HAL open science

The arrow of time in operational formulations of quantum theory

Andrea Di Biagio, Pietro Donà, Carlo Rovelli

To cite this version:

Andrea Di Biagio, Pietro Donà, Carlo Rovelli. The arrow of time in operational formulations of quantum theory. Quantum, 2021, 5, pp.520. 10.22331/q-2021-08-09-520 . hal-02981210

HAL Id: hal-02981210

https://hal.science/hal-02981210v1

Submitted on 2 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The arrow of time in operational formulations of quantum theory

Andrea Di Biagio[1], Pietro Donà[2], and Carlo Rovelli[2,3,4]

[1]Dipartimento di Fisica, Sapienza Università di Roma, 00185 Roma, Italy

[2]Aix-Marseille Université, Université de Toulon, CNRS, CPT, 13288 Marseille, France

[3]Perimeter Institute, 31 Caroline Street North, Waterloo Ontario N2L2Y5, Canada

[4]The Rotman Institute of Philosophy, 1151 Richmond St. N, London Ontario N6A5B7, Canada

arXiv:2010.05734v2 [quant-ph] 29 Jul 2021

The operational formulations of quantum theory are drastically time oriented. However, to the best of our knowledge, microscopic physics is time-symmetric. We address this tension by showing that the asymmetry of the operational formulations does not reflect a fundamental time-orientation of physics. Instead, it stems from built-in assumptions about the *users* of the theory. In particular, these formalisms are designed for predicting the future based on information about the past, and the main mathematical objects contain implicit assumption about the past, but not about the future. The main asymmetry in quantum theory is the difference between knowns and unknowns.

Introduction

Classical mechanics is invariant under time reversal: its elementary laws do not distinguish past from future. The observed arrow of time is a macroscopic phenomenon that depends on the use of macroscopic variables and the contingent fact that the entropy defined by these variables was lower in the past. Is it the same for quantum mechanics?

On the one hand, the Schrödinger equation is time reversal invariant and so is quantum field theory (up to parity transformation and charge conjugation). Elementary physics is time reversal invariant and the source of time orientation is again macroscopic and entropic. Elementary quantum phenomena do not carry a preferred arrow of time. On the other hand, however, the formalism of quantum theory is often defined in a markedly time oriented way.

Here we address this tension between the physics and the formalism. We investigate the reason for the time orientation of the quantum formalism and show that the tension can be resolved. The asymmetry in the formalism is due to the inherent directionality in the process of inference, which is related to the arrow of time only indirectly.

Andrea Di Biagio: andrea.dibiagio@uniroma1.it
Pietro Donà: pietro.dona@cpt.univ-mrs.fr
Carlo Rovelli: rovelli@cpt.univ-mrs.fr

Let us start by noting that, in any inferential problem, there is an asymmetry between what is known (the data), and what is unknown (the desiderata). Let us call this directionality the *arrow of inference*. The arrow of inference is not necessarily aligned with the entropic arrow of time. The arrow of inference may be pointed towards the past as well as towards the future. Quantum phenomena are such that we can only compute conditional probabilities, so quantum theory inherits the asymmetry between data and desiderata.

Quantum theory allows us to compute the probability of future events from past ones, but it also allows us to compute the probability of past events from future ones. As we illustrate in detail below, quantum theory does not distinguish between these two tasks. In contrast, the users of quantum theory are generally more interested in predicting the future than postdicting the past. We live in thermodynamically oriented world that has abundant macroscopic traces of the past but not of the future [1, 2]. Hence in most problems, the arrow of inference points in the same direction as the arrow of time. As a result we have designed formulations of quantum theory that conflate the two arrows. Ignoring the distinction may be a source of confusion.

We will focus on formalisms used in quantum information [3–5]. These are designed to study information processing tasks and the correlations that agents can achieve by sharing and manipulating quantum systems. This approach has lead to a wealth of insights, both of theoretical and technological value [6–20]. In particular, the information-theoretic reconstructions of quantum theory [21–30] derive the formal Hilbert space structure of quantum theory from simple physical principles (much like Einstein's two postulates of special relativity allowed to re-derive the Lorentz transformations [31, 32]). With the notable exception of Ding Jia's [29], the reconstructions either start by considering a space of theories that is intrinsically time-oriented [21–23, 25–27, 29] or introduce the time orientation explicitly as a postulate [24, 28, 30] ("no

signaling from the future").

There is nothing wrong with time-oriented formalisms designed to study time-oriented questions. But the success of this approach risks obscuring the time reversal invariance of elementary physics, and this could turn into an obstacle when extending it to the investigation of phenomena outside the laboratory. An example of such situations is quantum gravitational phenomena, in which the existence of a background spacetime cannot be taken for granted [33, 34].

During the development of this work, Schmid, Selby, and Spekkens [35], and Hardy [36] have proposed formalisms in which the physical and the inferential aspects of a theory can be separated and make space for time-symmetric physics. The present work can be seen as an additional motivation for these frameworks.

Our argument proceeds as follows. We start with the uncontroversial assumption that the Born rule gives *prediction* probabilities: conditional probabilities for future events, given past ones. We apply standard probability theory to find formulas for *postdiction* probabilities: probabilities about the past, given the future. Note that we do not postulate these probabilities: we derive them from the prediction probabilities.

In section 2, we show that for closed quantum systems the probabilities for prediction and postdiction are identical, a property we call *inference symmetry*. The Born rule can be used in both directions of time without modification, contrary to what is sometimes stated.

