Conservation of information and the foundations of quantum mechanics

Giulio Chiribella[1,a] and Carlo Maria Scandolo[1,b]

[1]*Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China*

Abstract. We review a recent approach to the foundations of quantum mechanics inspired by quantum information theory [1, 2]. The approach is based on a general framework, which allows one to address a large class of physical theories which share basic information-theoretic features. Here we first illustrate two very primitive features, expressed by the axioms of causality and purity-preservation, which are satisfied by both classical and quantum theory. We then discuss the axiom of purification, which expresses a strong version of the Conservation of Information and captures the core of a vast number of protocols in quantum information. Purification is a highly non-classical feature and leads directly to the emergence of entanglement at the purely conceptual level, without any reference to the superposition principle. Supplemented by a few additional requirements, satisfied by classical and quantum theory, it provides a complete axiomatic characterization of quantum theory in finite dimensions.

1 Introduction

A new approach to the foundations of quantum theory has emerged over the past three decades, drawing concepts and methods from the field of quantum information [3, 4]. This approach differs from that of many previous works in quantum foundations, which were primarily concerned with the interpretations of quantum mechanics and the measurement problem, maintaining the core of the Hilbert space formalism untouched. The recent works try instead to *derive* the Hilbert space framework from more basic principles regarding information-processing. In order to achieve this goal, one needs a more general framework capable of describing possible alternatives to quantum mechanics. These theories, called *general probabilistic theories* [1, 2, 5–12], describe the experiments can be performed with a given set of of physical devices, and provide a rule to assign probabilities to the outcomes of these experiments. Compared to the tradition of quantum logic [13–16], which also aimed at characterizing quantum theory into a larger landscape of theories, the new approach is characterized by the use of principles that are inspired by information-theoretic protocols, such as quantum teleportation [17]. All the recent axiomatizations of quantum theory [2, 6, 8, 12, 18–21] are clear examples of this new trend.

[a]e-mail: gchiribella@mail.tsinghua.edu.cn

[b]e-mail: scandolocm10@mails.tsinghua.edu.cn

In this paper we provide a non-technical introduction to general probabilistic theories, based on the framework established by D'Ariano, Perinotti, and one of the authors in Refs. [1, 2]. This framework is particularly apt to capture the operational aspects of a theory, making used of an intuitive graphical notation borrowed from the area of categorical quantum mechanics [22–24]. After introducing the framework, we discuss the axiomatization of quantum theory presented in Ref. [2]. We focus on three of the axioms, which we consider particularly fundamental. The first two axioms are causality and purity-preservation, which are satisfied by both classical and quantum theory. The third is the axiom of purification, which is not satisfied by classical theory and is responsible for many of the surprising features of quantum information. In particular, we show that purification leads directly to the No Cloning Theorem [25] and to the phenomenon of entanglement.

2 Why probabilistic theories?

Before entering into details, let us have a brief discussion on what the framework of general probabilistic theories aims to accomplish. After all, why going through the trouble of exploring more general theories, when quantum mechanics is already so successful in its predictions? In short, one can identify four reasons:

1. *Contribution to a deeper understanding of quantum mechanics.*
   Reconstructing a theory from basic physical principles, rather having just a mathematical description, helps building intuition and promoting the advancement of the theory itself. Think for example of Einstein's reconstruction of the Lorentz transformations from the principle of relativity and from the law of light propagation [26].

2. *Extensions and modifications of quantum mechanics.*
   Despite the present success of quantum mechanics, it is conceivable that the theory may need modifications in new regimes that have not been explored yet. The analysis of more general theories helps suggesting which quantum mechanical axioms can be modified to adapt the theory to new scenarios, such as those of a perspective theory of quantum gravity.

3. *Search for links amount quantum information protocols.*
   Quantum information theorists have devised a multitude of new protocols, which turn the counterintuitive features of quantum mechanics into advantages. A natural tendency is then to try to recognize the underlying patterns and to establish direct links between different quantum protocols. Besides the benefit of conceptual clarification, this may also help devising new protocols.

4. *Effective restrictions of quantum mechanics.*
   Suppose not all quantum states allowed by quantum mechanics are accessible with a particular experimental setup. For example, linear optics techniques can easily generate and manipulate Gaussian states of light, but are not able to access non-Gaussian states and operations. Given an effective theory describing a restricted subset of quantum states and operations, a natural question is "Which quantum features can be detected?". One way to provide an answer is to phrase the theory as a general probabilistic theory and check which axioms are satisfied.

