

Quantum behavior of the systems with a single degree of freedom and the derivation of quantum theory

Mehran Shaghghi

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

The number of independent messages a physical system can carry is limited by the number of its adjustable properties. In particular, the systems with just one adjustable property cannot carry more than a single message at a time. We demonstrate this is the case for the single photons in the double-slit experiment, and the root of the fundamental limit on measuring the complementary aspect of the photons. Next, we analyze the other “quantal” behavior of the systems with a single adjustable property. Finally, we formulate a mathematical theory to describe the dynamics of such systems and derive the standard Hilbert-space formalism of quantum mechanics. Our derivation demonstrates the underlying physical foundation of the quantum theory.

Introduction

The single-photon Young’s double-slit experiment is the ideal presentation of the dual wave-particle nature of the light. In the basic version of this experiment, a coherent light beam illuminates a plate pierced by two parallel slits, and the light passing through the slits is observed on a screen behind the plate forming an interference pattern. The experiment performed by sending single photons shows they also collectively construct an interference pattern [1-4] that is incompatible with the pattern of single particles that go through either of the slits. This apparent contradiction has served as the quintessential example of the wave-particle complementarity encountered in the quantum world. Quoting Feynman, this experiment is a phenomenon “which has in it the heart of quantum mechanics. In reality, it contains the only mystery” of the theory [5].

Many studies have confirmed this feature of the photons’ behavior in the double-slit experiment: an experiment designed to observe the wave-like interference necessarily gives up the option of observing the particle-like trajectories and vice versa [6, 7]. Any effort to force the observation of both effects introduces an element of randomness that makes the results non-conclusive [8]. The quantum mechanical statement of this happening is that it is only possible to observe one of the complementary wave or particle properties of the photons at a time, but not both, due to the complementarity principle. No further explanation is provided, and the complementarity principle is regarded as an essential feature of quantum mechanics (QM), demarcating between quantum and classical physics.

In this paper, we first show that the complementarity principle in the double-slit experiment results from information consideration of the interacting photons. We start by analyzing the properties of the physical systems to carry pieces of information. In particular, those with the capacity for only one message cannot carry more than an independent message at a time. We will show this is the case for the photons in the double-slit experiment and this limitation is the underlying reason behind the complementarity in the double-slit experiment. Next, we discuss other peculiarities in the behavior of such systems and their similarities with quantum systems. Finally, we develop a theory to analyze the behavior of the systems with a single messaging capacity under possible measurements and derive the standard formalism of quantum theory.

The communicational properties of physical systems

Consider a physical system with several adjustable variables. In principle, any property of a physical system that can be adjusted may be used as a substrate to load a message upon. For example, a token with 2 possible faces (“head” or “tail”), 8 possible colors, and 4 possible sizes to choose from, can be used to convey three independent messages: a 1-bit message using its face, a 3-bit message via its color, and finally a 2-bit message through its chosen size. We define the *messaging capacity* of a physical system as the total number of independent messages the system can contain. Each independently adjustable property of a physical system can convey one message; therefore, the *messaging capacity* of a system is determined by the number of its independently adjustable properties:

$$MC = J, \quad (1)$$

in which J is the total number of the independently adjustable properties of the system. This concept should not be confused with the *information capacity*¹ of a physical system, which is a mathematical measure of its number of adjustable states in bits

$$IC = \log_2(d_{tot}) \quad (2)$$

in which d_{tot} is the total number of adjustable states of the system. Accordingly, the bit value of a message loaded on an adjustable property of a system with d_j states is $\log_2(d_j)$ bits. In our example, the token can be in each of the $2 \times 8 \times 4 = 64$ different possible states, which leads to the total information capacity of $\log_2(64) = 6$ bits, which, in turn, is the sum of the bit values of the three messages it can convey.

The information capacity of a physical system is conceptually different from its messaging capacity. The information capacity of a system is a mathematical construct that measures the number of its possible states in the logarithmic scale of bits. In the physical world, however, the messages are not necessarily stored and conveyed in bits – the binary (0 or 1) states of the systems – as in nature the adjustable properties of the systems generally contain more than just two states. Meaning that unlike the abstract world of mathematics, where conventionally information is treated as units of bit, in the physical world, the messages get communicated in various non-integer chunks of bits. Furthermore, the messaging capacity of a physical system is a discrete-valued aspect of the system that directly represents the number of its independent variables. Information capacity, on the other hand, deals with the available states of those variables. That means for a variable with a continuous state-space – say the energy modes available to the photon – the information capacity is formally infinite. [A more detailed account of the difference between the information capacity and the messaging capacity of a physical system is presented in Appendix A.](#)

Generally, the capacity of the macroscopic physical systems for processing data, while bounded, is immense [9]. In particular, the number of independent messages that a physical system can communicate is limited to the number of its independent adjustable properties. A macroscopic object, like a token, has a large number of adjustable properties, such as mass, temperature, color, and width. In the microscopic realm, however, many such attributes cease to exist in a well-defined or independent manner.

¹ The word information in the term *information capacity* could be misleading; for example, a 5GB hard disk drive does not necessarily contain 5 gigabytes of information since it usually contains many sectors containing random bits of data, conveying no information. The term “*data storage capacity*” can be a better one for describing the concept. Here we followed the established term in the literature.

Furthermore, microscopic physical systems have a much smaller number of components compared to macroscopic systems. Therefore, in general, microscopic systems have a much smaller number of independently adjustable properties, and accordingly, a much lower messaging capacity. For example, attributes such as color or temperature are not well-defined for an elementary particle like an electron, and it has only a few independently adjustable properties: energy, position, and spin. Nevertheless, like in any other physical system, these attributes of the electrons can store and convey messages.

The case of photons in the double-slit experiment

Similar to other microscopic systems, photons possess only a few adjustable attributes, which makes their messaging capacity very limited. Individual photons have three independently adjustable properties that may be used to convey messages: direction, frequency, and polarization. These properties of the photons are all utilized in the 3D movie theatres using '*polarized 3D systems*' to convey the shape, color, and depth perception of the moving pictures. The cosmic microwave background (CMB) maps are also made by extracting these three pieces of information from the photons: direction (which way they were coming), frequency (energy), and polarization [10, 11]. In short, a free photon has the messaging capacity to carry only three independent pieces of information.

The photon's messaging capacity can be reduced further by putting constraints on its attributes. For example, in performing the double-slit experiment, the setup requires a coherent monochromatic light source. This means the incoming photons all should be highly directional –i.e., pointing to the same location– and all should have the same energies. This prerequisite fixes two out of the three adjustable attributes of the photons –the direction & frequency– and leaves the photons in the double-slit experiment with only one non-fixed attribute. Thus, the photons in the double-slit experiment have only a single messaging capacity.

This representation of the matter gives us a new perspective on the situation. The photons in the experiment are physical systems with just a single adjustable attribute that would be able to convey exclusively one message. This account explains why performing two independent measurements on them would not result in consistent outcomes and the reasoning behind the unique behavior of the photons in the double-slit experiment. In short, a system with a single messaging capacity cannot simultaneously carry two independent pieces of information. The case of the photons in the double-slit experiment is an example of such a situation, where the two independent measurements are either about extracting the which-way information (particle-like property) of the photon or its wave-like property. On such systems, once a measurement is performed, a second measurement associated with an independent proposition must necessarily result in an outcome with zero informational content.

The physical meaning of zero information

In physics, we gain information from a physical object with the act of measurement. By performing a measurement, we always get a reading. Even a 'zero' reading is still a result and obtained data. Therefore, acquiring 'zero information' cannot be assumed to be equivalent to obtaining no results, but it means that the obtained results represent zero information.

The information content of a message, composed of bits of data, can be evaluated mathematically. The expected information gain in a process involving possible outcomes of $X = \{x_1, x_2, x_3, \dots\}$, is mathematically defined as the change in the Shannon entropy

$$I = H_2(X) - H_1(X) \quad (3)$$

in which the Shannon entropy is defined as

$$H_t(X) = - \sum P_t(x_i) \log_2(P_t(x_i)) \quad (4)$$

where $P_t(x_i)$ are the probabilities of the outcomes [12].

We can easily verify the two cases in which the information gain is zero: the case that the obtained data is the same old reading, and the case of having a random reading as randomness is an expression of zero information. In the case of no new data, when one keeps getting the same result r , we have $P_t(r) = 1$ and hence $H_t(X) = 0$ and $I = 0$. In the case of randomness, the outcomes do not follow a deterministic pattern and the probability distribution of the outcomes remains the same regardless of the previous outcomes, that is, $P_1(x_i) = P_2(x_i)$. Accordingly, the Shannon entropy stays constant, $H_1(X) = H_2(X)$, and therefore the information gain, in this case, is also zero. The relationship between ‘true randomness’ and ‘no information’ has been known in information theory since the early stages of its development, epitomized as *perfect secrecy* in cryptography [13]. In layman’s terms, zero information means either no new result or receiving random results; basically, either no news or conflicting news.

This argument demonstrates mathematically why in the double-slit experiment, the result of a complementary measurement on the photons should always contain an element of randomness, and, hence, explains the basis of the complementary principle. The randomness in the results of the complementary measurement –associated with an independent proposition from the first– indicates that the measurement is performed on a system with zero informational content. Furthermore, this account entails that such randomness is ontological and not removable.

Here, with no need to postulate any concept outside classical physics, by recognizing the single messaging capacity of the photons in the double-slit experiment, we showed how information theory explains the complementary behavior of the photons. The case of the single variable physical systems, or single-messaging systems (hereafter, SM-systems), as we will explore, provides us with an unprecedented occasion in physics. For example, as demonstrated above, there is a fundamental indeterminism involved in performing independent measurements on such systems. This means even when there is no uncertainty in the state of the system, performing measurements in the SM-systems are generally associated with some intrinsic uncertainty. This characteristic indicates that the dynamics of the SM-systems should be fundamentally different from the familiar deterministic dynamics of classical physics.