In section 3, we discuss open quantum systems, where this symmetry is hidden. In that case, the prediction and postdiction probabilities differ. However the difference is dictated by the asymmetries in the inferential problem, not by the arrow of time, once again contrarily to what is often stated in the literature. Unitary quantum mechanics is both time-symmetric and inference-symmetric.

In section 4, we investigate the same question for *quantum channels*, the more general evolutions featuring in operational formulations of quantum theory used in quantum information. Quantum channels are not, in general, inference-symmetric. By shifting the Heisenberg cut to include part of the apparatus, we show that the inference-asymmetry of quantum channels stems from asymmetries in the inferential data.

In section 5, we relate the tasks of postdiction with passive and active time-reversals, and discover that quantum channels can also be seen as shorthands for calculations about the past.

We combine all insights in section 6 to show that the asymmetries of the quantum information formulations do not stem from an arrow of time intrinsic to all quantum systems, but from the asymmetry inherent in the process of inference. The time-asymmetry of the operational formalisms used in quantum information theory is that of the time-oriented macroscopic agents that set up the experiments.

At the end of the paper, we briefly discuss the time orientation of other formulations (Copenhagen, Everettian, de Broglie-Bohm) as well.

Previous work

The exploration of the time-symmetry of quantum uncertainty started early with Einstein, Tolman, and Podolski [37] already noting in 1931 that "the principles of quantum mechanics actually involve an uncertainty in the description of past events which is analogous to the uncertainty in the prediction of future events."

Aharonov, Bergmann, and Lebowitz [38] build a time-symmetric theory out of quantum theory by considering the frequencies of outcomes of a sequence of projective measurements in ensembles constructed by pre- and post-selection. They note that this theory is time-symmetric in the sense that the frequencies observed in one ensemble are the same as the ones observed in the ensemble prepared by swapping the pre- and post-selection and performing the measurements in the reversed order. They note that probabilities calculated based only on preselection are experimentally accurate, while probabilities calculated only on post-selection are insufficient. They argue that this time-asymmetry is not inherent to quantum mechanics but is a consequence of the asymmetry of the macroscopic world. Finally they advance that this asymmetry should be represented by adding a time-asymmetric postulate to the time-symmetric theory. In light of our work, we don't need to build a time-symmetric theory and then break the symmetry. Quantum mechanics is already symmetric in the sense that generally it does not distinguish prediction and postdiction, or the past from the future. Our work also has a different aim: to understand the time-asymmetry of the operational formulations of quantum theory.

While Aharonov *et. al.* [38] considered a sequence of measurements "sandwiched" between selection events, the setup in our work is similar to that in [39], where Watanabe was concerned with calculating the probabilities for a past or future event given a present event. Watanabe introduced *retrodictive quantum mechanics*, where a state is assigned to the system based on present data and then evolved to past times. Like Aharonov *et. al.*, Watanabe remarks that "blind retrodiction," (*i.e.* what we call postdiction: retrodiction with flat priors on the past events) does not work well in practice because agents in the past can decide to interrupt or go on with the experiments. Watanabe also recognises that inference is inherently asymmetric and can run in either direction of time, starting from data at a given instant.

Retrodictive quantum mechanics has been further developed over the decades by Barnett, Pegg, Jeffers and collaborators [40–42] who devise general postdiction formulas, including ones equivalent those we derive in sections 2 and 4.2 for postdiction in closed systems and for quantum channels respectively. Our contribution to retrodictive quantum mechanics is twofold. First, we propose the explanation of the asymmetry between prediction and postdiction in quantum channels in terms of implicit data about the past of a purifying system. Secondly, we prove *no-signalling from the unknown*, which is a property of quantum mechanics, and in particular of retrodictive quantum mechanics, that to our knowledge has not been recognised before. We also pay attention to the conceptual difference between retrodiction and time-reversal, and we relate these two concepts.

Leifer and Pusey [43] consider a similar prepare-and-measure scenario as us and investigate what time-symmetry can imply on possible ontological models for quantum theory. Their definition of an *operational time reverse* formally equivalent to what we call an active time-reversal. They define *operational time-symmetry* as the existence of an operational time reverse. They derive a fascinating no-go theorem for ontic extensions (aka hidden variable models) of operational time-symmetric processes. We share the belief that operational time-symmetry is an essential feature of quantum mechanics, but we do not concern ourselves with ontic extensions. Instead, we are interested in understanding why not all quantum channels are operationally time-symmetric. We also study the difference and relation between postdiction and time-reversal, prove that operational time-symmetry is equivalent to inference symmetry.

Oreshkov and Cerf [44] define an extension to operational quantum theory, allowing for a "notion of operation that permits realizations via both pre- and post-selection." Their motivation for building a new theory is stated in the abstract: "The symmetry of quantum theory under time reversal has long been a subject of controversy because the transition probabilities given by Born's rule do not apply backward in time." Our work shows that the Born rule applies equally well in both directions in time—as long as we treat prediction and postdiction on equal footing. We argue that the asymmetry of operational quantum theory reflects the asymmetry of the agents, not the asymmetry of quantum phenomena per se.

Our work was inspired by conversations at the QISS conference at HKU and at the QISS virtual seminars [45], where it transpired that the time-asymmetric operational formulation obfuscated the fundamental time-symmetry of quantum theory.

The importance of separating the physical from the inferential in quantum mechanics is a more modern idea, perhaps traceable to E.T. Jaynes who famously compared quantum theory to an omelette to be unscrambled [46]. The QBists see quantum theory as not much more than the correct probability calculus to use in our world [47]. Leifer and Spekkens [48] have formally developed the analogy between quantum probabilities and Bayesian inference. They introduce a notion of "quantum conditional states" representing sequential and parallel quantum experiments and prediction and postdiction on the same footing, just like in classical probability theory. In our work, we limit ourselves to classical probability theory. At first, we use the Born rule to obtain classical conditional probability distributions $P_{pre}(x|a, U)$ for prediction probabilities in sequential experiments. We show that the Born rule actually can be used to compute prediction and postdiction probabilities.

