurement is M^Q, the probabilities of the system for the measurement outcomes are:

$$P_j^Q = \delta_{j,i} \tag{4}$$

On the other hand, in the case that the next measurement is an independent measurement type M^R, then the outcomes $\{m_1^R, .., m_N^R\}$ are all equally likely: $P_1^R = P_2^R = \cdots = P_N^R$. Hence, the probabilities of the system for the outcomes of the measurement can be written as:

$$P_j^R = \frac{1}{N}, \forall j \in \{1, \cdots, N\} \tag{5}$$

where N is the total number of the outcomes.

Besides these two cases, an extensible framework is required to evaluate the probabilities of the SM system for the outcomes of a general type of measurement, i.e., for measurements that are neither the same as the previous one nor fully independent of it. Consider two measurements, M^K & M^L, which are not fully independent, meaning certain outcomes of the second measurement

can happen less or more likely based on the outcome of the first. The interdependence between the outcomes of the measurements can be defined as:

$$T_{j,i}^{L,K} = P(m_j^L | m_i^K) \tag{6}$$

which represent the conditional probability of obtaining the jth outcome in measurement M^L, given the ith outcome of measurement M^K. These conditional probabilities can be framed in an $N \times N$ "interdependency matrix" of the two measurements.

The interdependency matrix has several properties. For example, for any fixed i we have

$$\sum_j T_{j,i}^{L,K} = 1 \tag{7}$$

as the conditional probabilities for a certain given event should add up to 1. In other words, the sum of the elements in each column equals 1. Representing the probabilities of the SM system for the measurement M^K as P_i^K, the probabilities of the system for the measurement M^L can be determined using standard probability formulas as

$$P_n^L = \sum_i P(m_n^L | m_i^K) P_i^K = \sum_i T_{n,i}^{L,K} P_i^K \tag{8}$$

Above sums the probabilities for all possible ways that the system can result in m_n^L. The interdependency matrices need to conserve the total probability, i.e., $\sum_n P_n^L = 1$. This is ensured by:

$$\sum_n P_n^L = \sum_n \sum_i T_{n,i}^{L,K} P_i^K = \sum_i P_i^K \sum_n T_{n,i}^{L,K} = 1 \tag{9}$$

based on (3) & (7).

Alternatively, one may consider the system probabilities for the measurement M^L, that is P_i^L, and seeks the probabilities of the system for the dependent measurement M^K outcomes using their conditional probabilities through the interdependency matrix of:

$$S_{j,i}^{K,L} = P(m_j^K | m_i^L). \tag{10}$$

The interdependency matrices represent the probabilistic correlations between the outcomes of two measurements. Functionally, they map the probabilities of the system for one measurement to that of the other as shown in (8). In general, these mapping matrices should have certain properties. Firstly, the components represent probabilities, therefore they are positive numbers not greater than 1:

$$0 \leq T_{j,i} \leq 1$$

$$0 \leq S_{j,i} \leq 1 \tag{11}$$

Additionally, they must obey:

$$\sum_j T_{j,i} = 1$$

$$\sum_j S_{j,i} = 1 \tag{12}$$

for any fixed i, as measurements necessarily produce an outcome (cf. (7)).

The interdependency matrices map the probabilities of the SM system from one measurement-space to another according to (8):

$$P_j^L = \sum_i T_{j,i} P_i^K$$

$$P_j^K = \sum_i S_{j,i} P_i^L \tag{13}$$

Therefore, the consecutive application of these reciprocal transformations on any initial state should map it back to itself, which means the following identity must hold for the interdependency matrices:

$$TS = ST = \boldsymbol{I} \tag{14}$$

in which I is the identity matrix. However, the current construction of the interdependency matrices fails to satisfy this property in general, since the mapping components are positive probability value and their multiplication does not sum to zero for the non-diagonal components of the identity matrix (i.e., TS and ST matrices). And the only instance where $S = T = I$ is the trivial scenario of identical measurements.