For a more extended presentation of these arguments, we refer the reader to the insightful discussion by Hardy and Spekkens [27].

3 Systems and tests

We now provide a framework for general probabilistic theories, following the scheme of Ref. [1, 2] (see also [28]). In this framework there are two primitive notions, the notion of *physical system* and the notion of *test*.

A *test* represents one use of a physical device (for example, a beam-splitter, a polarimeter, or a Stern-Gerlach apparatus). Every device has an input system and an output system. We denote systems with capital letters, such as A, B, and so on. Among all systems, it is convenient to include the *trivial system*, which represents "nothing". A device with trivial system as input is a device with no input, and a device with trivial system as output is a device with no output.

In general, a device can have various outcomes, which can be e. g. a sequence of digits, or a spot in a photographic plate. The outcomes can be identified by the experimenter, and each outcome corresponds to a different process that can take place when the device is used. Hence, tests will be represented as collections of processes labelled by outcomes, such as $\{C_i\}$. We will also adopt a graphical language, which is intuitive and at the same time mathematically rigorous [22–24, 29]. Using this language, a test is represented as a box with incoming and outgoing wires that represent the input and output system, respectively. For example, the test $\{C_i\}$ will be represented as

$$\xrightarrow{A} \boxed{\{C_i\}} \xrightarrow{B} \quad.$$

If we want to specify which process occurred, we omit the braces, as in the following

$$\xrightarrow{A} \boxed{C_i} \xrightarrow{B} \quad.$$

Deterministic evolutions, like the unitary evolution induced by Schrödinger's equation, can be represented in this framework as tests with only one possible outcome (where the meaning of the outcome is just that the evolution took place).

The process of preparing a state can be also described as a test, specifically a test with the trivial system as input and the system that is being prepared as output. A test of this form, say $\{\rho_i\}$, is called *preparation-test* and represents a device which prepares the system in a state ρ_i, randomly chosen from the set $\{\rho_i\}$. We represent a preparation-test as

$$\left(\!\{\rho_i\} \xrightarrow{A} \right. \quad.$$

On the other hand, also destructive measurements can be represented as tests. These special tests have no output (trivial system as output), and destroy the input system while acquiring some information from it, as it happens e. g. when an electron is absorbed by a photographic plate, leaving a spot on it. A test of this form, say $\{a_i\}$, is called *observation-test* and each individual process a_i represents one way of destroying the input system. Using graphical language, we represent an observation-test as

$$\xrightarrow{A} \left. \boxed{\{a_i\}}\right) \quad.$$

4 Sequential and parallel composition

Having fixed how to represent devices, the next step is to describe how to connect them. Devices can be connected in sequence or in parallel. In sequential composition the two devices are connected one after the other. To do so, clearly the input of the second device must be the same as the output of the first device, as shown in the following example:

$$\left(\!\{\rho_i\} \xrightarrow{A} \boxed{\{C_j\}} \xrightarrow{B} \{b_k\}\!\right) \quad. \tag{1}$$

The above diagram gives instructions on how to build up an experiment: in this experiment one first initializes system A with the preparation-test $\{\rho_i\}$ on A, then performs the test $\{C_j\}$, which transforms system A into system B, and finally one acquires information from B by performing the observation-test $\{b_k\}$. If we wish to express which events actually occurred, we write

$$\left(\rho_i \;\overset{A}{\quad}\; \boxed{C_j} \;\overset{B}{\quad}\; b_k \right). \tag{2}$$

This means that the state ρ_i was prepared, the process C_j took place, and finally the system has been destroyed, producing the outcome k. In general, note that there is a strict ordering in sequential composition: some tests are performed first and other later. In the graphical language, the ordering goes from left to right. This ordering will be essential to phrase the causality axiom (section 8), which forbids signalling from the future to the past.

Let us move now to parallel composition. Parallel composition of tests arises when we apply two devices to two different systems independently. The parallel composition of two processes is simply represented as

$$\overset{A}{\quad} \boxed{C_i} \overset{B}{\quad}$$
$$\overset{C}{\quad} \boxed{\mathcal{D}_j} \overset{D}{\quad} \; .$$
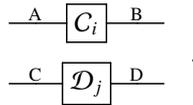
An important difference between sequential and parallel composition is that, when two processes are composed in parallel, the order in which they take place does not matter.