We summarize the main original contributions of this work.

- We interpret the inference-asymmetry of quantum channels as due to implicit information about the purifying system. (Section 4.3).
- We note that the asymmetry of the operational formulation of quantum mechanics is due to three factors.
    1. The asymmetry intrinsic in the process of inference (Section 6)
    2. The operational formulation is interested in prediction.(Section 4.1)
    3. Quantum operations contain built-in assumptions about the past and the ability of agents to pre-select states, and this is where the time-symmetry is broken. (Sections 4.3 and 6.4)
- The proof of the *no-signalling from the further unknown* property of quantum mechanics, a consequence of the conservation of probability and not an independent axiom. (Section 6.3) When applied to prediction, it reduces to the well-known "no-signalling from the future."
- The derivation that inference-symmetry is equivalent to bistochasticity (Section 4.4) and hence to operational time-symmetry (Section 5.2).

1 Prediction and postdiction

Quantum indeterminism is time-reversal invariant. In presenting the probabilistic nature of quantum theory, we often emphasise that the future of a quantum system is not entirely determined by its past. It certainly is true that the outcomes of future interactions with a quantum system are uncertain, given the details of past interactions. What is less often recognised is that the converse is equally true: given the details of present interactions, the past ones are uncertain. This was already pointed out a long time ago

[37]: the irreducible indeterminism of quantum phenomena cuts both ways, leaving both the past and the future uncertain, given data about the present. This has practical consequences such as the impossibility of deterministic state-discrimination and the no-cloning theorem [3]: interactions with a quantum system do not allow us to guess with certainty how it was prepared.

As we will see below, the past of a quantum system is *quantitatively* as uncertain as its future: the probabilities calculated using the Born rule can be applied to both predict and postdict. Quantum theory, in fact, does not distinguish *a priori* the tasks of prediction and postdiction and we might say that there is a fundamental "unpostdictability" of the behaviour of quanta.

Let us operationally define what we mean by prediction and postdiction using two related tasks. In both tasks, a friend prepares a quantum system in an initial configuration, allows it to undergo a given transformation, measures it, and finds it in some final configuration. The friend then gives us some information and about these events, and asks us to guess the rest. In the first task, we are asked to guess the outcome of the measurement, given the initial configuration and the details of the transformation and of the measurement. In the second task, we guess the outcome of the preparation, given the outcome of the measurement, the details of the transformation, and the set of possible initial states.

Definition 1 (Prediction task)**.** *Given a preparation, a map and a test and the outcome of the preparation, calculate the probabilities for the outcomes of the test.*

Definition 2 (Postdiction task)**.** *Given a preparation, a map and a test and the outcome of the test, calculate the probabilities for the outcomes of the preparation.*

These are inferential tasks, in which we use the available information to make educated guesses. If we want to think of probabilities in frequentist terms, we can imagine the friend repeating the setup many times, ensuring a uniform distribution in the initial or final configuration and interpret the calculated probabilities as relative frequencies of outcomes in an ensemble of trials, in the limit of infinite trials.

While the setting of these tasks might appear artificial at first, a moment of thought reveals that it serves as a useful shorthand for physically relevant situations. In fact, postdiction has been extensively studied before [40, 49, 41, 50] and has a number of practical applications; see [42, 51, 52] and references therein.

In the following, we study the relation between these two tasks, as it captures the role played by the arrow of time in quantum theory. First, we consider the case of closed systems, namely when the system under consideration is isolated between preparation and observation. Then we consider open systems, namely when we ignore some degrees of freedom, such as environmental degrees of freedom. Finally, we extend the analysis to the more general case in which the notions of preparation, evolution and measurement are subsumed in the more general idea of *operation* used in quantum information and quantum foundations.

2 Closed systems

In this section, we assume that the friend prepares the system by determining the values of a maximal set of compatible observables and does the same at observation. We also assume that the system under consideration is isolated between preparation and observation. Therefore, the preparation and test are represented by orthonormal bases of the Hilbert space associated with the quantum system and the transformation is represented by a unitary transformation.

In this case, the Born rule is equally good for predicting the future given the past and postdicting the past given the future [39]. Since this fact is not universally known, we derive it assuming only the uncontroversial fact that the Born rule can be used to predict the future.

We denote by $\{a_i\}_{i=1}^d$ and $\{x_i\}_{i=1}^d$ the bases of the preparation and test, respectively—although we drop the basis indices when they are not strictly needed, to keep the notation cleaner. The solution to the prediction task with the unitary evolution U, and the outcome of the preparation a is given by the Born rule: the probability

$$P_{pre}(x|a,U) = |\langle x|U|a\rangle|^2, \tag{1}$$

for the outcome x of the test. The solution to the postdiction task is obtained from the solution of the prediction game and standard probability theory. By Bayes' theorem

$$P_{post}(a|x,U) = \frac{P_{pre}(x|a,U)P(a)}{P(x)}, \tag{2}$$

where $P(a)$ is the prior probability on the initial configuration and $P(x)$ is the probability of the final configuration given the prior. Since all we know in the postdiction task is the basis of the preparation, we have

$$P(a_i) = \frac{1}{d} \tag{3}$$

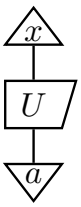
for all $i = 1, \dots d$. The *a priori* probability of the outcome of the test is computed summing over all possible initial states:

$$P(x) = \sum_{i=1}^{d} \frac{1}{d} P_{pre}(x|a_i,U) = \frac{1}{d}, \tag{4}$$

where we used the fact that the evolution is unitary. The postdiction probability is then computed from (2):

$$P_{post}(a|x,U) = |\langle x|U|a\rangle|^2 = P_{pre}(x|a,U), \tag{5}$$

which, in terms of the pictorial calculus of Ref. [5], reads

$P_{pre}(x|a,U) =$  $= P_{post}(a|x,U).$ (6)

Thus for a closed quantum system, the solution to the prediction and postdiction tasks is given by the same formula.