To address this issue, a different *probability measure* is needed that can accept non-positive inputs. To determine the probabilities in mappings between various measurement types, we need a probability measure that is a continuous function, accepting non-positive inputs and yielding values in the interval $[0,1]$, with $P(0) = 0$ and $P(1) = 1$. The power-function probability measures in the form of:

$$P(e) = |\rho(e)|^{\alpha}, \alpha \in \mathbb{R}^{+} \tag{15}$$

are based on the probability-intensities of events, $\rho(e)$. They represent the general form of probability measures that accommodate negative or complex-valued probability-intensities, as long as the resulting probabilities $P(e) = |\rho(e)|^{\alpha}$ fall within the unit interval $[0,1]$. The power function satisfies several important properties: it preserves probabilities at 0 and 1, it is a monotonic single-parameter function, and it maintains scaling invariance in analysis.

Using these probability measures, the updated transformations are as follows:

$$\sigma_j^L = \sum_i \Gamma_{j,i}\, \sigma_i^K$$

$$\sigma_j^K = \sum_i \Delta_{j,i}\, \sigma_i^L \tag{16}$$

These transformations map probability-intensities instead of probabilities, with $P_i^L = \left|\sigma_i^L\right|^{\alpha}$, $P_i^K = \left|\sigma_i^K\right|^{\alpha}$, $P(m_j^L|m_i^K) = \left|\Gamma_{j,i}\right|^{\alpha}$, and $P(m_j^K|m_i^L) = \left|\Delta_{j,i}\right|^{\alpha}$. We seek probability measures that allow invertible transformations conserving total probabilities in the bi-directional mappings. The value of the power α can be determined by imposing these requirements as follows. The bi-directional invertibility of the transformation matrices means:

$$\sum_i^N \Gamma_{a,i}\, \Delta_{i,b} = \sum_n^N \Delta_{a,n}\Gamma_{n,b} = \delta_{a,b} \tag{17}$$

Furthermore, the conditional probabilities should add up to 1 for any specific given event, (cf. (7)):

$$\sum_{n}^{N} P(m_n^K | m_i^L) = 1 = \sum_{n}^{N} |\Gamma_{n,i}|^{\alpha}$$

$$\sum_{i}^{N} P(m_i^L | m_n^K) = 1 = \sum_{i}^{N} |\Delta_{i,n}|^{\alpha}$$

(18)

In other words, the sum of the αth powers of the elements in each column equals 1. Writing that the i, i component of $\Delta\Gamma$ is equal to 1 along with the above relation leads to:

$$1 = \sum_{n}^{N} \Delta_{i,n} \Gamma_{n,i} = \sum_{n}^{N} |\Gamma_{n,i}|^{\alpha}$$

(19)

This should hold term by term in general, therefore we get:

$$\Delta_{i,n} = |\Gamma_{n,i}|^{(\alpha-1)}$$

(20)

Using the above equation along with the n, n component of $\Gamma\Delta$ leads to:

$$1 = \sum_{i}^{N} \Gamma_{n,i} \Delta_{i,n} = \sum_{i}^{N} \Gamma_{n,i} |\Gamma_{n,i}|^{(\alpha-1)} = \sum_{i}^{N} |\Gamma_{n,i}|^{\alpha}$$

(21)

This means that not only do the αth powers of the columns of the transformation matrix elements add up to 1, (see Eq.(18)), but also the sum of the αth powers of the rows is also 1. Inserting the result from Eq.(2020) into Eq.(18) leads to

$$1 = \sum_{i}^{N} |\Delta_{i,n}|^{\alpha} = \sum_{i}^{N} \left| |\Gamma_{n,i}|^{(\alpha-1)} \right|^{\alpha} = \sum_{i}^{N} |\Gamma_{n,i}|^{(\alpha^2-\alpha)}$$

(22)

The last two equations hold only if $\alpha^2 - \alpha = \alpha$, which determines that $\alpha = 2$. This means that in each row and each column of the mapping, the squares of elements add up to 1, indicating that the mappings are unitary matrices.