The canonical role of Measurement in physics

The main goal of physics, as an applied science, is to be able to predict the outcomes of future events while explaining past phenomena. We need a theory to describe the future state of a system before the planned experiment is performed and realized. This feature is, in a sense, similar to the analysis we are interested in the financial market or politics. In such fields, one can provide explanations of past events. However, the descriptive aspect of a theory does not have much significance if it is not associated with the theory’s predictive power. The laws of physics are those formulations that explain natural phenomena in the past and the future. The primary interest of a physicist is to figure out the future state of a system under an experiment that is not yet performed; after performing the experiment the state of the system and the results are known, in reality, and principle.

One can also look at the situation either retroactively or prospectively: either we know the state of a system and want to know the state it will be in after performing a specified experiment, or we know the current state of a system and are interested to know its previous state before performing the experiment. In either case, we deal with a known state and are interested to understand the effect of the process already performed or to be performed. The system at hand, in any case, is in a defined state, as the result of some previous measurement (the state of an informationally isolated system is indefinite). Regardless of whether a measurement process changes the state of the system, we know that if we immediately repeat the same measurement on a system, we'll get the same result. Therefore, the state of the system can be defined according to the performed measurement on it. In physics, we deal with systems with definite states² and are interested to know about the state that the system would be in after a measurement. In other words, the state of a system is not an abstract a priori concept, as opposed to in philosophy or epistemology. We, therefore, need to base our discussions on pieces of evidence as they would appear in the outcomes of the realizable measurements.

Hence, the measurement is the cornerstone of physics. The information gain after a measurement, however, needs some scrutiny. Mathematically the information gain can be quantified according to equation (3) which is a measure of “surprise” after the process. In deterministic theories –including all the classical physics– however, the results of the measurements are predictable in principle (for example, where a projectile lands, given its initial velocity, gravity, etc.). Therefore, in the final analysis, performing a measurement does not involve producing a new piece of information, in the strict mathematical sense. Considering the act of measurement however, the SM-systems are in a different category. As discussed earlier, in performing independent measurements on the SM-system, the perfect secrecy rule affirms that the outcomes contain randomness and cannot be determined from the previous state of the system in principle. Therefore, in the SM-systems performing a measurement generally involves an element of surprise and unpredictability, and accordingly information gain (For instance, measuring the X component of an electron spin that is in Z+ state). That only in non-deterministic theories performing measurements involve information gain can sound counterintuitive; however, once we realize that information is a measure of surprise rather than collecting data, it becomes clear that only in theories that are not deterministic measurements involve producing new pieces of information.

Measurements and the SM-systems

In discussing measurements, the essential difference between the physics of SM-systems and classical physics becomes evident. Since an SM-single cannot have more than a definite property at a time, any measurement performed to measure an independent property of the system entails erasing the previous content of the system and putting the system in a new unpredictable state –corresponding to an outcome of the measurement being performed–. In general, performing measurements in SM-systems involves information gain and producing new pieces of information, and therefore, an unpredictable change to the state of the system and an element of surprise. This is in contrast to classical physics, which has a deterministic structure. In classical physics, measurement outcomes are essentially predictable and do not involve information gain. This makes classical physics unsuitable for explaining systems in which measurements intrinsically involve producing new pieces of information. Therefore, due to the fundamental impossibility of prediction in SM-systems, these systems require a new kind of physics; a

² Note that the knowledge of the state of a system does not mean the knowledge of the states of its subsystems; e.g., a system composed of two coupled particles with zero total momentum.

general theory that would relate the initial state of the SM-system to the possible measurement outcomes, but not in a deterministic manner. If constructed successfully, such a theory of “SM-physics” would essentially be a generalized probability theory. Consequently, it should be of no surprise that the theory’s predictive power would only be probabilistic, as opposed to the deterministic nature of classical physics.

Particularities of the SM-systems in measurements

The limited messaging capacity of the SM-systems gives them some characteristics radically different from that of classical objects. We already discussed the ontological randomness in performing independent measurements. Here we identify some other unique situations involved in performing measurements on the SM-systems.

Fundamental Uncertainty: unless performing the same measurement, the outcome of a new measurement on an SM-system is fundamentally uncertain. This is since an SM-system cannot hold more than one piece of information, so a measurement either should produce the previous result (which is the case in repeating the same measurement and no information gain) or erase that and produce a new result (an information gain) that in the latter case necessarily involves surprise and an unpredictable change of the state of the system.

Complete information erasure: an SM-system cannot hold more than one piece of information; hence once a new independent measurement is performed on the system and a new piece of information results, the previous information content of the system is necessarily lost forever. That means in SM-systems, the information content of the system can be erased, in principle, by performing a new independent measurement. This situation contrasts with the classical systems, in which information is conserved, say when you erase a file from a hard drive, the process can be reverted in principle by looking where the energies dissipate, photons escaped, etc., and reversing the involved physical processes.

Projection: performing an independent measurement on SM-systems entails an abrupt unpredictable change to the state of the system from the previous one to a new unpredictable one. This non-classical situation is the same as what is referred to as the projection postulate, the reduction of the wave packet, or the collapse of the state vector in QM. Note that we cannot expect a real information gain from a system without a real abrupt and unexpected change in the description of the system. Hence, it is reasonable that information gain disrupts the course of nature.

Notation: Here, we use the following notation to represent the measurements on the SM-system. Performing the measurements type S , \tilde{M}^S , on the SM-systems can result in any of the mutually exclusive outcomes \tilde{M}_i^S , $i \in \{1, \dots, N\}$. As we discussed, if a system is already in a state that corresponds to an outcome of the same type of measurement, \tilde{M}_j^S , performing the measurement does not change the state of the system and we’ll have

$$M^S \tilde{M}_j^S = \tilde{M}_j^S \quad (5)$$

Performing an independent P type measurement, \tilde{M}^P , however, will change the state of the system to an outcome of the new measurement

$$M^P \tilde{M}_n^S \rightarrow \tilde{M}_r^P \quad (6)$$

which is a random outcome among the possible outcomes of the new measurement. In short:

$$M^I \tilde{M}_n^I = \tilde{M}_n^I \quad (7)$$

$$M^{II} \tilde{M}_n^I \rightarrow \tilde{M}_r^{II}$$

On the “quantal” properties of the SM-systems

The SM-systems exhibit several other features, such as no-cloning, non-commutation, and non-contextuality, which are considered peculiarities of the quantum systems. Below, we propound some of the “quantal” properties of the SM-systems in an axiomatic manner.

Axiom 1 (*Existence of systems with a single adjustable property*). There exist physical systems with only one adjustable property. In practice, the other independent properties of such a system either are fixed by various screenings in the experimental setup (e.g., the photons in the double-slit experiment) or can be abstracted away since the specified property of the system can be considered isolated and in total seclusion from the rest (e.g., electron spin).

Theorem 1 (*Single-messaging systems*). A system with single adjustable property can hold no more than one message.

The proof is obvious by contradiction. Note that an SM-system can have a huge information capacity. In theory, one can devise methods to code more than one message on the sequence of the data carried by the system. In practice, this can be possible if the coding algorithm is shared with the other party. However, this includes separate communications and a wealth of shared background knowledge between the parties which none are in the single physical message transmitted. Thus, physically it cannot carry more than one message. (See Appendix A, for a discussion about our notion of information –which includes a lot of implied shared knowledge– versus physical information which bears no tag).

Corollary 1 (*Single context recording*). The SM-system at any time can contain only a single message, one piece of information relating to the last measurement performed on the system.

Corollary 2 (*No simultaneous encoding*). The SM-system cannot simultaneously contain more than one message.

Corollary 3 (*No Counter-factual reasoning*). Since at any time, SM-systems cannot contain more than one content, in discussing SM-systems one cannot reason about properties of the system that have not been measured and the results of different measurements that have not been performed. This feature contrasts with the case of classical systems that properties of objects can always be assumed to have values even when they have not been measured. Counter-factual reasoning thus cannot be applied in the SM-systems, as no more than one property can have a meaning and definiteness at any time. This means the level of reality that can be attributed simultaneously to physical quantities in SM-systems is very confined.

Corollary 4 (*Contextuality*). In discussing SM-systems, unless already performed, a measurement contains no pre-existing values. It follows since an SM-system cannot have more than a single content (corollary 1). Therefore, in SM-systems a question like “what is the X value of the system when its P value is p” is meaningless; neither has it admitted an operational definition, nor does it reflect an understanding of the SM-systems.

Corollary 5 (*No hidden content*). The content of the SM-system is fully determined as the result of the performed measurement. Note that the single content of the SM-system indicates there is no extra content to the SM-system besides what can be inferred from the outcome of the measurement. In other words, the description that the SM-system is in the state that measurement \hat{Q} would result in \check{q}_2 , fully characterizes the state of the system. Any other information that can be deduced from that description, usually from the setup of the experiment, also may characterize the state. For example, in the Stern–Gerlach experiment setup the outcome that “the particle is deflected upward” could mean that the spin of the particle is in \check{S}_z^+ state; therefore, one can also describe that state as “ \check{S}_z^+ ”.

Axiom 2 (*No null measurement*). An observation always yields a result. In performing a measurement, obtaining a zero reading is also a piece of data and an outcome.

Theorem 2 (*physical representation of zero information*). In the physical world, obtaining zero information means either getting the same old result or collecting random readings.

See the paragraph following equation (3) for proof.

Corollary 6 (*consistency in repeated measurements*). An immediate repeated measurement on a system yields the same outcome. Besides expecting accord in the physical world, we don’t expect information gain in the sequential measurements of the same property. Accordingly, after performing a measurement a state –which corresponds to the outcome of the measurement– can be attributed to the SM-system.

Theorem 3 (*ontological randomness*) In SM-systems performing measurements associated with independent propositions results in random readings.

From Axiom 1, SM-systems can contain only one message. Once a measurement performed on the system produces an outcome, there is no more informational content to the system (corollary 5). A later measurement associated with an independent proposition should yield an outcome (axiom 2). The zero informational content of the system means that the outcome either should be the same old one, or a random reading (theorem 2). The first case is rejected by corollary 4, thus, the outcome of the new measurement must contain randomness.

Corollary 7 (*No history*) The SM-system cannot contain its past. The only information an SM-system holds is its current state, as it can’t hold more than one piece of information. This means the result of the previous measurements on the SM-system gets erased right after performing an independent measurement. Accordingly, the future state of the SM-system following an independent measurement does not depend on its history (cf. ontological randomness).