Note that the flat prior (3) is crucial in the derivation above. Had the prior been different, the postdiction probabilities would be different from the prediction probabilities. However, assuming a different prior would introduce an inappropriate asymmetry between the two tasks. Our objective is to treat the prediction and postdiction tasks on equal footing. When we predict the result of a measurement using the Born rule, we do not assume any prior knowledge on the result of the experiment besides the space of alternatives. Therefore, in a postdiction task we do not assume any prior knowledge of the result of the preparation besides the space of alternatives, hence the the flat prior. The flat prior does not imply that the input system was prepared in the maximally mixed state, it is simply the probability distribution that represents the prior knowledge in the postdiction task.

For a system evolving unitarily between the preparation and the observation events, later events are uncertain given the earlier event *and* earlier events are uncertain given later events, and the probabilities are given by the same formula. The Born rule does not distinguish the past from the future: it allows to calculate the probability of an event given another event, no matter their order in time.

Let us formalise this property:

Definition 3 (Inference symmetry)**.** *A transformation Φ is inference symmetric if for any two orthonormal bases $\{a_i\}_{i=1}^{d_A}$ and $\{x_i\}_{i=1}^{d_X}$ for the input and output spaces respectively, the prediction and postdiction probabilities are identical for given i and j:*

$$P_{pre}(x_i|a_j,\Phi) = P_{post}(a_j|x_i,\Phi) \tag{7}$$

Quantum unitary evolution is inference-symmetric. Why is this time-symmetric aspect of unitary evolution rarely emphasised?

In most practical situations, we don't need to use the flat prior when guessing the past. We can do better, because there are macroscopic traces of the past, like our memories or entries in a notebook. Additionally, in most laboratory experiments, the initial distribution is known or chosen so that frequency of events in the ensemble is far from uniform.

The fact that macroscopic traces are records of the past and not the future and the fact that an experimenter's choice can affect the future and not the past are both macroscopic phenomena that pertain to the irreversible physics of the macroscopic world surrounding the experiment [2, 53], not to the quantum dynamics, which by itself does not know the arrow of time.

Some authors go so far as to say that the Born rule does not work 'backward in time' and see it as a fundamental asymmetry in the theory [44] and that quantum theory needs to be modified, or extended, to make it symmetric. But this is too quick. If we do not assume any knowledge or bias in the past, (5) is indeed the correct formula to use according to quantum theory. In turn, the validity of the Born rule in predicting the future relies on the same assumptions about the future. Namely, if we did have some knowledge of the future, then (1) would not be the best formula to make predictions. For example, if the detector does not detect certain states, the Born rule fails.

The presence of records of the past is not a property of quantum theory *per se*, or the behaviour of a single quantum, but a property of what surrounds the quantum. See also [39, 38] for early examples of this argument.

3 Open systems

Let us now consider the case when the system we deal with is not isolated. An open quantum system can always be seen as a part of a larger closed quantum system. To study this case, consider a tensor decomposition of the input $A\otimes B$ and output $X\otimes Y$ Hilbert spaces. The tasks we consider now regard computing probabilities restricted to some of these subspaces. Denote by d_A, d_B, d_X and d_Y the dimensions of the respective spaces and with $\{a_i\}$, $\{b_i\}$, $\{x_i\}$, and $\{y_i\}$ bases on them. The evolution between input and output space is represented by the unitary U. By the results of the previous subsection, this process is inference-symmetric with the solution

$$P_{pre}(xy|ab,U) = |\langle xy|U|ab\rangle|^2 = P_{post}(ab|xy,U). \tag{8}$$

Suppose that we agree with our friend to ignore the subspace Y of the outcome space and compute only the probability of finding x as the outcome of the test on X. This simulates the situation in which our system gets entangled with some other system that is subsequently ignored, like when information leaks into the environment. Note that the difference between a closed and an open system is in the inferential data,

not in the physical system. We can solve the prediction task by computing the marginal probability

$$P_{pre}(x|ab,U) = \sum_{i=1}^{d_Y} P_{pre}(xy_i|ab,U), \tag{9}$$

where we sum over the space Y we decided to neglect. Similarly we solve the postdiction task with the same unitary evolution, with only knowledge on the outcome of the test on the space X, weighting the original postdiction probabilities with a flat prior:

$$P_{post}(ab|x,U) = \sum_{i=1}^{d_Y} \frac{1}{d_Y} P_{post}(ab|xy_i,U). \tag{10}$$

We can use the inference symmetry of the closed system (8) to relate the two expressions above:

$$P_{post}(ab|x,U) = \frac{1}{d_Y} P_{pre}(x|ab,U). \tag{11}$$

Thus the prediction and postdiction probabilities are no longer equal once part of the output system is ignored.