In conclusion, the only probability measure that ensures consistent mappings between different measurements probability spaces is:

$$P(e) = |\rho(e)|^2 \tag{23}$$

This square probability measure is based on probability amplitudes. It implies that the interdependency matrices are unitary matrices, which transform the probability amplitudes of the SM system for measurements, σ_j^I, defined as:

$$P_j^I = \left|\sigma_j^I\right|^2 \tag{24}$$

and conserve the sum of their squares:

$$\sum_j \left|\sigma_j^I\right|^2 = \sum_j P_j^I = 1 \tag{25}$$

This expression represents the conservation of total probability. The components of the mappings are also confined as follows:

$$0 \leq \left|\,\Gamma_{j,i}\right|^2 = \left|\rho\big(m_j^L\big|m_i^K\big)\right|^2 = P\big(m_j^L\big|m_i^K\big) \leq 1$$

$$0 \leq \left|\,\Delta_{j,i}\right|^2 = \left|\rho\big(m_j^K\big|m_i^L\big)\right|^2 = P\big(m_j^K\big|m_i^L\big) \leq 1 \tag{26}$$

These conditions allow the components to take negative or complex values. Similar to the earlier construct, these interdependency matrices are used to pursue the probabilities of the SM system. However, instead of operating on the SM system probabilities, P_j^I, these mappings operate on the SM system probability amplitudes, σ_j^I.

With these adjustments, the updated interdependency matrices can be used as before to determine the probabilities of the system for a second measurement. The interdependency matrices $\Gamma^{L,K}$ an $\Delta^{K,L}$ transform the SM system's probability amplitudes between the two measurements M^K and M^L. The $\Gamma^{L,K}$ matrix maps the probability amplitudes from measurement M^K to measurement M^L, and $\Delta^{K,L}$ maps those from measurement M^L to measurement M^K, according to:

$$\sigma_j^L = \sum_i \Gamma_{j,i} \sigma_i^K$$

$$\sigma_j^K = \sum_i \Delta_{j,i} \sigma_i^L$$

(27)

These unitary mappings conserve the total probability, i.e., $\sum_j |\sigma_j^L|^2 = \sum_j |\sigma_j^K|^2 = 1$. Furthermore, these unitary interdependency matrices are inverse of each other and are conjugate transposes of each other:

$$\Delta = \Gamma^{-1} = \Gamma^*$$

$$\Gamma = \Delta^{-1} = \Delta^*$$

(28)

In other words, the two interdependency matrices between the measurements are conjugate transposes of each other. This implies that for SM systems, the conditional probability amplitudes between any pair of measurement I and II outcomes are related as follows:

$$\rho(m_b^{II} | m_a^I) = \rho^*(m_a^I | m_b^{II}).$$

(29)

In summary, in the above analysis by investigating the constraints on transforming probabilities of SM systems between measurements, the characteristics of probability mappings in these systems were derived. The standard probability measure used in classical physics was found to be inadequate for consistent determination of probabilities in SM system mappings. To address that, alternative probability measures were sought that could accommodate negative or complex probability-intensities. Demanding conservation of total probability in the bi-directional mappings led to the development of a probability measure based on probability amplitudes and the unitarity of the interdependency matrices. Throughout the analysis, the primary constraint was the conservation of total probability in invertible mappings, with the transformation properties adhering to this consistency requirement.