A particular case of parallel composition is the composition of preparation devices. When a preparation device prepares system A in the state ρ and another device prepares system B in the state σ, we say that the composite system AB is in a *product state*, denoted by $\rho \otimes \sigma$ and graphically represented as

$$\left(\rho \;\overset{A}{\quad} \right.$$
$$\left(\sigma \;\overset{B}{\quad} \; .$$
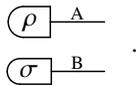
It is important to note that the operations that can be performed on a composite system are not restricted to product operations. In general, one can also use a joint device which processes the component systems together. Such a process represents the result of an *interaction*, such as e. g. the interaction between two beams of particles in an accelerator. Joint devices will be represented as boxes with multiple wires, one wire for each system, as in the following example:

$$\overset{A}{\quad} \overset{C}{\quad} \boxed{\; C_i \;} \overset{}{\underset{B \quad D}{}} \; .$$
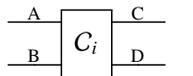
5 A consistent rule for predicting probabilities

When we have a diagram with no external wires, like in equation (2), we interpret it as a *probability*. This is a shorthand notation to mean that a process that starts with the preparation of a state and ends with the destruction of the system yields a probability. For example, diagram (2) represents the *joint probability* that the state ρ_i is prepared, the transformation C_j takes place, and the process b_k destroys the system.

The rule for specifying the probabilities of all possible diagrams with no external wires is assumed as part of the specification of the theory. The only requirements on this rule are *i)* that the sum of the probabilities for all the outcomes produced in an experiment should be equal to one and *ii)* that the outcome probabilities for experiments performed in parallel should be of the product form.

6 Purity of states and transformations

In both classical and quantum statistical mechanics, it is common to distinguish between pure states and mixed states. For example, in the quantum case the pure states are described by rays in the Hilbert space of the system, while mixed states are described by density operators. The distinction between pure and mixed states, however, is not specific to quantum or classical theory: in fact, it makes sense in every probabilistic theory. More generally, one can define also pure and mixed transformations.

The idea at the basis of the definition is *coarse-graining*. Let us clarify it with an easy example. In the roll of a die, there are six basic outcomes, identified by numbers from 1 to 6. However, we can consider a coarse-grained outcome, which results from join together some of the basic outcomes, neglecting some information. This is the case, for example, when we just say that the outcome the roll was an odd number. This coarse-grained outcome is the union of the basic outcomes 1, 3, and 5.

Clearly, doing a coarse-graining, we lose some information. We can say that a transformation is *pure* if it does not arise as a coarse-graining of other transformations; in this way, a pure transformation represents a process of which we have maximal information. An example of pure transformation in quantum theory is the unitary evolution resulting from Schrödinger's equation. The definition of pure transformation also applies to states, which are a particular type of transformations, namely the transformations implemented by preparation devices. A state that is not pure is called *mixed*: when a system is described by a mixed state, one has only partial information about the preparation. For instance, we may know that some pure states ψ_i's are prepared with given probabilities p_i's. If we ignore which state ψ_i has been prepared, we describe the system with the mixed state $\rho = \sum_i p_i \psi_i$, which can be regarded as a sort of "expectation state" of the system. Here we are using the symbol of sum just as a notation for probabilistic coarse-graining. However, with a little bit of work, one can actually define a suitable notion of sum of transformations, using the rule for probabilities that is provided by the theory [1].

7 Preservation of purity

Equipped with the notions of pure state and pure transformation, we can now discuss one of the axioms of Ref. [2]. The axiom provides an answer to the following question: Is the composition of two pure transformations still pure? Intuitively, when we have maximal knowledge of two processes, we should also have maximal knowledge of the process that results from their composition. This intuitive requirement is formalized by the axiom of *purity-preservation*:

Axiom 1 (Purity-preservation)**.** *Sequential and parallel compositions of pure transformations are pure transformations.*

Purity-preservation is a very primitive requirement. Think of the theory as an algorithm, used by the physicist to make deductions from known facts: given as datum that system A undergoes the process C from time t_0 to t_1 and the process $\mathcal{D}$ from time t_1 to t_2, the algorithm deduces that system A undergoes the process $\mathcal{D} \circ C$ from time t_0 to t_2. The lack of purity-preservation would mean that the algorithm is not able to determine what really happened to the system after a sequence of time-steps, even when provided with the most precise input about each individual step. Not even quantum mechanics is so random: for example, the composition of two unitary evolutions is still a unitary evolution, and not a stochastic process with multiple outcomes. More generally, it is easy to see that both classical and quantum theory satisfy the axiom of purity-preservation.