Corollary 8 (*Collapse*) In SM-systems, performing an independent measurement involves erasing the previous content of the system and putting it into a new state. This entails the change of the previous state of the system to a state corresponding to an unpredictable outcome of the new measurement (theorem 3). This abrupt unpredictable change of the state due to performing measurement is commonly referred to as collapse in the QM literature.

Theorem 4 (*noncommutativity*) In SM-systems, the result of performing two different measurements depends on the order in which the measurements are performed.

This is a manifestation of the physical fact that the system cannot hold more than one message at a time and follows from theorem 3 and corollary 7. Given that performing independent measurements on the SM-system results in erasing the previous state and getting a random outcome, the final state of the

SM-system depends on the order that the measurements get performed (cf. corollary 4-contextuality). We, therefore, expect the underlying algebra of SM-systems to be non-commutative.

Symbolically we may represent the two situations as

$$\widehat{M}^{II} \widehat{M}^I \widetilde{M}_j^0 \rightarrow \widehat{M}^{II} \widetilde{M}_k^I \rightarrow \widetilde{M}_m^{II} \quad (8)$$

$$\widehat{M}^I \widehat{M}^{II} \widetilde{M}_j^0 \rightarrow \widehat{M}^I \widetilde{M}_b^{II} \rightarrow \widetilde{M}_w^I \quad (9)$$

in which the final outcomes are not the same.

Theorem 4 (*no-cloning*) An SM-system cannot be cloned in general.

The proof we present here is similar to the one provided by Dieks [14]. Start with an SM-system in a certain state. Assume we have a way to clone the SM-system. We make n copies of the state. When we repeat the original measurement on the copies, all consistently yield the same result as the original state. Next, on the copies, we perform an independent measurement. This measurement on the copies results in different random readings (theorem 2). Therefore, the final states of the copied ones are not all identical as required by the definition of the cloning apparatus. This indicates we cannot make copies of an SM-state in a consistent way. The envisioned cloning assumption therefore should be false.

Symbolically we may represent the cloning apparatus as

$$\widehat{U}(0_0 | \widetilde{M}_n^S) \mapsto (\widetilde{M}_n^S | \widetilde{M}_n^S) \quad (10)$$

where 0_0 is the “neutral” state of the cloning apparatus before the procedure. \widetilde{M}_j^S is the initial state of the system, which, in turn, is an outcome of the measurement type \widehat{S} , that is

$$\widehat{M}^S \widetilde{M}_j^S = \widetilde{M}_j^S \quad (11)$$

Suppose we perform the measurement \widehat{S} on the system

$$\widehat{M}^S \left(\widehat{U}(0_0 | \widetilde{M}_j^S) \right) \mapsto \widehat{M}^S (\widetilde{M}_j^S | \widetilde{M}_j^S) = (\widetilde{M}_j^S | \widetilde{M}_j^S) \quad (12)$$

The effect should be –and is– the same if we had performed the measurement initially on the state to be cloned:

$$\widehat{U} \left(\widehat{M}^S(0_0 | \widetilde{M}_j^S) \right) = \widehat{U}(0_0 | \widehat{M}^S \widetilde{M}_j^S) = \widehat{U}(0_0 | \widetilde{M}_j^S) \mapsto \widetilde{M}_j^S | \widetilde{M}_j^S \quad (13)$$

On the other hand if we perform an independent measurement \widehat{Q} on the system

$$\widehat{M}^Q \left(\widehat{U}(0_0 | \widetilde{M}_j^S) \right) \mapsto \widehat{M}^Q (\widetilde{M}_j^S | \widetilde{M}_j^S) \rightarrow \widetilde{M}_t^Q | \widetilde{M}_t^Q \quad (14)$$

the outcome is not consistent with the final result acquired by the effect of the cloning apparatus:

$$\widehat{U} \left(\widehat{M}^Q(0_0 | \widetilde{M}_j^S) \right) = \widehat{U}(0_0 | \widehat{M}^Q \widetilde{M}_j^S) \rightarrow \widehat{U}(0_0 | \widetilde{M}_r^Q) \mapsto \widetilde{M}_r^Q | \widetilde{M}_r^Q \quad (15)$$

Since the outcomes are not the same, the cloning fails.

As demonstrated, the SM-systems comprise many unique features of quantum physics, with intuitively understandable descriptions. These shared characteristics strongly suggest that the SM-systems are quantum systems. Next, we construct a mathematical theory to describe the dynamics of the SM-systems. Whether the SM-systems are indeed QM systems should be decided once we have the mathematical theory of the SM-systems and can compare the algebraic structures of the two theories.

Stating the framework and constructing the theory of SM-systems

We discussed the fundamental role of measurement in physics and how the unique standing of the SM-systems about conducting measurements distinguishes these systems from classical ones; the unprecedented situation that performing an independent measurement indeed changes the state of the SM-system to a new unpredictable outcome of the measurement. Measurements are usually performed in the

experiments and sometimes are called observations. We don't use the latter term as it implies the existence of an 'observer' or an 'operator'. In our description, measurements can be performed using measuring apparatuses. (Note that by referring to measurement, we in no way imply that it is being performed in the experiment, and certainly we don't refer to a measuring apparatus itself.)

A measuring apparatus is a physical system that consists of at least three parts: a fixed state (the zero of the reading parts, the "zero state"), a pointer that its state can change due to interacting with the system being measured, and finally a register part (the memory) that saves the reading of the pointer compared to the fixed state (the outcome). Any system with such a tripartite structure can be considered as a measuring apparatus. A molecule interacting with another, for example, cannot be considered as a measuring apparatus unless it has enough degrees of freedom to register the state of the other system independently. A measuring apparatus, in the sense used here, however, does not need to be, say complex, or being operated by a conscious agent.

With this description, measuring devices are not SM-systems. If we try to use an SM-system as a measuring apparatus (ignoring the tripartite structure for now) the trouble is evident: upon coupling, the correlation formed between the two systems, i.e., the 'SM-measuring device' and the 'measured SM-system', entails that the combined system (measuring device + SM-system) is just another SM-system, as it has no more than a single independent variable.

Schematically one can represent the tripart measuring apparatus in its expanded form as

$$|0\rangle_{\text{pointer}}|\widehat{\text{register}}\rangle \quad (16)$$

however, for all practical purposes, the only variable that one needs to report to convey the measurement outcome is the registered pointer reading

$$0|0\rangle_{\check{0}} \rightarrow 0|n\rangle_{\check{n}} \equiv \check{n} \quad (17)$$

since in performing measurements on SM-systems, the reading of the apparatus fully specifies the state of the SM-system after the measurement (cf. Corollary 5-No hidden content).

In building our framework, we start with an SM-system in a certain state. Our goal is to construct a theory to explain the possible behavior of the SM-system under any of the realizable measurements, a theory that would be probabilistic. Here we use the propensity interpretation of probability in discussing the dynamics of the SM-systems. This interpretation thinks of probability as being resulted from an existing propensity, disposition, or tendency of a physical system in a given situation to result in a specified outcome. Propensities are not relative frequencies but can be adjudged as the causal basis for the observed long-run frequencies. For example, consider a particular biased coin with the propensity of 0.42 to land heads every time it is tossed. The frequentist approach to probability cannot assign a chance for single tosses of the coin since relative frequencies do not exist for individual observations, but only for large ensembles. On the flip side, propensities can be used along with the law of large numbers to explain the long-run frequencies. Apart from this interpretational distinction, we use the terms probability and propensity in the same general sense.

We argued that SM-physics is not a deterministic theory. We envisage the SM-physics to, given the present state of an SM-system, provide us with the machinery for computing the probabilities of the system for the outputs of different possible measurements. Our aim is thus to specify the system's propensities for the outcomes of realizable measurements. By realizing that: 1) the state of the SM-system

is an outcome of an already performed measurement (Corollary 4) and 2) SM-systems have single-content (Corollary 5), the problem reduces to specifying the propensities of the outcome of the first measurement for those of the next measurement. Therefore, we have a clear roadmap to follow: we need to find out how the relationship between the two measurements determines the propensities of the SM-system. Our goal is to develop a mathematical theory to associate the relationship between two measurements with the propensities of an outcome of one measurement for the outcomes of the other measurement.

Construction of formalism

In discussing the possible outcomes of an SM-system for a specified measurement, as mentioned above, we can consider the relationship between two measurements: the one that has defined the state of the system, and the one that we are interested in to calculate the propensities of the system for its different outcomes. Two measurements could be either dependent or independent. For two independent measurements, the outcome of the second one is independent of the outcomes of the first one, while for the dependent measurements, certain outcomes of the second one can be more probable or less probable depending on the outcome of the previous measurement. An example of the dependent measurement can be measuring the spin of an electron already in \check{S}_+^z state, in the direction that is tilted 20° from the z -axis in the zx -plane.

Without losing generality, let's assume measurements with N distinguishable outcomes on an SM-system (generalization to the infinite case should not be difficult.). We represent a measurement type K symbolically as \widehat{M}^K , which its possible outcomes are the independent members of the set $S(\widehat{M}^K)$ defined as:

$$S(\widehat{M}^K) = \{\check{M}_1^K, \dots, \check{M}_N^K\} \quad (18)$$

Let us write P_j^I to indicate the propensity of the SM-system to end up in the final state \check{m}_j^I due to measurement \widehat{M}^I . Since the measurement eventually yields a result, we have

$$\sum_{j=0}^N P_j^I = 1 \quad (19)$$

In a few cases P_j^I can be evaluated readily. For an already performed measurement \widehat{Q} on the system with the outcome \check{M}_i^Q , we know that $M^Q \check{M}_i^Q = \check{M}_i^Q$; therefore, for the case of the second measurement being \widehat{M}^Q we have

$$P_j^Q = \delta_{j,i} \quad (20)$$

At the same time, the outcomes of an independent realizable measurement \widehat{S} are all equally possible, and therefore the system has an equal propensity for those

$$P_j^S = \frac{1}{N} \quad (21)$$

with N being the total number of possible outcomes of the measurement.