State Vectors, Operator Algebra, Hilbert Space Representation, and the Probability Rule

The Hilbert-space formalism of SM systems theory is easily recognizable in the above construct, with clear denotations of its elements. The probability amplitude of the SM system, denoted as σ^I, can be regarded as a vector of length 1 in the N-dimensional space defined by the N

independent outcomes of the measurement M^I, namely $m_1^I, m_2^I, .., m_N^I$ (as in (1)). Employing the conventional bra-ket notation, the state-vector of the SM system can be expressed as $|\sigma^I\rangle = \sum_i \sigma_i^I |m_i^I\rangle$ with the following normalization (Eq.(25)):

$$\langle \sigma^I | \sigma^I \rangle = 1. \tag{30}$$

With this representation of the SM states as unit vectors in the complex vector space of the probability amplitudes, the algebraic structure of SM systems is apparent. Once a measurement basis is chosen to represent the SM state, e.g., $|\sigma^I\rangle$, the probability amplitude of the system for another measurement can be expressed as a linear combination of those bases, using the interdependency of the two measurements, according to (27) as:

$$|\sigma^{II}\rangle = \rho^{II,I} |\sigma^I\rangle \tag{31}$$

in which $\rho^{II,I}$ is the unitary transformation portrayed in (29). In other words, mappings of the state of SM systems between measurements are carried out by the interdependency matrices of the measurements. The logic is simple: the dependence of the states can be determined by the dependence of the measurements that would produce those states since SM states are defined by the outcome of measurements.

The Born rule for calculating probabilities is also clear in this representation. The probability amplitude of a measurement outcome, m_b^{II}, given the initial state of the SM system, $\sigma^I = m_a^I$, determined from (31) is

$$\rho_{b,a}^{II,I} = \langle \sigma_a^I | \sigma_b^{II} \rangle = \langle m_a^I | m_b^{II} \rangle \tag{32}$$

and applying (23) leads to the probability of the described event as:

$$P_{b,a}^{II,I} = \left| \rho_{b,a}^{II,I} \right|^2 = \left| \langle \sigma_a^I | \sigma_b^{II} \rangle \right|^2 = \left| \langle m_a^I | m_b^{II} \rangle \right|^2 \tag{33}$$

This result is the Born probability rule; it is a built-in part of the theory rooted in the conservation of total probability in transformations.

In summary, the Hilbert space formalism of SM systems theory allows us to represent the state of the system as a unit vector in the complex vector space of probability amplitudes. The state-vector can be expressed in terms of a chosen measurement basis and can be transformed between measurements probability spaces using the unitary interdependency matrices. The Born rule for calculating probabilities is naturally derived from the probability amplitudes, and total probability is conserved in all transformations. Overall, this formalism provides a clear and powerful framework for understanding the behavior of SM systems.

Superposition of Possibilities and the Interference Effect

In the above formulation, the transformation of the system probability amplitudes under a series of measurements can be described by consecutively applying the interdependency matrices of those measurements, as expressed in:

$$|\sigma^{III}\rangle = \rho^{III,II}|\sigma^{II}\rangle = \rho^{III,II}\rho^{II,I}|\sigma^{I}\rangle$$
$$= \rho^{III,I}|\sigma^{I}\rangle \tag{34}$$

where

$$\rho_{f,i}^{III,I} = \sum_{n} \rho_{f,n}^{III,II}\rho_{n,i}^{II,I} \tag{35}$$

describes the relationship between the interdependency matrices of the measurement.

This chain rule allows the determination of the system's probability amplitude for a measurement based on the interdependencies of measurements. Importantly, this sum accounts for the interference effects: the probability of events, calculated according to:

$$P\left(m_f^{III}\big|m_i^I\right) = P_{f,i}^{III,I} = \left|\rho_{f,i}^{III,I}\right|^2 = \left|\sum_{n} \rho_{f,n}^{III,II}\rho_{n,i}^{II,I}\right|^2 \tag{36}$$

can include extra terms, "the interference terms," that lead to results different from what the classical method of calculating probabilities predicts. For example, using the above relation, the interference in the double-slit experiment follows as the photons have two options (slit1 and slit2) to go from the source (O) to the screen (S):

$$P\left(m_S^{Screen}\big|m_O^{Source}\right) = \left|\rho_{S,O}^{Screen,Source}\right|^2$$
$$= \left|\rho_{S,slit1}^{Screen,Slit}\rho_{slit1,O}^{Slit,Source} + \rho_{S,slit2}^{Screen,Slit}\rho_{slit2,O}^{Slit,Source}\right|^2 \tag{37}$$