8 Causality

Another very primitive requirement about physical theories is causality. One can phrase the axiom as follows:

Axiom 2 (Causality). *The probability of an outcome at a certain step does not depend on the choice of experiments performed at later steps.*

The "steps" mentioned here are the steps in a sequence of laboratory operations, such as those depicted in the example (1). In that particular example, the causality axiom ensures that the probability that the system A is prepared in the state ρ_i does not depend on the choice of test $\{C_j\}$ or on the choice of the destructive measurement $\{b_k\}$. Informally, the causality axiom states that it is impossible to signal from the future to the past. It is easy to see that both classical theory and quantum theory satisfy this requirement.

The impossibility of signalling from the future to the past implies the impossibility of instantaneous signalling across space [1]. Suppose that two distant parties, conventionally called Alice and Bob, perform two independent tests $\{\mathcal{A}_i\}$ and $\{\mathcal{B}_j\}$ on two systems in their laboratories. Since the two tests are performed in parallel, the order does not matter, i. e. the probability of the outcomes should not depend on whether Alice performs her test first or Bob performs his test first. Combining this observation with the causality axiom, we have that the probability that Alice finds the outcome i should not depend on the choice of test $\{\mathcal{B}_j\}$ performed by Bob, and vice-versa. Of course, there can be correlations between Alice's and Bob's outcomes, if the state of the two systems ρ_{AB} is not of the product form $\rho_A \otimes \rho_B$. These correlations will be described by some joint probability distribution, $p_{AB}(i, j)$, but the marginal probability distribution $p_A(i) = \sum_j p_{AB}(i, j)$ is independent on the choice of test $\{\mathcal{B}_j\}$, and the marginal probability distribution $p_B(j) = \sum_i p_{AB}(i, j)$ is independent on the choice of test $\{\mathcal{A}_i\}$. In other words, the correlations that arise form a causal theory do not allow for signalling across space. As a consequence, if Alice wants to send a message to Bob, she has to send some physical system. This also implies that the speed of every message is limited by the maximum speed at which physical systems can be transferred in space. For example, if we assume that the maximum speed coincides with the speed of light in the vacuum, we have that faster-than-light communication is ruled out in every general probabilistic theory that satisfies the causality axiom.

9 The purification principle and the conservation of information

So far we discussed axioms that are satisfied both by classical and quantum theory. However, in order to identify quantum theory one needs at least one axiom that is not satisfied by classical theory. Such an axiom should capture what makes the quantum world so radically different from the classical one, and possibly, should allow one to deduce the key protocols of quantum information theory. One axiom that possesses these features is the purification axiom [1, 2]:

Axiom 3 (Purification). *Every physical process can be simulated in an essentially unique way as a reversible evolution of the system interacting with a pure environment.*

Let us unpack the content of this statement. First, the axiom tells us that every process, even an irreversible one, can be modelled as the result of a reversible process, where the system interacts with the environment. The origin of the irreversibility is only in the fact that the environment is discarded: at least in principle, if the experimenter were able to maintain full control of the degrees of freedom that are interacting during the experiment, the overall evolution would be reversible. This fact expresses the principle of Conservation of Information, which states that, at some fundamental

level, information cannot be destroyed, but can only be discarded. The Conservation of Information, *per se*, is not a distinctive quantum feature: for example, it is satisfied also by Newton's equations of motion and by other dynamical equations in classical physics. What is distinctive about quantum theory is the combination of the Conservation of Information with the "pure environment" part of the purification axiom: when we model an irreversible process as a reversible evolution of the system together with an environment, we can always start with the environment in a pure state. In classical physics, it is possible to simulate a stochastic process through a reversible process, but the price for such a simulation is that one needs an external source of randomness, provided by the environment. Instead, the purification axiom imposes that even the stochastic processes can be simulated without the need of initial randomness.

Consider, for example, the process of preparing a system A in a mixed state ρ. In this case, the purification axiom ensures that we can prepare ρ with the following procedure:

1. prepare A in a pure state α and prepare another system E in a pure state η

2. let the composite system AE evolve with a suitable reversible process $\mathcal{U}$, thus obtaining the pure state $\Psi_{AE} = \mathcal{U}(\alpha \otimes \eta)$

3. discard system E.