Besides these two specific cases, we need to construct an extensible framework to evaluate the propensities of the system for the outcomes of a general type of the second measurement, i.e., the case that the measurement is neither the same as the previous one nor is fully independent of that. Consider

two dependent measurements \widehat{M}^K & \widehat{M}^L with the inter-probability between their possible outcomes, written as

$$T_{j,i}^{L,K} = P(\widetilde{M}_j^L | \widetilde{M}_i^K) \quad (22)$$

(the probability of obtaining the j^{th} outcome in measurement \widehat{L} , given the i^{th} outcome of measurement \widehat{K}). The inter-probability of the two measurements can be framed in an $N \times N$ ‘interrelation matrix’. One can recognize a number of properties for this matrix. For example, for any fixed i we should have

$$\sum_j T_{j,i}^{L,K} = 1 \quad (23)$$

since the \widehat{M}^L measurement is presumed to produce one and only one result. Now, representing the propensity of the SM-system for the measurement \widehat{M}^K as P_i^K , it is straightforward to calculate the propensity of the system for the measurement \widehat{M}^L :

$$P_n^L = \sum_i P(\widetilde{M}_n^L | \widetilde{M}_i^K) P_i^K = \sum_i T_{n,i}^{L,K} P_i^K \quad (24)$$

which is summing all the possible ways the system can end up in \widetilde{M}_n^L . We expect these matrices to conserve the total probability, i.e., we should have $\sum_n P_n^L = 1$. This is guaranteed since

$$\sum_n P_n^L = \sum_n \sum_i T_{n,i}^{L,K} P_i^K = \sum_i P_i^K \sum_n T_{n,i}^{L,K} = 1 \quad (25)$$

given Eqs. (19) & (23).

Now assume we start with \widehat{M}^L as our first measurement, which the system has the propensity of P_i^L . In the same way, for the dependent measurement \widehat{M}^K we can write the inter-probability matrix of

$$S_{j,i}^{K,L} = P(\widetilde{M}_j^K | \widetilde{M}_i^L). \quad (26)$$

In practice, these interrelation matrices indicate how the probabilities of getting a specific outcome in a first measurement transfer to that of having a specific outcome in the second measurement. In other words, these matrices are mappings from the probability-space of the first measurement to that of the second.

These mapping matrices should possess certain properties. Their components are probabilities which should be positive values less than or equal to one:

$$0 \leq T_{j,i} \leq 1 \quad (27)$$

$$0 \leq S_{j,i} \leq 1 \quad (28)$$

We also have

$$\sum_j T_{j,i} = 1 \quad (29)$$

$$\sum_j S_{j,i} = 1 \quad (30)$$

for any fixed i , as the measurements will result in one and only one outcome (cf. Eq. (23)). Furthermore, these matrices map the SM-propensities from one measurement-space to another.

$$P_j^L = \sum_i T_{j,i} P_i^K \quad (31)$$

$$P_j^K = \sum_i S_{j,i} P_i^L \quad (32)$$

which means the consecutive action of these reciprocal maps should take an initial state back to itself. Therefore, we should have

$$TS = ST = \mathbf{I} \quad (33)$$

in which \mathbf{I} is the identity matrix. However, the current construct of the interrelation matrices does not allow such an identity. The problem arises since the components of the matrices, i.e., $P(\tilde{M}_j^K | \tilde{M}_i^L)$ are positive value probabilities, which would not add up to zero for the non-diagonal components of the multiplied TS and ST matrices. And the sole possible instance of $S = T = \mathbf{I}$ refers to the trivial case of identical measurements.

To overcome this issue, we can try using probability-amplitudes, ρ , instead of the probabilities, P , defined according to

$$P(\tilde{M}_j^K | \tilde{M}_i^L) = |\rho_{j,i}^{K,L}|^2 \quad (34)$$

The new inter-measurement matrices can be then constructed accordingly as

$$\Gamma_{j,i}^{L,K} = \rho_{j,i}^{L,K} \quad (35)$$

and

$$\Delta_{j,i}^{K,L} = \rho_{j,i}^{K,L} \quad (36)$$

For these interrelation matrices we now have

$$\sum_j |\Gamma_{j,i}|^2 = 1 \quad (37)$$

$$\sum_j |\Delta_{j,i}|^2 = 1 \quad (38)$$

with the same logic that a measurement eventually yields a result and the probabilities should add up to 1 (cf. Eqs. (29) & (30)). Moreover, similar to (27) & (28) the components are confined according to

$$0 \leq |\Gamma_{j,i}|^2 \leq 1 \quad (39)$$

$$0 \leq |\Delta_{j,i}|^2 \leq 1 \quad (40)$$

which mean in this construct the components can be negative and complex too.

We can employ these interrelation matrices to pursue the SM-system's propensities analogous to the previous construct, but with an adjustment. Instead of acting on the propensities, P_j^I , of the SM-system, these mappings should act on the propensity-amplitudes of the SM-system to keep the calculations consistent. Accordingly, we define the propensity-amplitudes of the SM-system, σ_j^I , in terms of its propensities, P_j^I , as

$$P_j^I = |\sigma_j^I|^2 \quad (41)$$

in which σ_j^I indicates the propensity-amplitude of the SM-system to end up in the outcome \tilde{M}_j^I for measurement \hat{M}^I . For the propensity-amplitude of the system we have

$$\sum_j |\sigma_j^I|^2 = \sum_j P_j^I = 1 \quad (42)$$

To evaluate the propensities of the system for a second measurement we proceed as before. The interrelation matrices $\Gamma^{L,K}$ and $\Delta^{K,L}$ are mappings between the possible outcomes of the two measurements \hat{L} and \hat{K} , but for the propensity-amplitudes

$$\sigma_j^L = \sum_i \Gamma_{j,i} \sigma_i^K \quad (43)$$

$$\sigma_j^K = \sum_i \Delta_{j,i} \sigma_i^L \quad (44)$$

Similar to the previous construct, these mappings should conserve the total probability, i.e.,

$$\sum_j |\sigma_j^L|^2 = \sum_j |\sigma_j^K|^2 = 1 \quad (45)$$

This means Γ and Δ matrices have to be unitary. (We reject the antiunitary case since these matrices describe continuous transformations between the propensity-amplitudes. see also Eqs. (37) & (38)).

The second property of these interrelation matrices comes from the fact that they are mappings between the two measurement spaces in opposite directions, and therefore their consecutive actions should map any state back to itself (cf. Eq. (33)). Thus, we should have

$$\Gamma\Delta = \Delta\Gamma = I \quad (46)$$

This immediately results that these unitary matrices are conjugate transpose of each other

$$\begin{aligned} \Delta &= \Gamma^{-1} = \Gamma^* \\ \Gamma &= \Delta^{-1} = \Delta^* \end{aligned} \quad (47)$$

That means the interrelation matrices between the two measurements are conjugate transpose of each other. In other words, in the SM-systems the propensity-amplitudes between any pair of measurements I and II outcomes are related as

$$\rho(\tilde{M}_b^{II} | \tilde{M}_a^I) = \rho^*(\tilde{M}_a^I | \tilde{M}_b^{II}). \quad (48)$$

Here, we derived the specificities of the SM-systems' state-spaces by analyzing the constraints on the transformation of the propensities of the SM-systems between measurements. Interestingly, the principal constraint that we observed was the conservation of the total probability in the mappings, and from the consistency requirements, the properties of the transformations followed.

State vectors, operator algebras, and Hilbert spaces

It is now straightforward to recognize the Hilbert-space formalism of the SM-systems theory and to identify the denotations of the elements of our SM-theory. From our construct, the rescaled probabilities of the SM-system, i.e., the probability-amplitude, σ^I can be used to describe the state of the system as a vector of length one in the N -dimensional space defined by the N independent outcomes of the measurement I (Eq. (18)). Using the common bra-ket notation and Eq. (42), for the SM-system's state-vector we have

$$\langle \sigma^I | \sigma^I \rangle = 1. \quad (49)$$

With this representation of the SM-state as vectors in the complex vector space of propensity-amplitudes, the algebraic structure of the complex vector space of SM-systems is evident. In this representation, once we fix a basis to present the SM-state in, $|\sigma^I\rangle$, the system's propensity-amplitude for any other possible measurement can be described according to Eq. (43) as a linear combination of these bases

$$|\sigma^{II}\rangle = \rho^{II,I}|\sigma^I\rangle \quad (50)$$

where $\rho^{II,I}$ are the unitary transformations between the measurements, portrayed in Eq. (48). In our construct, the transformations of the state of the SM-system between different measurements are carried out by the interrelation matrices of the measurements. The logic is simple: an SM-state can be thought to be the result of a certain measurement, so the dependency of two SM-states can be judged by the dependence of the two measurements that would produce those two SM-states.

One can easily realize that the Born rule for probability is already embedded in our construct. To calculate the propensity of the SM-system for a specified measurement outcome, \tilde{M}_b^{II} , given its initial state, $\sigma^I = \tilde{M}_a^I$ one can use Eq. (50) to get

$$\rho_{b,a}^{II,I} = \langle \sigma_a^I | \sigma_b^{II} \rangle = \langle \tilde{M}_a^I | \tilde{M}_b^{II} \rangle \quad (51)$$

Therefore, using Eq. (34), the propensity of obtaining a specified outcome in the measurement given an initial state, is

$$P_{b,a}^{II,I} = |\rho_{b,a}^{II,I}|^2 = |\langle \sigma_a^I | \sigma_b^{II} \rangle|^2 = |\langle \tilde{M}_a^I | \tilde{M}_b^{II} \rangle|^2 \quad (52)$$

This is the Born probability rule, resulted as an inherent consequence of the theory.

In sum, in this framework, the “state-vectors” of Hilbert space are complex-valued conditional propensity-amplitudes of the SM-system for different realizable measurements. The vector algebra of the SM-systems' Hilbert space accordingly describes the transformations of these propensity-amplitudes of the system between the realizable measurements.