Suppose now that we agree with our friend to neglect the B part of the input space. This corresponds to the situation in which a system in an unknown state interacts with our original system. Again, the difference is in the inferential data, not the physical setup. The prediction probabilities are obtained by assigning a flat prior to the system B:

$$P_{pre}(xy|a,U) = \sum_{i=1}^{d_B} \frac{1}{d_B} P_{pre}(xy|ab_i,U), \tag{12}$$

while the postdiction probabilities are

$$P_{post}(a|xy,U) = \sum_{i=1}^{d_B} P_{post}(ab_i|xy,U). \tag{13}$$

The two are again related using (8):

$$P_{post}(a|xy,U) = d_B P_{pre}(xy|a,U). \tag{14}$$

Again, the two probabilities are different.

We can similarly analyse the case where we agree with our friend to neglect the result of the preparation in B and the outcome of the test on Y, which simulates a situation in which the system is open to influences from an unobserved quantum system. The prediction and postdiction probabilities again differ by a simple normalisation constant:

$$d_Y P_{post}(a|x,U) = d_B P_{pre}(x|a,U). \tag{15}$$

Note that the probabilities are equal only when $d_B = d_Y$, so that $A \equiv X$. When the input and output spaces are treated symmetrically, prediction and postdiction tasks are symmetric.

Crucially, the normalisation factor that makes the two kinds of probabilities different does not depend on time: if we neglect a subsystem B at time t_1 and another system Y at time t_2, then the probabilities for the values at t_2 given the values at t_1 are d_Y/d_B times the probabilities of guessing the values at t_1 given the values at t_2. This has nothing to do with a pre-established direction of time. Indeed, it is true regardless of whether $t_1 < t_2$ (as in the example above) or $t_1 > t_2$.

In the general case, the asymmetry between the prediction and postdiction tasks arises because of an asymmetry in the inferential data, not because of an intrinsic asymmetry in the physics or the evolution of the system. Indeed, in all these cases, the underlying process U is inference-symmetric. The inferential tasks are asymmetric only when the inferential data are asymmetric.

The normalisation of the identity is determined by the arrow of inference

We can rephrase all the probability calculations above in terms of density operators. Under unitary evolution a density operator transforms as $\rho \mapsto U[\rho] := U\rho U^\dagger$. A pure state ψ can be represented as a density operator by the projector $|\psi\rangle\langle\psi|$ and the Born rule can be recast as a trace

$$|\langle xy|U|ab\rangle|^2 = \mathrm{tr}\left(\,|xy\rangle\langle xy|\;\; U[|ab\rangle\langle ab|]\right). \tag{16}$$

Let us start with prediction, the more familiar task. In this language the prediction probability (9) can be rewritten as

$$\begin{aligned} P_{pre}(x|ab,U) &= \sum_{i=1}^{d_Y} P_{pre}(xy_i|ab,U) \\ &= \sum_{i=1}^{d_Y} \mathrm{tr}\left(\,|xy_i\rangle\langle xy_i|\;\; U[|ab\rangle\langle ab|]\right) \\ P_{pre}(x|ab,U) &= \mathrm{tr}\left((|x\rangle\langle x| \otimes \mathbb{I}_Y)\; U[|ab\rangle\langle ab|]\right). \end{aligned} \tag{17}$$

The decision to ignore part of the output system is represented by the insertion of the identity operator $\mathbb{I}_Y$, which in this role is called the discard operator of the subsystem Y. This is the classic technique of 'tracing out' a subsystem to ignore its future. Equation (12) can be similarly recast as

$$P_{pre}(xy|a,U) = \mathrm{tr}\left(|xy\rangle\langle xy|\;\; U\left[|a\rangle\langle a| \otimes \frac{1}{d_B}\mathbb{I}_B\right]\right), \tag{18}$$

where the flat prior is represented by the maximally mixed state, the density operator $\mathbb{I}_B/d_B$. This is also the well known result of doing prediction based on partial information. Note that the two formulas above make it clear that the choice of basis of the ignored system is irrelevant to the computed probabilities.

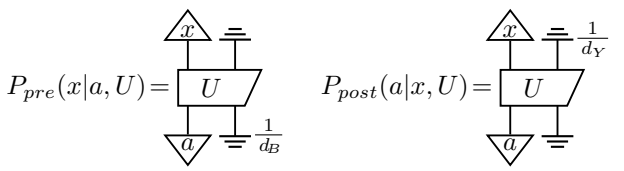
Let us now look at the postdiction probabilities. Using the two equations above together with (11) and (14), we can immediately write

$$P_{post}(ab|x,U) = \mathrm{tr}\left(\left(|x\rangle\langle x| \otimes \frac{1}{d_Y}\mathbb{I}_Y\right) U[|ab\rangle\langle ab|]\right), \tag{19}$$

$$P_{post}(a|xy,U) = \mathrm{tr}\left(\,|xy\rangle\langle xy|\;\; U[|a\rangle\langle a| \otimes \mathbb{I}_B]\right). \tag{20}$$

We see that the identity operator appears again in the ignored systems. However, the normalisation is the opposite of the predictive case.

The discard operator and the maximally mixed state are well known, and are normally only applied to the output and input side respectively. But the normalisation of the identity operator does not reflect the direction of time, the past or the future, input or output. It reflects the direction of inference. This is particularly obvious rewriting in the pictorial calculus:



(21)

with $\overline{\top}$ and $\underline{\perp}$ representing the identity operator as an output and input respectively. We discard in the direction we guess, and we have the maximally mixed state on the side of the data.

Thus the way the inference-symmetry appears to be broken in open systems reflects an asymmetry on the inferential data that is, *a priori*, independent on the direction of time.