This differs from the classical method of calculating probabilities, which predicts:

$$P\left(m_S^{Screen}\big|m_O^{Source}\right) = P(S|slit1)P(slit1|O) + P(S|slit2)P(slit2|O)$$
$$\equiv \left|\rho_{S,slit1}^{Screen,Slit}\rho_{slit1,O}^{Slit,Source}\right|^2 + \left|\rho_{S,slit2}^{Screen,Slit}\rho_{slit2,O}^{Slit,Source}\right|^2 \tag{38}$$

The transformation of the probability amplitudes between the measurements, as described by the chain rule in (34), highlights interference as a fundamental characteristic of SM systems. It can

be seen that it is the connection between the probability amplitudes of the outcomes of different measurements, rather than their probabilities, that modifies the result from classical probability. The above derivation provides a clear understanding of the basis of interference and demonstrates the essential role of the presence of the intermediate measurement. In the physics of SM systems, interference results from the superposition of possible outcomes of the intermediate measurement, indicating that the presence of intermediate non-performed measurements (the "interaction-free" measurements) cannot be neglected. It is also evident that interference results from the mathematical record-keeping of SM systems' probability amplitudes for possible outcomes of intermediate measurements rather than the physical occurrence of those outcomes.

Time Evolution and Derivation of the Schrödinger Equation

The analysis of the mappings between the probability spaces of different measurements for SM systems led to the algebraic structure of their state-space, specifically, the Hilbert space and the Born probability rule. Incorporating time evolution in this framework is a straightforward process, as discussed in various mathematical physics textbooks (see for example [30] Sec. 3.3).

A measurement can be labeled by the time variable t, denoting the time at which it is performed. Since the relationship between measurements does not depend on time, their interdependency matrices must maintain the same structure at different times; hence, there must exist a unitary transformation $\mathbb{T}(t_2 - t_1)$ such that:

$$\rho(t_2) = \mathbb{T}^{-1}(t_2 - t_1)\,\rho(t_1)\mathbb{T}(t_2 - t_1) \tag{39}$$

Under common assumptions about time evolution, the transformation can be written as:

$$\mathbb{T}(t_2 - t_1) = e^{-iH(t_2 - t_1)} \tag{40}$$

where H is a self-adjoint matrix that defines the Hamiltonian in the Hilbert space. Borrowing conventional quantum terminology, thus far, our discussion was in the "Heisenberg picture," in which the states of isolated systems remain fixed, but the interdependency matrices that represent the observables change with time. Transposing to the "Schrödinger picture" shifts the focus to the time evolution of the state, and the Schrödinger equation is obtained as follows:

$$i\frac{d}{dt}|\sigma^I(t)\rangle = H|\sigma^I(t)\rangle \tag{41}$$

****

Here, we showed how the analysis of the constraints imposed by the probabilistic nature of SM systems under measurements leads to the derivation of the standard Hilbert-space formalism of

quantum mechanics, as well as the Born probability rule. The full framework of quantum theory emerged from studying the general properties of the mappings that transform the state of the SM system from one measurement probability-space to another. Using this perspective, the elements of the theory can be properly understood.

The state-vector, σ^I, contains information about the outcome of the last measurement performed on the system. The unitary interdependency matrices, $\rho^{II,I}$, transform the system state-vector, σ^I, between different measurement probability spaces while conserving the total probability. These matrices embody information on how the outcomes of distinct types of measurements correlate with one another probabilistically. Accordingly, the state-vector can be mapped into different measurements probability spaces via the interdependencies of the measurements. The transformed state-vectors represent the probability amplitudes of the system for those measurements. The construct of the theory ensures that the total probability is conserved in these transformations. Finally, when a measurement is performed, the state of the system adjusts accordingly to reflect the result of the observation.