Superposition of possibilities

In our construct, one can write the propagation of the propensity-amplitude of the system under a series of measurements by consecutively applying the interdependency matrices of the measurements³

$$\begin{aligned} |\sigma^{III}\rangle &= \rho^{III,II}|\sigma^{II}\rangle = \rho^{III,II}\rho^{II,I}|\sigma^I\rangle \\ &= \rho^{III,I}|\sigma^I\rangle \end{aligned} \quad (53)$$

in which

$$\rho_{f,i}^{III,I} = \sum_n \rho_{f,n}^{III,II} \rho_{n,i}^{II,I} \quad (54)$$

describes the interrelation between the measurements.

We can use this chain rule to determine the system's propensity-amplitude for a third measurement from the previous measurements' interrelation matrices. Importantly, this sum accounts for the “interference effects”. The joint probability of the final and initial state, calculated according to

³ Note that by the term ‘measurement’ we don't imply ‘observation’ or ‘collapse’; we merely peruse the transformation of the system's propensity-amplitude between different types of measurements.

$$P(\tilde{M}_f^{III}|\tilde{M}_i^I) = P_{f,i}^{III,I} = |\rho_{f,i}^{III,I}|^2 = \left| \sum_n \rho_{f,n}^{III,II} \rho_{n,i}^{II,I} \right|^2 \quad (55)$$

can include extra terms (the interference terms) that make the result different from the classical method of calculating joint probabilities. For example, 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):

$$\begin{aligned} P(\tilde{M}_S^{Screen}|\tilde{M}_O^{Source}) &= |\rho_{S,O}^{Screen,Source}|^2 \\ &= |\rho_{S,slit1}^{Screen,Slit} \rho_{slit1,O}^{Slit,Source} + \rho_{S,slit2}^{Screen,Slit} \rho_{slit2,O}^{Slit,Source}|^2 \end{aligned} \quad (56)$$

which the result is clearly different from the classical expectation of

$$\begin{aligned} P(\tilde{M}_S^{Screen}|\tilde{M}_O^{Source}) &= P(S|slit1)P(slit1|O) + P(S|slit2)P(slit2|O) \\ &\equiv |\rho_{S,slit1}^{Screen,Slit} \rho_{slit1,O}^{Slit,Source}|^2 + |\rho_{S,slit2}^{Screen,Slit} \rho_{slit2,O}^{Slit,Source}|^2 \end{aligned} \quad (57)$$

The chain rule for the propagation of the propensity-amplitudes between the measurements, Eq. (53), is a rather fundamental feature of the SM-systems. One can see here that it is the connection between the propensity-amplitudes of outcomes of different measurements –instead of their propensities–, that modifies the result from that of classical probability. The structure of our derivation permits a clear statement about the basis of interference and demonstrates the essential role of the presence of the intermediate measurement. In SM-physics, where reality emerges as the outcome of the measurements, the interference –which is resulted from the superposition of the possible outcomes of the intermediate measurement– indicates that the presence of the intermediate non-performed measurements (the “interaction-free” measurements) cannot be neglected in the SM-systems. Furthermore, we can discern that interference results as a consequence of mathematical bookkeeping of the SM-systems’ propensity-amplitudes for the possible measurements, not the physical reality of those.

Time evolution and derivation of the Schrödinger equation

So far, analysis of the mappings between different measurements’ propensity spaces for the SM-system allowed us to obtain the algebraic structure of the operators in their state-spaces, i.e., the Hilbert space and the Born probability rule. The inclusion of time evolution in this framework is straightforward, as discussed in many mathematical physics textbooks (see for example [15] Sec 3.3).

A measurement can be labeled by the time variable t , indicating the time at which it is performed. Taking the relationship between two measurements does not depend on time, their interrelation matrices must have the same structure at different times; therefore, there should be a unitary transformation $\mathbb{T}(t_2 - t_1)$ such that

$$\rho(t_2) = \mathbb{T}^{-1}(t_2 - t_1) \rho(t_1) \mathbb{T}(t_2 - t_1) \quad (58)$$

With common assumption about time evolution, one can write

$$\mathbb{T}(t_2 - t_1) = e^{-iH(t_2-t_1)} \quad (59)$$

in which H is a self-adjoint matrix that can be used as a definition of the Hamiltonian.

Borrowing the conventional QM terminology, so far, our discussion has been in the “Heisenberg picture”, in which the state of an isolated system remains fixed, whereas the matrices which represent the observables vary in time. Transposition into the “Schrödinger picture” of operators renders the evolution of the state, and the Schrödinger equation follows:

$$i \frac{d}{dt} \sigma^I(t) = H \sigma^I(t) \quad (60)$$

We found the operator algebra of Hilbert space can express the relations between the outcomes of measurement in SM-systems, by mapping the propensity-amplitudes of various realizable measurements to each other. Note that these universal relations contain no randomness. We can now properly identify the elements of the theory. The state of the SM-system, σ^I , codes the information from the past measurement, and in the propensities of the system for future realizable measurements. The SM-states is a representation of conditional probability-amplitudes that can be specified for different measurements, using the interrelations between the measurements. The interrelation matrices, $\rho^{II,I}$, are unitary matrices transforming the SM-state, σ^I , between the measurements. The construct of the theory guarantees that the total propensity is conserved in transforming from one possible measurement to another. Finally, when a measurement is performed, the state of the system is adjusted as the result of the observation.

In sum, by analyzing the constraints that the probabilistic nature of the SM-systems under measurements imposes, we found the general properties of the mappings that transform the state of an SM-system from one measurement-state-space to another. The algebra of the theory follows from the transformations that describe the deterministic relationships between measurements. Significantly, the standard formalism of QM, as well as the Born probability rule, was entirely obtained from our construction.

Equivalence of SM- and QM-systems

In discussing the photons' behavior in the double-slit experiment, we recognized the case of SM-systems, which explained the complementary dual behavior of the photons. We further realized that the SM-systems comprise features such as noncommutativity and no-cloning, which are uniquely peculiarities of the quantum systems. We discussed that classical physics cannot explain the SM-systems, and a distinct theory is needed to describe these systems. Finally, we derived the mathematical theory of SM-systems, which is equal to the standard formalism of QM. This consideration leads the SM-systems and QM-systems to be the same. This judgment needs a bit of discussion as we have had no clear demarcation criteria to discern quantum systems from the other physical systems.

The quantum world has been thought to be related to the microscopic world, with QM describing the physics of atomic and subatomic particles. On the other hand, the SM-systems are not intrinsically bound to any specific physical size. Nevertheless, since macroscopic systems usually have a large number of independent properties and, accordingly, a large capacity to simultaneously hold different pieces of information, we don't expect to find SM-systems among them. Moreover, since the messaging capacity of the systems scales with their size, it is understandable that in the microscopic world, where the systems have numerable properties that in certain situations could get confined to just one, we would expect predominantly more cases of SM-systems. The same argument is valid for the system's behavior in the cryogenic temperatures, another area where quantum behavior such as in superfluidity [16, 17], superconductivity [18], quantum Hall effect [19], and laser-cooled trapped ions [20, 21] emerges.

Our description provides a clear benchmark for the systems that would not obey classical physics: a physical system exhibits quantum behavior when it possesses only one adjustable property. This criterion has an easily testable corollary: it is not the case that the photons –or in general other microscopic

systems— should always behave quantum mechanically. Corpuscular theories of classical physics indeed explain many aspects of electronics and optics. Even most aspects of Magnetic Resonance (MR) physics, which is based on the spin—a quantum concept—are perfectly understandable from a classical perspective [22]. Therefore, it is not the size that defines quantum behavior, and accordingly, identifying QM as the physics of the microscopic world is an oversight. On the other hand, there is no physical constraint, in principle, to prevent having macroscopic systems with just a single adjustable variable. Macroscopic quantum phenomena, as in Bose-Einstein condensate or as in superconducting quantum interference device (SQUID) [23], indicate that quantum systems are not size-dependent, but it is the number of independently adjustable properties of a system that specifies its behavior.

In interference experiments with single large molecules [24, 25], a similar type of argument can be made for each of the macroscopic molecules in the meticulously prepared homogeneous beam of the macromolecule selected for performing the experiment. Observing “quantumness” in such coherence experiments in effect is not feasible without a dynamical screening of the molecules in the preparation phase that puts a portion of the molecules in a coherent monochromatic beam [26], informationally isolated enough to exhibit signs of interference. In these experiments, the presence of interference is inferred by observing whether the amplitude of the collected signal shows sinusoidal modulation that exceeds the experimental uncertainty, i.e., whether there is some interference effect on top of the classical expectation. In practice, in the molecular experiments, unlike the optical experiments that use highly coherent monochromatic beams, the signal minimum does not reach zero, which is due to the molecules in the beam that behave classically.

Following the derivation of standard QM formalism based on properties of the SM-systems, we, therefore, have a clearer picture of the world. There are two main types of physical objects in the world: the SM-systems which QM describes their behavior, and the rest which are classical systems. Accordingly, a comprehensible account of “the physical systems with a single adjustable property” fully defines the quantum system. In what follows, therefore, we use the terms SM-systems and QM-systems equivalently.

Interpretation of some controversial concepts

In the light of our derivation, we can elucidate a number of the controversial issues in quantum literature and, as well, provide clear interpretations for some concepts like “state” and “measurement”.

Irreducible randomness vs. hidden variable: We argued that the physics of the SM-system, unlike classical physics, cannot be deterministic in principle. At first, this could look counterintuitive since in the SM-systems the exact property of the system can be determined entirely by a single observation. However, as we demonstrated, this single contextuality of the SM-systems leads also to the inevitable uncertainty in the outcome of a later independent measurement. Our presentation indicates that the physics of the SM-systems is intrinsically indeterministic, and no deeper reality is hidden beneath the unpredictable behavior of the SM-systems. Our account is illustrative of von Neumann’s irreducible randomness [27] for the observed quantum randomness.

Physical state and the state-function: In classical physics, the state of physical systems is the list of their physical configurations with their values. The SM-systems are then in a peculiar situation, as they only possess a single physical property which limits the level of reality attributed to them. This means the state of an SM-system has a single entry, portraying the result of the last measurement on the system.