An example illustrates why it is natural that the normalisation depends on the direction of inference and is independent on the direction of time. Consider a system evolving with the identity, *i.e.* nothing happens to it. If we are told the state a of the system and we are not asked to guess anything, then all the probabilities are trivially $1 = \mathrm{tr}\,|a\rangle\langle a|$. Conversely, if we are only told the system is in one state out of a orthonormal basis, then the probability that it is in a given state a is $\frac{1}{d} = \frac{1}{d}\,\mathrm{tr}\,|a\rangle\langle a|$. When we don't guess, all the probabilities are 1. When we are told to guess but we have no clue, our only option is to assume a uniform distribution. This is true regardless of whether we are predicting or postdicting

We often use quantum theory to predict, which is why we generally associate the normalisation factor $1/d$ to the identity operator in the input space. In practice, we are normalising our data. If we were postdicting, we would associate the normalisation factor to the identity in the output space. The operator $\mathbb{I}/d$ does not represent a physical fact, but the probability distribution we use to weight the conditional probabilities.

The results of these two sections show that the irreducible quantum uncertainty applies equally to both directions of time. Indeed, when dealing with a closed quantum system, the Born rule gives both the prediction and postdiction probabilities directly. When the prediction and postdiction probabilities differ, they do so because of an asymmetry in the inferential data.

4 Quantum Operations

The transformations considered above might seem limited in scope to researchers in quantum information and quantum foundations. In these communities, the notions of preparation, evolution and measurement are subsumed by the more general notion of *operation*, which reflects their more elaborate needs: a more coarse-grained description of quantum processes, independent on the underlying dynamics, the capacity of melding classical and quantum information processing, dealing with classical uncertainty and so on. However, since agents and labs are made of atoms and photons, and the interactions between atoms and photons is satisfyingly described by the unitary evolution and pure state approach, the results of the previous section have bearing on quantum operations too.

After a brief survey of the notion of quantum operation, we solve the prediction and postdiction tasks for a general quantum channel and explain their postdiction asymmetry.

4.1 Operations

Here we provide a short description of the notion of operation (for a more careful introduction see for example [4, 5, 3]). We then note that this notion is time-oriented by design and recall how it relates to the more basic notions.

An *operation* $\mathcal{O}^{A\to X}$, also known as an *instrument*, from an input Hilbert space A to an output Hilbert space X is represented is a set $\{O_i\}$ of completely positive (CP) trace non-increasing linear maps (aka *quantum maps*) from the space $\mathcal{L}(A)$ of linear operators on A to $\mathcal{L}(X)$, satisfying the *completeness equation* (aka the *causality condition*):

$$\forall \rho \in \mathcal{L}(A): \quad \sum_i \mathrm{tr}\, O_i[\rho] = \mathrm{tr}\,\rho. \tag{22}$$

An operation $\mathcal{O}^{A\to X} = \{O_i\}$ also defines a completely positive, trace-preserving (CPTP) map $\rho \mapsto \mathcal{O}[\rho] := \sum_i O_i[\rho]$. When the operation $\mathcal{O}^{A\to X}$ is applied to a system in state ρ, the outcome i happens with probability given by the generalised Born rule

$$P(i|\rho,\mathcal{O}) = \mathrm{tr}\, O_i[\rho], \tag{23}$$

and the state of the system after this outcome is

$$\rho_i = \frac{O_i[\rho]}{\mathrm{tr}\, O_i[\rho]}. \tag{24}$$

We see that the completeness requirement (22) amounts to a statement of the conservation of probabilities. If the outcome i is unknown, then the state of the system is a mixture of the states above, weighed by the relevant probability:

$$\mathcal{O}[\rho] = \sum_i P(i|\rho, \mathcal{O})\rho_i = \sum_i O_i[\rho]. \tag{25}$$

If the output space of an operation $\mathcal{O}^{A\to X}$ coincides with the input space of another operation $\mathcal{M}^{X\to Y}$, these two operations can be sequentially composed forming a new operation $(\mathcal{M}\circ\mathcal{O})^{A\to Y} = \{M_j \circ O_i\}$. If the state of the input system is ρ, the probability of the outcome ij is

$$P(ij|\rho, \mathcal{M}\circ\mathcal{O}) = \operatorname{tr} M_j[O_i[\rho]]. \tag{26}$$

Operations can also be composed in parallel using the tensor product structure of the underlying Hilbert spaces.

A *preparation* of a system associated with a Hilbert space A is an operation $\mathcal{P}^{I\to A}$. By a simple mathematical isomorphism, quantum maps from I to A can always be associated with positive linear operators on A, so that any preparation can be represented by a subset $\{\rho_i\} \subset \mathcal{L}(A)$ such that $\sum_i \operatorname{tr}\rho_i = 1$. Above, we have considered only preparations of the form $\{|a_i\rangle\langle a_i|/d_A\}_{i=1}^{d_A}$, where $\{a_i\}$ is an orthonormal basis for A.

A *test* on the same system is an operation $\mathcal{T}^{A\to I}$ from the Hilbert space A to the trivial Hilbert space. By the isomorphism above, and the Riesz representation theorem, tests are often also represented by a collection of positive operators $\{\sigma_j\} \subset \mathcal{L}(A)$, such that $\sum_j \sigma_j = \mathbb{I}_A$ with their actions on the state given by $\rho \mapsto \operatorname{tr}\sigma_j\rho$.