4 Discussion and Conclusion

In this report, we have presented a systematic derivation of the standard formalism of quantum theory from a physical foundation. The underlying physical idea in our derivation is the recognition of physical systems with a single adjustable variable and their inherently probabilistic nature due to their limited capacity to carry messages. This is whence the indeterministic nature of quantum mechanics arises. Based on this physical foundation, we have derived quantum theory in a transparent and intuitive manner. Similar postulates have been suggested in the past [12, 14], however, they did not succeed in deriving the full formalism of the theory.

This derivation shows that quantum mechanics describes the physics of systems that possess only a single adjustable variable. The theory describes the transformations of the probability amplitudes of such systems for different measurements, with the algebraic structure of the theory rooted in the conservation of total probability in these transformations. The current derivation of the quantum formalism for finite-dimensional individual systems can easily be extended to the general case. It is worth noting that this derivation places the ensemble interpretation of quantum mechanics [31] as a secondary interpretation of the theory, rather than its primary one.

In addition, this derivation clarifies the fundamental concept of state in quantum theory. The state of a quantum system is a mathematical representation of the propensity of the physical system for various measurement outcomes, which all depend on the outcome of the most recent measurement performed on the system. The theory provides a way to calculate the system's probabilities for future measurement outcomes from its current state, taking into account the interdependencies of the measurements. This perspective resolves misconceptions about the state of a quantum system, namely that it stores all information about the system's representations in

various measurements [32]. Rather, this information is contained in the interdependencies of measurements, not in the quantum system itself.

The interpretation of quantum mechanics becomes clear from our derivation. The foundation of quantum theory lies in the concept that a quantum system is a physical system with no more than one independent adjustable variable. Since such systems can only contain one proposition of objective reality, the theory is inherently probabilistic rather than deterministic. At its essence, quantum theory is a mathematical framework for calculating the probabilities of a measurement's outcomes. It describes how the probabilities of single-variable systems transform among different measurements probability spaces. Rather than describing the measurement process, the theory focuses on what can be known about the potential results of measurements.

Our work does not change the existing formalism of quantum theory. Instead, it provides a comprehensive framework for interpreting quantum phenomena and represents a significant step towards a deeper understanding of the theory. Quantum mechanics is the probability theory for physical systems that possess a single adjustable variable. The core mathematical structure of the theory is based on consistent record-keeping of probabilities between different measurements. Understanding the physical foundation of quantum theory allows us to revisit the phenomena described by conventional quantum mechanics and gain deeper insights into the nature of our world.

References

[1]   P. A. M. Dirac, *The principles of quantum mechanics* (no. 27). Oxford university press, 1981.
[2]   J. Von Neumann, *Mathematical foundations of quantum mechanics.* Princeton university press, 2018.
[3]   G. Birkhoff and J. Von Neumann, "The logic of quantum mechanics," *Annals of mathematics,* pp. 823-843, 1936.
[4]   G. W. Mackey, *Mathematical foundations of quantum mechanics.* Courier Corporation, 2013.
[5]   C. Piron, "Axiomatique quantique," *Helvetica physica acta,* vol. 37, no. 4-5, pp. 439-&, 1964.
[6]   C. Piron, "Survey of general quantum physics," *Foundations of physics,* vol. 2, no. 4, pp. 287-314, 1972.
[7]   G. Ludwig, "Attempt of an axiomatic foundation of quantum mechanics and more general theories. III," *Communications in Mathematical Physics,* vol. 9, no. 1, pp. 1-12, 1968.
[8]   B. Mielnik, "Geometry of quantum states," *Communications in Mathematical Physics,* vol. 9, no. 1, pp. 55-80, 1968.
[9]   A. Landé, "Albert Einstein and the quantum riddle," *American Journal of Physics,* vol. 42, no. 6, pp. 459-464, 1974.
[10]   D. I. Fivel, "How interference effects in mixtures determine the rules of quantum mechanics," *Physical Review A,* vol. 50, no. 3, p. 2108, 1994.
[11]   L. Accardi, "Can mathematics help solving the interpretational problems of quantum theory?," *Il Nuovo Cimento B (1971-1996),* vol. 110, no. 5, pp. 685-721, 1995.
[12]   C. Rovelli, "Relational quantum mechanics," *International Journal of Theoretical Physics,* vol. 35, no. 8, pp. 1637-1678, 1996.
[13]   N. Landsman, "Simple new axioms for quantum mechanics," *International journal of theoretical physics,* vol. 37, no. 1, pp. 343-348, 1998.