Also, in classical physics, the state of a system can be used to determine its future behavior, a task that is fundamentally impossible for the SM-systems. In the SM-systems, a theory can only probabilistically describe the outcomes of measurement, i.e., provide their propensities. However, as we saw, in SM-physics, one may use the state of the SM-system to list the propensities of the system for any realizable measurements using a mathematical construct, referred to as the state-function of the system. This wealth of information may look contradictory at first, given that the state of the SM-system has a single entry and the single context recording of the SM-systems. How can an SM-state contain all information for propensities of the system for any realizable measurement?

From our derivation, we can see that such information is not buried in the system's state-vector, but in fact, it is embedded in the mappings which directly reflect the relations between the measurements⁴. The state is a mathematical description of the physical state of an SM-system, which is an outcome of the last performed measurement on the system. It can be used to describe the probabilities of the outcomes of future realizable experiments. The SM-theory provides the machinery to calculate the propensities of the system for different measurements. As it is clear from our derivation, the transformations which map the state-function between different measurements are the interrelation matrices of the measurements. We use the dependency of two measurements (one that the state of the SM-system is in, and the other one to be realized) to calculate the propensity of the system for the new measurement. This resolves the misconceptions about the state of an SM-system, that it contains all the information about its representation in various measurements (for example, this statement from WIKIPEDIA 'a quantum state "carry" a larger amount of (classical) information – thanks to quantum superposition'). That information exists in the relationship between the two measurements, not the SM-system. The "extra information" is the interdependency of the possible measurements, not the information contained in the SM-system⁵. One should note that the decomposition of an SM-state in the state-space of another measurement, say, $\check{S}_+^X = \frac{1}{\sqrt{2}} (\check{S}_+^Z + \check{S}_-^Z)$, is a mathematical description of the dependency of the two measurements' outcomes, not the actual portrayal of the SM-system in an unperformed measurement. The mathematical expansions of the SM-states should not be interpreted physically as if the SM-systems have more than one objective reality –which is, by the way, its pure state description–. One should always keep in mind the limited realism of the SM-system and its single context recording.

Unlike in many non-individualist views of the quantum state (e.g., [28, 29]), we advocate that a state is assigned to individual SM-systems, not just to the ensembles of similarly prepared SM-systems. The state-function, in its broad sense, represents the knowledge of possibilities and their probabilities, rather than the state of reality. We can perceive the state-function as a list of propensities, which while being an objective property of the system, has no reality. Note that not all objective entities are necessarily real; borrowing the terms used in philosophy (for example see [30]), we can discern *epistemic realities*, such as money, the banking systems, and other social constructs, versus *ontological objective realities*, like a paper-money(bill) and a bank building. In this account, the SM-state is epistemically objective. The idea of intrinsic quantum probabilities of individual systems (calculated from a state vector

⁴ Maybe an example can make the difference between the information buried in a statement, and that which is embedded in the background clear. From the statement "Janice is 6-years old" one may infer that, "Janice is not married" or "Janice cannot drive a car", but none such information is contained in the original statement, rather contained in the background information about human societies.

⁵ In a similar sense, a bill does not carry the information about its purchase power for different market items. That information is contained in the market dynamics, not in the paper money.

or density matrix) has been around in QM literature [31-33]. The notion of objective quantum propensities, not only can explain the statistical interpretation of quantum theory [31], but also enables us to make sense of single-case probability attributions in quantum events, say the probability of decay of a decaying atom at a particular time.

Schrödinger equation and the measurement problem: In our construct, we analyzed the transformation of the SM-system's propensities between the measurements. From our derivation of the Schrödinger equation, which is a deterministic, linear equation, it is clear that the equation describes the dynamics of the propensity-amplitude state of the SM-systems between the measurements. Two points about this equation should be elaborated: First, the obtained formulation is concerned with the SM-systems and not all physical systems. Second, the equation describes how the state of the SM-systems transforms between different types of realizable measurement, and, accordingly, how the propensities of the system can be evaluated for those measurements. This formalism is complete, linear, and deterministic in describing the evolution of the propensity state until a measurement is performed, and as the result the state of the system changes.

In our construct, we clearly defined and distinguished the SM-systems from classical ones. Therefore, there is no confusion about the divide between quantum and classical, as well as many other situations that arise because of no clear definition of that line. For example, in our framework with the provided description of the measurement apparatus as a tripartite classical system (Eq. (16)), there are no problems parallel to the "measurement problem" discussed in quantum literature (insisting on describing measurements by QM-physics), as well as Schrödinger's cat [34, 35], Wigner's friend [36], the observer and observed, and many other situations that essentially are the extension of the measurement problem. With this clear distinction, when it comes to applying the Schrödinger equation to physical situations, we should regard the definition and applicability of the theory, which is to the systems that have an extremely limited messaging capacity, one to be exact. This denotes that while SM-theory is a fundamental theory, it is not a universal physical theory.

When considering the role of measurement in physics, there is a fundamental difference between classical physics (which is a deterministic theory) and SM-physics. As discussed, (e.g., in Corollary 8 (*Collapse*)), in SM-physics, performing a measurement generally involves an unpredictable outcome and an element of surprise. Therefore, the linearity of SM formalism cannot extend to after performing a measurement and the act of observation. The reason is clear: the Schrödinger equation fully determines the evolution of the SM-state, and with deterministic equations, there is no surprise and information gain. The described dynamics of the SM-system mirror the fact that in these systems the physical reality is shaped by the choice of the performed measurement, and the information gain always involves surprise and unpredictability.

Our construct reveals there is no need to reconcile the reduction of the SM-system's state with the Schrödinger linear evolution. The question of specifying exactly how, when and why collapse (as opposed to unitary dynamics) occurs in observations ("the analysis of the measurement process") is predicated on a false assumption: that the Schrodinger unitary evolution is still valid in performing a measurement, as the SM-system interacts with a system which is not an SM-system; and that the mathematical decompositions of SM-states denote the factual presentation of these single-content systems.

In sum, the measurement problem loses its problematic aspect by noting: 1) the measuring apparatus is not an SM-system, 2) an SM-system has no content rather than the outcome of the performed measurement; the mathematical expansion of an SM-system's state in the non-measured state-spaces has no physical meaning and is not an element of reality, and 3) for an SM-system, observation generally involves information gain from the system and thus comprises an element of surprise and sudden change (collapse).

Some past approaches to the problem of deriving QM

In contrast to most physical theories, the quantum theory had not been grounded yet on fundamental physical principles. Since its beginning, an underlying axiomatic mathematical formalism of the theory (based on Hermitian operators and their corresponding eigenvectors and eigenvalues, mainly developed by Dirac and von Neuman [27, 37]) has been known; however, a conceptual physical foundation in terms of simple axioms and meaningful principles on which the entire theory could be built upon had been missing. There have been many attempts to find physical principles behind quantum theory [38-54]. In general, however, they either fail in deriving the complete formalism of the theory or are based on abstract mathematical assumptions with no clear physical basis. Our elucidation of the single-photon double-slit experiment situation resolved the challenge of what should be designated as an axiom to derive QM from. By recognizing the SM-systems, and realizing their “quantal” characteristics, we developed a theory of conditional propensity-amplitudes for those systems, which presents the complete standard formulation of quantum mechanics.

Most of the past works to derive, or to reconstruct QM from some principles, in general, have been, in the final analysis, mathematical endeavors to analyze the internal structure of the theory and its mechanism in the hope to find the cornerstones of the theory. In simple words, by analyzing how the different characteristics of the quantum systems relate to each other, there were studies aiming to provide new insights into the characteristics that structure QM, and accordingly, to shine a light on its foundations. For example, Clifton, Bub, and Halvorson [52] used an approach in C*-algebraic formalism to demonstrate how several information-theoretic constraints on physical systems (to be specific: no Superluminal information transfer, no perfect broadcasting, no bit commitment) suffice to deduce some of the features of QM (to be specific: kinematic independence, noncommutativity, nonlocality). Another approach by Hardy focuses on general properties of probability theories and discusses the reasons which distinguish quantum theory from classical probability theories [51].

There have been some conceptually similar works to the current work, aiming to understand the physical origin of the mathematical structure of the theory based on the informational properties of the quantum systems (most noteworthy [47, 55]). These studies commonly base their axioms on the finite information capacity of an “elementary system”. While such works present some interesting conceptual results, one should note that postulating statements like ‘the information capacity of a quantum system is finite’ as the antecedent already affirms the consequent that ‘a quantum system cannot carry enough information to provide answers to all different experiments’, and hence, in the final analysis, does not provide new physics. Thus, even if such works had been successful in deriving the quantum theory, without explaining what an “elementary system” means ontologically, the employed assumption remains ineligible in grounding QM on a fundamental physical principle. (Rovelli [47] presented a possible reconstruction scheme, which parts of its ingredients were implemented in our work).

In general, the reconstruction of quantum theory starts by formulating the foundational principles which the authors believe plausible and then translating them into mathematical axioms. One expects the first principle to be simple physical statements with easily understandable meaning, not highly abstract mathematical assumptions (for example [39, 41]). Compared to our approach, it appears that the main shortcoming of such approaches could be constructing their theory based on mathematics rather than the physics of the situation. In particular, it appears customary to base the approach on information and bits, which are abstract mathematical concepts, rather than on physical messages that are pieces of information conveyed physically. In our approach, we primarily based our argument on the physics of the situations. After recognizing and defining the messaging capacity of a system, we used information theory's perfect secrecy rule to explain unavoidable randomness in the physical systems with a single messaging capacity. Following the construct on this discrete physical property, we finally could derive the full theory.