Here the system E plays the role of the environment and the pure state Ψ is called a *purification* of ρ. For example, in quantum theory it is possible to prepare a two-level system in the mixed state $\rho = (|0\rangle\langle 0| + |1\rangle\langle 1|)/2$ by first preparing two systems in the pure product state $|0\rangle \otimes |0\rangle$ and then letting them evolve with a suitable unitary evolution that transforms the state $|0\rangle \otimes |0\rangle$ into the singlet state $|S\rangle = (|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle)/\sqrt{2}$. By discarding the second system, one remains with the first system in the mixed state $\rho = (|0\rangle\langle 0| + |1\rangle\langle 1|)/2$. By contrast, in classical theory it is impossible to simulate the preparation of a mixed state using only pure states and reversible evolutions.

More generally, the purification axiom guarantees that every physical process C, transforming the state of system A from ρ_A to $\rho'_A = C(\rho_A)$, can be simulated according to the following procedure

1. combine system A with another system E, prepared in a pure state η

2. let the composite system AE evolve with a suitable reversible process $\mathcal{U}$, thus obtaining the state $\rho'_{AE} = \mathcal{U}(\rho_A \otimes \eta)$

3. discard system E.

The third and last important point about the purification axiom is that the reversible simulation of a physical process is "essentially unique": if two reversible evolutions of system AE, say $\mathcal{U}$ and $\mathcal{U}'$, simulate the same process, then there exists a reversible evolution $\mathcal{V}_E$, acting only on the environment, such that $\mathcal{U}' = (\mathcal{I}_A \otimes \mathcal{V}_E) \circ \mathcal{U}$. In other words, this means that all the reversible evolutions that simulate the same process on system A must be equivalent, up to a "gauge transformation" on the environment. In quantum theory this means that two unitary operators U and U', acting on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_E$ and simulating the same process on A, must be equal up to a change of basis in the Hilbert space $\mathcal{H}_E$. The uniqueness up to reversible transformations is a very important feature, because it guarantees that all the models that we can invent to account for the irreversibility of a process are physically equivalent.

10 The no-cloning theorem

Among the appealing features of the purification axiom there is the fact that it gives direct access to many of the key structures of quantum information [1]. To give the flavour of how this is accomplished, here we show how the purification axiom can be used to derive the No-Cloning Theorem [25] in the context of general probabilistic theories.

Suppose that we want to construct a copy machine, which takes as input a system A_1 in some pure state α and another identical system A_2 in a fixed state α_0, providing as output systems A_1 and A_2 in the state $\alpha \otimes \alpha$. If the machine works for every possible pure state α, then we call it a *universal copy machine*. While in classical physics there is no limitation in principle about the construction of universal copy machines, in quantum physics one has the No-Cloning theorem [25], stating that no physical process can make two perfect replicas of an arbitrary pure state.

Let us see how this result can be derived directly from our axioms. The proof is by contradiction: suppose that there exists a process that transforms $\alpha \otimes \alpha_0$ into $\alpha \otimes \alpha$ for every pure state α. By the purification axiom, this process can be realized by combining the input systems with another system E (the environment) in some pure state η, and letting them evolve reversibly via some process $\mathcal{U}$, thus obtaining the state $\mathcal{U}(\alpha \otimes \alpha_0 \otimes \eta)$. Now, by definition of copy machine, after discarding system E, the systems A_1 and A_2 should be in the state $\alpha \otimes \alpha$. This implies that, in particular, system A_1 should be in the state α. Hence, from the input to the output, the state of system A_1 undergoes the identity transformation $\alpha \mapsto \alpha$. Summarizing, the identity transformation on A_1 can be realized by 1) combining system A_1 with the system A_2E in the state $\alpha_0 \otimes \eta$, 2) applying the reversible evolution $\mathcal{U}$, and 3) discarding system A_2E. On the other hand, another (trivial) way to realize the identity transformation on A_1 is to combine it with the system A_2E in the state $\alpha_0 \otimes \eta$, apply the reversible transformation $\mathcal{U}' = \mathcal{I}_{A_1} \otimes \mathcal{I}_{A_2} \otimes \mathcal{I}_E$, and discard system A_2E. Since the reversible simulation of physical processes is unique up to reversible transformations on the environment, one must have $\mathcal{U}' = (\mathcal{I}_{A_1} \otimes \mathcal{V}_{A_2E})\mathcal{U}$ for some reversible transformation $\mathcal{V}_{A_2E}$ acting only on A_2E. Equivalently, this means that $\mathcal{U} = (\mathcal{I}_{A_1} \otimes \mathcal{V}_{A_2E}^{-1})\mathcal{U}' \equiv \mathcal{I}_{A_1} \otimes \mathcal{V}_{A_2E}^{-1}$. Using this relation, we obtain