[14]   A. Zeilinger, "A foundational principle for quantum mechanics," *Foundations of Physics,* vol. 29, no. 4, pp. 631-643, 1999.

[15]   B. Coecke, D. Moore, and A. Wilce, "Current research in operational quantum logic," in *Series Fundamental Theories of Physics*, vol. 111: Springer, 2000.

[16]   L. Hardy, "Quantum theory from five reasonable axioms," *arXiv preprint quant-ph/0101012,* 2001.

[17]   R. Clifton, J. Bub, and H. Halvorson, "Characterizing quantum theory in terms of information-theoretic constraints," *Foundations of Physics,* vol. 33, no. 11, pp. 1561-1591, 2003.

[18]   G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Informational derivation of quantum theory," *Physical Review A,* vol. 84, no. 1, p. 012311, 2011.

[19]   L. Masanes and M. P. Müller, "A derivation of quantum theory from physical requirements," *New Journal of Physics,* vol. 13, no. 6, p. 063001, 2011.

[20]   H. F. Hofmann, "Derivation of quantum mechanics from a single fundamental modification of the relations between physical properties," *Physical Review A,* vol. 89, no. 4, p. 042115, 2014.

[21]   J. A. Wheeler, "World as system self-synthesized by quantum networking," in *Probability in the Sciences*: Springer, 1988, pp. 103-129.

[22]   J. A. Wheeler, "Toward "it from bit"," *Quantum Coherence,* p. 281, 1992.

[23]   J. A. Wheeler, "Information, physics, quantum: the search for links," in *Feynman and computation: exploring the limits of computers*, 1999, pp. 309-336.

[24]   C. Brukner and A. Zeilinger, "Malus' law and quantum information," *acta physica slovaca,* vol. 49, pp. 647-652, 1999.

[25]   C. Brukner, M. Zukowski, and A. Zeilinger, "The essence of entanglement," *arXiv preprint quant-ph/0106119,* 2001.

[26]   C. Brukner and A. Zeilinger, "Young's experiment and the finiteness of information," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences,* vol. 360, no. 1794, pp. 1061-1069, 2002.

[27]   Č. Brukner and A. Zeilinger, "Information and fundamental elements of the structure of quantum theory," in *Time, quantum and information*: Springer, 2003, pp. 323-354.

[28]   Č. Brukner, M. Aspelmeyer, and A. Zeilinger, "Complementarity and information in "delayed-choice for entanglement swapping"," *Foundations of Physics,* vol. 35, no. 11, pp. 1909-1919, 2005.

[29]   C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal,* vol. 28, no. 4, pp. 656-715, 1949.

[30]   W. Thirring and E. M. Harrell, *A course in mathematical physics 3: Quantum mechanics of atoms and molecules*. Springer, 1981.

[31]   L. E. Ballentine, "The statistical interpretation of quantum mechanics," *Reviews of modern physics,* vol. 42, no. 4, p. 358, 1970.

[32]   C. M. Caves and C. A. Fuchs, "Quantum information: How much information in a state vector?," *arXiv preprint quant-ph/9601025,* 1996.