The perfect secrecy rule has been utilized in a similar context in other works too. Predominantly in the works initiated by Zeilinger [49] and extended later in a series of papers [55-59]. It is worth analyzing how using the abstract notion of information, rather than the physical concept of the message, could lead to a construct stray. In that work, Zeilinger puts forward a Foundational Principle (FP) for QM that postulates a quantum system as an elementary system that carries one bit of information. The FP postulate then provides an explanation for the observed randomness in quantum measurements for the spin of a spin $1/2$ particle case, as was analyzed in the aforementioned paper. The proposed principle, however, fails to apply to quantum particles in general: the spin of a spin $3/2$ particle, for example, carries 2 bits of information, or a photon has a continuous energy state-space, which makes its information capacity infinite in principle [60]. Motivated by the FP postulate, Brukner and Zeilinger [61] tried to avoid this inconsistency in the information capacity of quantum particles by discussing the inapplicability of the Shannon information in quantum measurements and defining their own measure of information. Their argument against the suitability of the Shannon information in quantum measurements, however, turned out to be erroneous; besides the notion that the Shannon information only makes sense for the systems with pre-existing values, their reasoning against the applicability of the Shannon information in quantum measurements is based on an incorrect application of conditional probability in quantum physics (several textbook examples of this occasional misunderstanding was mentioned years earlier by Ballentine [62]). Incidentally, the FP postulate, which Brukner and Zeilinger based their picture upon, utilizes the perfect secrecy rule, which, in turn, is based on the Shannon information that they argued against its applicability in quantum physics. Not long after publication, Brukner and Zeilinger's argument on the inadequacy of the Shannon information in QM was comprehensively refuted by Timpson [63, 64] and the consistency of Shannon information in both the classical and quantum settings was emphasized in several papers [65-67]. Zeilinger and Brukner's measure of information has been criticized for several other reasons and still calls for careful analysis [68-70]. While Brukner and Zeilinger suggest their interpretation provides an explanation for a number of quantum phenomena, their speculative approach leaves an important point moot: that a quantum particle can be regarded as having just a single bit of information content.

Discussion and Conclusion

In classical physics, the dynamics of a particle's evolution is governed by its position and momentum. In QM, the uncertainty principle forbids simultaneous determination of the position and momentum of a quantum particle. No further explanation is offered for this fundamental limitation. Here we recognized and discussed the case of the SM-systems in which their single context recording limits

simultaneous determination of their independent properties. We demonstrated the “quantal” behavior of these systems is the natural consequence of their limited messaging capacity. Finally, we developed a theory to explain the dynamics of such SM-systems and derived the conventional formalism of QM.

We started primarily by recognizing that such a theory cannot be deterministic and would be a probabilistic theory. The task of the theory is then to describe how to evaluate the system's propensities for possible measurements outcomes. The input is the state of the SM-system, which, in turn, is the result of an already performed measurement. The derivation followed based on considering how the propensity of an SM-system transforms between different types of measurements. In particular, we used conservation of probability to find the properties of the operators acting on the state space of the SM systems. Significantly, the standard Hilbert-space formalism of QM, as well as the Born probability rule, was completely obtained from our construction.

When properly identified in terms of measurement procedures, the algebra of the theory describes the transformation of complex conditional propensity-amplitudes. This, in turn, originates from the correct computational tools to predict the corresponding probabilities for the measurements, rooted in the deterministic relations between physical properties. Two considerations helped us in configuring the formalism of the theory: 1) the relationships between measurements are invariant, and 2) one should be able to change her measurement decisions in principle and get the same propensities. The latter means that it should be mathematically possible to render the propensities of the system from one type of measurement to another, and in principle, that process should be reversible so that total propensity remains conserved. These observations yield us to derive the full formalism of the theory.

Our derivation of the formal machinery of QM gives us a clear interpretation of its constituent. In particular, the state describes the system's propensity-amplitudes for the outcomes of realizable measurements. Therefore, the state is not an abstract *a priori* concept but takes its reality from being the result of an already performed measurement on the system, with its objectivity being with respect to another realizable measurement. We also have a clear divide between classical physics and quantum physics, and thus the meaning of the linear Schrödinger equation and its limits of applicability, which in turn resolves the measurement problem.

Traditionally QM has been understood as an operator algebra over a Hilbert space of quantum states, with a few interpretation rules. While the mathematical premises of the theory lack physical grounding, the success of the theory in practice has been astonishing. Our derivation helps the theory to eventually acquire a precise meaning in virtue of the first principles. The only new physical concept that we introduced is the number of the independently adjustable properties of a physical system, which indicates the number of independent messages that the system can simultaneously carry. The premise that a quantum system is a physical system with a single adjustable property explains where the structure of the theory comes from. At the levels of the single content SM-systems, not more than a single proposition of objective reality can be defined. This leads to quantum theory not being a deterministic theory, but a general probabilistic theory. At its core, quantum theory is a mathematical tool for computing probabilities for the outcomes of the measurement that will be our experimental interventions to the system. The theory describes the transformation of the SM-systems' propensities for different measurements relative to each other. It is not a theory about the processes internal to the systems being measured that give rise to the measurement outcome. In other words, the theory is not as much about the system (which does not have complex internal mechanisms) but rather about what could be known about

the possible outcomes of realizable measurements. Quantum theory is not about the ingredients of the world or a theory of space or time. It is a theory of mechanics of the SM-systems whose role is to provide a framework for computing the probabilities of events given that other events have occurred by accounting for the general structure of the correlations in the physical world.

With this perspective, QM essentially is a probability theory for the SM systems that can be derived completely based on the relationships between the possible outcomes of different measurements. Its formalism is the consequence of the consistency of the bookkeeping of the system's propensities for different measurements. It is un-exotic and free of underlying metaphysical groundings mentioned in some complex theories or interpretations, such as hidden variables [71-74], multiple worlds [75-77], many minds [78-80], etc. Such works, while claiming to provide a better "explaining" of QM, have not served us with a better understanding of the underlying mechanism of the theory.

QM does have a simple and intuitive principle at its heart. Effectively, the principle behind its formalism is that the total sum of probability values should be one, i.e., the fact that experimenting yields a result, and therefore the sum of the propensities of the system for different outcomes should be one. This principle is not only simple but also has an easily understandable physical meaning: independently of a measurement's type, the propensities for the outcomes should add up to one. Therefore, in describing the transformation of the SM-system's propensities for different types of measurement, the total should be conserved. Once you posit this criterion, the rest of the theory follows from the consistency requirement. The extent of the structure of the theory – basically the whole standard Hilbert space formalism– that was derived solely from such a constraint is interesting. The complete formalism of QM can be derived from one simple constraint that the total probability of a set of related events should always add up to one.

Our derivation would not cause a huge revision in quantum theory from the familiar formalism we learned from our textbooks. However, it places the results of quantum physics into a much larger context. The presented picture explains the physical origins of the theory, clarifies where the formalism of QM comes from, and provides us with the vivid meanings of its elements. With this new perspective, we can revisit the scenarios described by conventional quantum mechanics and resolve the misunderstandings in our understanding of physics and the world around us. This work could be a beginning for us to understand quantum mechanics finally; a theory that perplexed generations of physicists, a puzzle that Feynman famously had described as "impossible, absolutely impossible, to explain in any classical way" which he suspected "very strongly that it is something that will be with us forever" [5]. In the light of the present derivation, now we have a clear and precise interpretation of the quantum theory.

References

- [1] P. Grangier, G. Roger, and A. Aspect, "Experimental evidence for a photon anticorrelation effect on a beam splitter: a new light on single-photon interferences," *EPL (Europhysics Letters)*, vol. 1, no. 4, p. 173, 1986.
- [2] F. Jelezko, A. Volkmer, I. Popa, K. Rebane, and J. Wrachtrup, "Coherence length of photons from a single quantum system," *Physical Review A*, vol. 67, no. 4, p. 041802, 2003.
- [3] V. Jacques *et al.*, "Single-photon wavefront-splitting interference," *The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics*, vol. 35, no. 3, pp. 561-565, 2005.

-
- [4] A. Zeilinger, G. Weihs, T. Jennewein, and M. Aspelmeyer, "Happy centenary, photon," *Nature*, vol. 433, no. 7023, pp. 230-238, 2005.
- [5] R. P. Feynman, "Feynman lectures on physics. Volume 3: Quantum mechanics," *Reading*, 1965.
- [6] M. O. Scully, B.-G. Englert, and H. Walther, "Quantum optical tests of complementarity," *Nature*, vol. 351, no. 6322, pp. 111-116, 1991.
- [7] V. Jacques *et al.*, "Experimental realization of Wheeler's delayed-choice gedanken experiment," *Science*, vol. 315, no. 5814, pp. 966-968, 2007.
- [8] W. K. Wootters and W. H. Zurek, "Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle," *Physical Review D*, vol. 19, no. 2, p. 473, 1979.
- [9] S. Lloyd, "Ultimate physical limits to computation," *Nature*, vol. 406, no. 6799, p. 1047, 2000.
- [10] D. Hanson *et al.*, "Detection of B-mode polarization in the cosmic microwave background with data from the south pole telescope," *Physical Review Letters*, vol. 111, no. 14, p. 141301, 2013.
- [11] E. Leitch *et al.*, "Measurement of polarization with the degree angular scale interferometer," *Nature*, vol. 420, no. 6917, pp. 763-771, 2002.
- [12] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, no. 3, pp. 379-423, 1948.
- [13] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, pp. 656-715, 1949.
- [14] D. Dieks, "Communication by EPR devices," *Physics Letters A*, vol. 92, no. 6, pp. 271-272, 1982.
- [15] W. Thirring and E. M. Harrell, *A course in mathematical physics 3: Quantum mechanics of atoms and molecules*. Springer, 1981.
- [16] F. London, "The λ -phenomenon of liquid helium and the Bose-Einstein degeneracy," *Nature*, vol. 141, no. 3571, pp. 643-644, 1938.
- [17] L. Tisza, "Transport phenomena in helium II," *Nature*, vol. 141, no. 3577, pp. 913-913, 1938.
- [18] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, "Theory of superconductivity," *Physical review*, vol. 108, no. 5, p. 1175, 1957.
- [19] T. Ando, Y. Matsumoto, and Y. Uemura, "Theory of Hall effect in a two-dimensional electron system," *Journal of the Physical Society of Japan*, vol. 39, no. 2, pp. 279-288, 1975.
- [20] C. Monroe, D. Meekhof, B. King, and D. J. Wineland, "A "Schrödinger cat" superposition state of an atom," *Science*, vol. 272, no. 5265, pp. 1131-1136, 1996.
- [21] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, "Quantum dynamics of single trapped ions," *Reviews of Modern Physics*, vol. 75, no. 1, p. 281, 2003.
- [22] L. G. Hanson, "Is quantum mechanics necessary for understanding magnetic resonance?," *Concepts in Magnetic Resonance Part A: An Educational Journal*, vol. 32, no. 5, pp. 329-340, 2008.
- [23] R. Jaklevic, J. Lambe, A. Silver, and J. Mercereau, "Quantum interference effects in Josephson tunneling," *Physical Review Letters*, vol. 12, no. 7, p. 159, 1964.
- [24] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. Van der Zouw, and A. Zeilinger, "Wave-particle duality of C₆₀ molecules," *nature*, vol. 401, no. 6754, p. 680, 1999.
- [25] S. Gerlich *et al.*, "Quantum interference of large organic molecules," *Nature communications*, vol. 2, p. 263, 2011.
- [26] S. Gerlich *et al.*, "A Kapitza-Dirac-Talbot-Lau interferometer for highly polarizable molecules," *Nature Physics*, vol. 3, no. 10, pp. 711-715, 2007.
- [27] J. Von Neumann, *Mathematical foundations of quantum mechanics*. Princeton university press, 2018.
- [28] L. E. Ballentine, "The statistical interpretation of quantum mechanics," *Reviews of modern physics*, vol. 42, no. 4, p. 358, 1970.