$$\mathcal{U}(\alpha \otimes \alpha_0 \otimes \eta) = \alpha \otimes \Psi_{A_2E}, \qquad \Psi_{A_2E} := \mathcal{V}_{A_2E}^{-1}(\alpha_0 \otimes \eta).$$

Note that, by definition, the state Ψ_{A_2E} is independent of α. Hence, we reached a contradiction: if we discard systems A_1 and E on both sides, the l.h.s. is equal to α (by the hypothesis that $\mathcal{U}$ realizes a copying process) and the r.h.s. is independent of α. In conclusion, we proved that the process $\alpha \otimes \alpha_0 \mapsto \alpha \otimes \alpha$ cannot be realized in any theory satisfying the purification principle.

Conceptually, proving the No-Cloning theorem directly from purification is an important result: it tells us that the fact that the existence and essential uniqueness of a reversible simulation of physical processes implies the impossibility of universal copying machines. It is also important to recall that the No-Cloning Theorem is the working principle at the basis of the security of quantum cryptographic protocols, such as the BB84 key distribution protocol [30]. Having derived this theorem from first principles suggests that one may be able to provide also an axiomatic proof of the security of key distribution based on the purification axiom.

11 Entanglement

Entanglement is one of the weirdest features of quantum mechanics [31, 32]. It gives rise to correlations that cannot be explained by any local realistic model [33] and deeply challenge our intuition about the microscopic world. Besides being puzzling, however, entanglement is also a precious resource for quantum communication protocols [34, 35], such as quantum teleportation [17]

In quantum mechanics, entanglement appears as a mathematical consequence of the superposition principle. But is there a deeper reason for its existence? In order to answer the question, we need first to provide a theory-independent definition of entanglement. In section 4, we introduced product states as the result of independent preparation operations performed in parallel for different systems. When two systems A and B are in a pure product state, say $\alpha \otimes \beta$, one can assign to each system a pure state: system A is in the state α and system B is in the state β. By contrast, we say that a pure state Ψ is *entangled* if it is not a product state. When systems A and B are in an entangled state, one has a maximal knowledge about the composite system AB, without having maximal knowledge of its parts. The above definition of entanglement captures an idea expressed by Schrödinger [32], who famously wrote

> "the best possible knowledge of a whole does *not necessarily* include the best possible knowledge of all its parts".

It is easy to see that every probabilistic theory satisfying the purification axiom must have entangled states [1]. Indeed, suppose that the pure states of system AB are only of the product form $\alpha \otimes \beta$ for some pure states α and β of systems A and B, respectively. Then, when we discard system B, the remaining state of system A is pure. In conclusion, the composite system AB cannot be used to purify any mixed state of A. In order for the purification principle to hold, there must exist at least one system B such that the composite system AB contains one entangled state.

In summary, the purification axiom leads directly to entanglement. Conceptually, this is also an important point, because entanglement gives rise to the most dramatic differences between quantum and classical theory. In the same paper quoted above, Schrödinger expressed the intuition that entanglement is at the centre of the structure that characterizes quantum mechanics, writing

> "I would not call that *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought".

In a sense, the axiomatization of finite dimensional quantum theory presented in Ref. [2] fulfills this intuition: the basic rules of the Hilbert space framework can be derived from the purification axiom (which arguably captures the structures that Schrödinger was highlighting in his paper), in combination with five other axioms that are satisfied also by classical theory, such as e. g. the axioms of causality and purity-preservation.