An operation $\Phi^{A\to X}$ with a single outcome is also called a *quantum channel*, and is represented by a CPTP map. Quantum channels are also called *deterministic* quantum maps, as they have only one outcome. Above, we considered only unitary quantum channels, of the form $\rho \mapsto U\rho U^\dagger$, but much more general ones are possible. For example, consider the *dephasing* channel on a qubit:

$$\rho \longmapsto \langle 0|\rho|0\rangle\ |0\rangle\langle 0| + \langle 1|\rho|1\rangle\ |1\rangle\langle 1| \tag{27}$$

The dephasing channel can be seen as the result of applying the nondestructive projective measurement

$$\left\{\rho \mapsto |i\rangle\langle i|\,\rho\,|i\rangle\langle i|\right\}_{i=0}^{1} \tag{28}$$

and ignoring the result.

A preparation with a single outcome is also called a *state*. States are by definition represented as unit-trace positive operators, *i.e.* density matrices. A test with a single outcome is also called a *deterministic effect*. There is only one deterministic effect, represented by the identity operator.

Note that this formalism is time-asymmetric *by construction*. The time asymmetry shows up in two ways, reminiscent of the Copenhagen-type interpretations. First, the outcomes $\{i\}$ depend probabilistically on the state of the system in the past: the probabilities calculated in this setting are invariably *prediction* probabilities. Second, the state of a system at any point in time reflects events in the past, and it is independent of the events in the future. The only data assumed to be available is data about the past.

The spaces of states and effects are not isomorphic. This was identified as the main source of time-asymmetry of operational quantum theory in [44], where it was proposed to enlarge the space of effects by not requiring that operations sum up to trace-preserving maps. However, from the perspective of our present work, we understand this asymmetry between states and effects as being the difference between known and unknown in the process of inference. Preparations represent our assumptions in the inferential problem, while tests represent the different propositions about the unknowns. There is no need to remove the distinction between preparations and tests to make quantum theory time-symmetric, all is needed is to recognise that operational quantum theory is geared for prediction: a situation where the data is in the past of the unknowns.

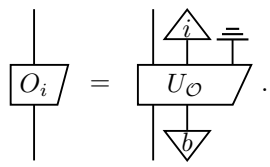
Operational quantum theory is connected to the simpler setting of pure states and unitary evolutions by the concept of *purification*. Any quantum channel $\Phi^{A\to X}$ can be purified [54], meaning it can be represented by a unitary channel $U_\Phi : A\otimes B \to X\otimes Y$ and a pure state b for system B such that

$$\forall\rho\in\mathcal{L}(A):\ \Phi[\rho] = \operatorname{tr}_Y U_\Phi[\rho\otimes|b\rangle\langle b|]. \tag{29}$$

In other words, any quantum channel can always be understood as a unitary interaction with an ancilla quantum system prepared in a specific way and where part of the output is ignored. In fact, any operation $\mathcal{O}^{A\to X} = \{O_i\}$ can be purified [55], meaning that it is mathematically equivalent to a unitary evolution of the system A in the presence of an ancilla B, in which part of the output system is ignored and part of it is measured on an orthonormal basis. That is, there exists a unitary operator $U_\mathcal{O}$ on the Hilbert space $A\otimes B$ and a decomposition $A\otimes B \equiv X\otimes Y\otimes Z$ and a pure state $|b\rangle\langle b| \in \mathcal{L}(B)$ such that

$$\begin{aligned}&\forall\rho\in\mathcal{L}(A), \forall i:\\ &\quad O_i[\rho] = \operatorname{tr}_{YZ}\left((|i\rangle\langle i|_Y \otimes \mathbb{I}_Z)\circ U_\mathcal{O}[\rho\otimes|b\rangle\langle b|]\right),\end{aligned} \tag{30}$$

where i now labels an orthonormal basis of the Hilbert space Y. Pictorially,



(31)

It is a well-known property of quantum theory that one can always shift the Heisenberg cut to include part of the apparatus in the quantum system under description. This is the physical content of the two mathematical results above. The label i that distinguishes the various outcomes of the operation $\mathcal{O}$ is now seen as labelling the possible values of a *pointer variable*, a part of the apparatus that gets entangled with the system and that, when observed, allows the determination of the state of the system. If one cared enough, one could model every quantum operation explicitly in these terms. But countless purifications are consistent with the same operation, and these low-level details are not important when one is only interested in the effect on the given quantum systems. Quantum operations are useful shorthand.

4.2 Prediction and postdiction with quantum channels

We now solve the prediction and postdiction tasks using the generalised Born rule, similarly to how we did for closed quantum systems. Unitary evolution is replaced by a CPTP map Φ. The input and output spaces are A and X respectively, which do not need to be isomorphic. Preparation and measurement are performed on bases $\{a_i\}_{i=1}^{d_A}$ and $\{x_i\}_{i=1}^{d_X}$. The solution of the prediction task is given by the generalised Born rule:

$$P_{pre}(x|a,\Phi) = \operatorname{tr} |x\rangle\langle x|\, \Phi[|a\rangle\langle a|]. \tag{32}$$

The solution to the postdiction task is found again by Bayesian inversion

$$P_{post}(a|x,\Phi) = \frac{P_{pre}(x|a,\Phi)P(a)}{P(x)}. \tag{33}$$

Assigning a flat prior for $P(a)$, we can compute the probability of the data

$$P(x) = \sum_{i=1}^{d_A} \frac{1}{d_A} P_{pre}(x|a_i,\Phi) = \operatorname{tr} |x\rangle\langle x|\, \Phi\left[\frac{1}{d_A}\mathbb{I}_A\right], \tag{34}$$

which this time is not uniform. Equation (33) then becomes

$$P_{post}(a|x,\Phi) = \frac{\operatorname{tr} |x\rangle\langle x|\, \Phi[|a\rangle\langle a|]}{\operatorname{tr} |x\rangle\langle x|\, \Phi[\mathbb{I}_A]}. \tag{35}$$