- [29] A. Peres, *Quantum theory: concepts and methods*. Springer, 2002.
- [30] J. R. Searle and S. Willis, *The construction of social reality*. Simon and Schuster, 1995.
- [31] J. B. Hartle, "Quantum mechanics of individual systems," *American Journal of Physics*, vol. 36, no. 8, pp. 704-712, 1968.
- [32] R. N. Giere, "Objective single-case probabilities and the foundations of statistics," in *Studies in Logic and the Foundations of Mathematics*, vol. 74: Elsevier, 1973, pp. 467-483.
- [33] L. E. Ballentine, "Propensity, probability, and quantum theory," *Foundations of physics*, vol. 46, no. 8, pp. 973-1005, 2016.
- [34] J. D. Trimmer, "The present situation in quantum mechanics: A translation of Schrödinger's "Cat Paradox" paper," *Proceedings of the American Philosophical Society*, pp. 323-338, 1980.
- [35] E. Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik," *Naturwissenschaften*, vol. 23, no. 50, pp. 844-849, 1935.
- [36] E. P. Wigner, "Remarks on the mind-body question," in *Philosophical reflections and syntheses*: Springer, 1995, pp. 247-260.
- [37] P. A. M. Dirac, *The principles of quantum mechanics* (no. 27). Oxford university press, 1981.
- [38] G. Birkhoff and J. Von Neumann, "The logic of quantum mechanics," *Annals of mathematics*, pp. 823-843, 1936.
- [39] G. W. Mackey, *Mathematical foundations of quantum mechanics*. Courier Corporation, 2013.
- [40] C. Piron, "Axiomatique quantique," *Helvetica physica acta*, vol. 37, no. 4-5, pp. 439-&, 1964.
- [41] C. Piron, "Survey of general quantum physics," *Foundations of physics*, vol. 2, no. 4, pp. 287-314, 1972.
- [42] 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.
- [43] B. Mielnik, "Geometry of quantum states," *Communications in Mathematical Physics*, vol. 9, no. 1, pp. 55-80, 1968.
- [44] A. Landé, "Albert Einstein and the quantum riddle," *American Journal of Physics*, vol. 42, no. 6, pp. 459-464, 1974.
- [45] D. I. Fivel, "How interference effects in mixtures determine the rules of quantum mechanics," *Physical Review A*, vol. 50, no. 3, p. 2108, 1994.
- [46] 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.
- [47] C. Rovelli, "Relational quantum mechanics," *International Journal of Theoretical Physics*, vol. 35, no. 8, pp. 1637-1678, 1996.
- [48] N. Landsman, "Simple new axioms for quantum mechanics," *International journal of theoretical physics*, vol. 37, no. 1, pp. 343-348, 1998.
- [49] A. Zeilinger, "A foundational principle for quantum mechanics," *Foundations of Physics*, vol. 29, no. 4, pp. 631-643, 1999.
- [50] B. Coecke, D. Moore, and A. Wilce, "Current research in operational quantum logic," in *Series Fundamental Theories of Physics*, vol. 111: Springer, 2000.
- [51] L. Hardy, "Quantum theory from five reasonable axioms," *arXiv preprint quant-ph/0101012*, 2001.
- [52] 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.
- [53] 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.
- [54] 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.

- [55] 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.
- [56] C. Brukner and A. Zeilinger, "Malus' law and quantum information," *acta physica slovacica*, vol. 49, pp. 647-652, 1999.
- [57] C. Brukner, M. Zukowski, and A. Zeilinger, "The essence of entanglement," *arXiv preprint quant-ph/0106119*, 2001.
- [58] 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.
- [59] 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.
- [60] P. P. Rohde, J. F. Fitzsimons, and A. Gilchrist, "Information capacity of a single photon," *Physical Review A*, vol. 88, no. 2, p. 022310, 2013.
- [61] Č. Brukner and A. Zeilinger, "Conceptual inadequacy of the Shannon information in quantum measurements," *Physical Review A*, vol. 63, no. 2, p. 022113, 2001.
- [62] L. E. Ballentine, "Probability theory in quantum mechanics," *American Journal of Physics*, vol. 54, no. 10, pp. 883-889, 1986.
- [63] C. G. Timpson, "On a supposed conceptual inadequacy of the Shannon information in quantum mechanics," *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, vol. 34, no. 3, pp. 441-468, 2003.
- [64] C. G. Timpson, "The applicability of Shannon information in quantum mechanics and Zeilinger's foundational principle," *Philosophy of Science*, vol. 70, no. 5, pp. 1233-1244, 2003.
- [65] M. J. Hall, "Comment on" Conceptual inadequacy of Shannon information..." by C. Brukner and A. Zeilinger," *arXiv preprint quant-ph/0007116*, 2000.
- [66] A. Duwell, "Quantum information does not exist," *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, vol. 34, no. 3, pp. 479-499, 2003.
- [67] P. G. L. Mana, "Consistency of the Shannon entropy in quantum experiments," *Physical Review A*, vol. 69, no. 6, p. 062108, 2004.
- [68] A. E. Rastegin, "On the Brukner–Zeilinger approach to information in quantum measurements," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2183, p. 20150435, 2015.
- [69] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, "On mutually unbiased bases," *International journal of quantum information*, vol. 8, no. 04, pp. 535-640, 2010.
- [70] A. Shafiee, F. Safinejad, and F. Naqsh, "Information and the Brukner-Zeilinger interpretation of quantum mechanics: A critical investigation," *Foundations of Physics Letters*, vol. 19, no. 1, pp. 1-20, 2006.
- [71] D. Bohm, "A suggested interpretation of the quantum theory in terms of" hidden" variables. I," *Physical review*, vol. 85, no. 2, p. 166, 1952.
- [72] D. Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II," *Physical Review*, vol. 85, no. 2, pp. 180-193, 01/15/ 1952, doi: 10.1103/PhysRev.85.180.
- [73] L. De Broglie, *Une tentative d'interprétation causale et non linéaire de la mécanique ondulatoire:(la théorie de la double solution)*. Gauthier-Villars, 1956.
- [74] L. De Broglie, *Non-linear wave mechanics: A causal interpretation*. Elsevier Publishing Company, 1960.
- [75] H. Everett III, "" Relative state" formulation of quantum mechanics," *Reviews of modern physics*, vol. 29, no. 3, p. 454, 1957.

- [76] J. A. Wheeler, "Assessment of Everett's" relative state" formulation of quantum theory," *Reviews of modern physics*, vol. 29, no. 3, p. 463, 1957.
- [77] B. S. DeWitt, "Quantum mechanics and reality," *Physics today*, vol. 23, no. 9, pp. 30-35, 1970.
- [78] D. Albert and B. Loewer, "Interpreting the many worlds interpretation," *Synthese*, pp. 195-213, 1988.
- [79] M. Lockwood, "'Many Minds'. Interpretations of Quantum Mechanics," *The British journal for the philosophy of science*, vol. 47, no. 2, pp. 159-188, 1996.
- [80] M. J. Donald, "Progress in a many-minds interpretation of quantum theory," *arXiv preprint quant-ph/9904001*, 1999.

Appendix A: Information vs. inference

One may wonder about a claim such that a photon can convey at most three distinct messages on its direction, energy, and polarization. Is it not possible for example, to load two messages just on the two independent components of its polarization, regardless of its other degrees of freedom?

Here we should differentiate between our notion of information (semantic information) in contrast to the physical information (information carried by physical objects). In short, our concept of semantic information is what gets inferred and interpreted in a web of shared background knowledge and implied assumptions, while physical information bears no such references and comes with no tags.

The following example can help show the difference: one may use a pulse of laser with adjustable wavelength –say with exact values between 100 to 999 *nm*– to send a message to the other party. The information to be sent will be encoded on the wavelength of the light. With this setup, the physical carrier can convey a message worth up to $\log_2(899)=9.8$ bits of data. In the receiving end, however, the wavelength measured in *nm* can be converted to binary (a number between 0001100100 and 1111100111) in which the value on each placing can represent the answer to a different pre-assigned yes/no question, 9 in total. With such encodings and with a wise choice of questions, one can thus transmit a wealth of information just by using a pulse of the laser. So it seems, that the wavelength of the photons can convey a lot more than just one message.

One should, however, not confuse this human ability to encode many pieces of information on the wavelength of a beam of light –that benefits from many shared background knowledge– with the physical message that bears the information. In such encodings, the agents at the ends implicitly or explicitly benefit from myriads of shared background knowledge between themselves, e.g., they both know what a wavelength, a *nm*, number bases, the choice of the questions, their order, etc., are, which none are part of the transmitted physical message.

Similarly, a single electron does not carry on itself per se the background knowledge of, for example, spatial directions, so that the observer at the receiving end would be able to project and read the momentum in specific directions. In practice, physical information is distinguishable only based on the distinctive physical properties of the message carrier. Meaning, that a single physical property with a huge number of states, physically only bears one message containing a big chunk of data equal to $\log_2(\#states)$ bits.