12 Conclusions

In this paper we reviewed a basic language of general probabilistic theories and illustrated three operational axioms that can be expressed in this language, namely purity-preservation, causality, and purification. In a nutshell: Purity-preservation is the requirement that maximal knowledge of the processes happening in a sequence of time-steps implies maximal knowledge of the overall process from the input to the output. Causality is the requirement that no signal can be sent from the future to the past. Purification is the requirement that every physical process can be simulated in an essentially unique way as a reversible process where the system interacts with an environment, initially prepared in a pure state. Purification expresses a strengthened form of the principle of Conservation of Information: it guarantees that one can always account for irreversibility by formulating a model where, at the fundamental level, information is preserved, and, in addition, it guarantees that such a model is essentially unique and does not require a source of randomness in the environment. Purification leads directly the some key structures in quantum mechanics, like the No-Cloning Theorem and the existence of entangled states. Combined with purity-preservation, causality, and other three axioms

satisfied also by classical theory, purification leads to a complete axiomatization of quantum theory in finite dimensions. The big picture emerging from this derivation is that the core of quantum theory can be identified in the ability to simulate irreversible and stochastic processes using only pure resources and reversible evolutions.

References

[1] G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A **81**, 062348 (2010)

[2] G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A **84**, 012311 (2011)

[3] C.A. Fuchs, Journal of Modern Optics **50**, 987 (2003)

[4] G. Brassard, Nature Physics **1**, 2 (2005)

[5] S. Popescu, D. Rohrlich, Foundations of Physics **24**, 379 (1994)

[6] L. Hardy, eprint arXiv (2001), `quant-ph/0101012`

[7] J. Barrett, Phys. Rev. A **75**, 032304 (2007)

[8] G.M. D'Ariano, in *Philosophy of Quantum Information and Entanglement*, edited by A. Bokulich, G. Jaeger (Cambridge University Press, Cambridge, 2010), pp. 85–126

[9] H. Barnum, J. Barrett, M. Leifer, A. Wilce, Phys. Rev. Lett. **99**, 240501 (2007)

[10] H. Barnum, A. Wilce, Electronic Notes in Theoretical Computer Science **270**, 3 (2011), proceedings of the Joint 5th International Workshop on Quantum Physics and Logic and 4th Workshop on Developments in Computational Models (QPL/DCM 2008)

[11] L. Hardy, in *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, edited by H. Halvorson (Cambridge University Press, Cambridge, 2011), pp. 409–442

[12] L. Hardy, eprint arXiv (2011), `1104.2066`

[13] G.W. Mackey, American Mathematical Monthly pp. 45–57 (1957)

[14] G. Ludwig, *Foundations of quantum mechanics* (Springer, 1985)

[15] C. Piron, *Foundations of quantum physics*, Mathematical Physics Monograph Series (Benjamin-Cummings Publishing Company, 1976)

[16] D.J. Foulis, C.H. Randall, in *Interpretations and foundations of quantum theory* (1981)

[17] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993)

[18] P. Goyal, K.H. Knuth, J. Skilling, Phys. Rev. A **81**, 022109 (2010)

[19] B. Dakić, C. Bruckner, in *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, edited by H. Halvorson (Cambridge University Press, Cambridge, 2011), pp. 365–392

[20] L. Masanes, M.P. Müller, New Journal of Physics **13**, 063001 (2011)

[21] L. Masanes, M.P. Müller, R. Augusiak, D. Pérez-García, Proceedings of the National Academy of Sciences **110**, 16373 (2013)

[22] B. Coecke, eprint arXiv (2005), `quant-ph/0510032`

[23] B. Coecke, Contemporary Physics **51**, 59 (2010)

[24] P. Selinger, in *New Structures for Physics*, edited by B. Coecke (Springer, Berlin, 2011), pp. 289–356

[25] W. Wootters, W. Zurek, Nature **299**, 802 (1982)

[26] W. Rindler, *Introduction to Special Relativity*, Oxford science publications (Clarendon Press, 1991)

[27] L. Hardy, R. Spekkens, Phys. Can. **66**, 73 (2010)

[28] G. Chiribella, G.M. D'Ariano, P. Perinotti, Entropy **14**, 1877 (2012)

[29] S. Abramsky, B. Coecke, in *Handbook of Quantum Logic and Quantum Structures: Quantum Logic*, edited by K. Engesser, D.M. Gabbay, D. Lehmann (Elsevier, 2008), pp. 261–324

[30] C.H. Bennett, G. Brassard, *Public key distribution and coin tossing*, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing* (1984), pp. 175–179

[31] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. **47**, 777 (1935)

[32] E. Schrödinger, Mathematical Proceedings of the Cambridge Philosophical Society **31**, 555 (1935)

[33] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, 2nd edn. (Cambridge University Press, 2004)

[34] M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information*, 1st edn. (Cambridge University Press, 2004)

[35] M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013)