This formula was also derived in [40]. At first, it looks quite different from the formula for the prediction probabilities. However, it is also a remarkably simple solution. Note that the postdiction probabilities are related to the prediction probabilities by a simple multiplicative factor

$$P_{post}(a|x,\Phi) = f_\Phi(x) \cdot P_{pre}(x|a,\Phi), \tag{36}$$

where

$$f_\Phi(x)^{-1} = \sum_{i=1}^{d_A} P_{pre}(x|a_i,\Phi) = \operatorname{tr} |x\rangle\langle x|\, \Phi[\mathbb{I}_A]. \tag{37}$$

For a given measurement outcome x, the prediction probabilities are proportional to the postdiction probabilities, up to a fixed normalisation factor $f_\Phi(x)$. Once one has calculated the set of prediction probabilities, one already has the postdiction probabilities, up to this normalisation factor.

Inference symmetry is thus broken in general, and this seems to support the idea there is a fundamental difference between the past and the future in quantum theory. But it is broken in a simple way, reminiscent to the situation in the previous section when we were concerned with partial data in a closed system. Indeed, when we "look under the hood" using purification, we see that this is essentially what is going on. From the purified point of view, there is an asymmetry between inputs and outputs because the purifying ancilla's state is assumed known in input and ignored in output.

4.3 Purified task

Let Φ be a quantum channel from Hilbert space A to Hilbert space X. Then there exists a unitary channel $U_\Phi : A \otimes B \to X \otimes Y$ and a pure state b for system B such that

$$\Phi[|a\rangle\langle a|] = \operatorname{tr}_Y U_\Phi[|a\rangle\langle a| \otimes |b\rangle\langle b|]. \tag{38}$$

The prediction probabilities for Φ are, by definition, just those of the corresponding pure open system:

$$P_{pre}(x|a,\Phi) = P_{pre}(x|ab,U_\Phi), \tag{39}$$

as given in (9). However, in general

$$P_{post}(a|x,\Phi) \neq P_{post}(ab|x,U_\Phi). \tag{40}$$

This is readily explained by the fact that the initial state of the ancilla is assumed known, or "held fixed." Since we know the ancilla's input state is b, the probability of the data is in fact

$$P(x) = P_{pre}(x|b,U_\Phi). \tag{41}$$

Thus, by Bayes' theorem,

$$P_{post}(a|x,\Phi) = \frac{P_{pre}(x|ab,U_\Phi)}{d_A P_{pre}(x|b,U_\Phi)} = \frac{P_{pre}(x|a,\Phi)}{d_A P_{pre}(x|b,U_\Phi)}, \tag{42}$$

which indeed immediately translates to (35). We now have a different perspective on the normalisation factor $f_\Phi(x)$: it quantifies the implicit knowledge about the input ancilla system and how this knowledge affects the postdiction task. The specification of the quantum channel Φ contains information about the past of the purifying system system so that postdiction on the quantum channel Φ is equivalent to postdiction on the purified system but with some added information about the input system.

Indeed, we can arrive at the same formula by using the postdiction probabilities for the purified open system and the simple formula $P(a|b) = P(ab)/P(b)$. Indeed one can verify that

$$P_{post}(a|x,\Phi) = \frac{P_{post}(ab|x,U_\Phi)}{P_{post}(b|x,U_\Phi)}, \tag{43}$$

by using (11) and (15).

While we have used an arbitrary purification, the two equations above hold for any purification of the quantum channel Φ. So, whatever the physically appropriate purification might be, the lesson is the same: the inference asymmetry for a quantum channel derives not from an asymmetry of quantum mechanics, but from an asymmetry in the questions asked.

4.4 Inference-symmetric channels

We have seen that not every quantum channel is inference-symmetric. Thanks to the solution (35) to the postdiction task in the case of a general quantum channel, we immediately see that a channel Φ is inference-symmetric if and only if $f_\Phi(x) = 1$ for all pure states x, namely, if and only if it is identity-preserving:

$$\Phi[\mathbb{I}_A] = \mathbb{I}_X. \tag{44}$$

Since quantum channels are trace-preserving, it follows that the input and output spaces are isomorphic. The trace and identity preserving maps are known as the bistochastic quantum maps or unital channels [56, 57], and they are the free operations of the resource theory of quantum thermodynamics [58] with trivial hamiltonians, and the resource theory of purity [59].

Every unitary channel is obviously bistochastic. The noisy operations:

$$\rho \longmapsto \mathrm{tr}_B\, U\left[\rho \otimes \frac{1}{d_B}\mathbb{I}_B\right]. \tag{45}$$

form a strict subset of the bistochastic channels [58, 60]. A noisy operation represents the evolution of a system that undergoes unitary interaction with an ancilla about which nothing is known. These channels are exactly those that are simulated by the tasks considered in the previous section when $Y \equiv B$ and both Y and B are left out of the task.

Since the bistochastic channels are the only channels that admit an active time-reversal [61], the equivalence of inference-symmetry and bistochasticity further connects inference-symmetry and time-reversal invariance, as we will see in the next section.

4.5 More general preparations

Until now, we limited the preparation to the random element of an orthonormal basis. The reason we considered this is that it reflects the simplest way to interact with a system, namely, to couple to one of its non-degenerate observables. But what if our friend tells us that they prepared one out of n, not necessarily orthogonal, states? This situation can also be accounted for using purification.

Say the set of possible initial states states is $\{\psi_i\}_{i=1}^n$, in the d-dimensional Hilbert space A, let U be the unitary evolution and $\{x_j\}_{j=1}^d$ the orthonormal basis of the measurement. Then the prediction probabilities are

$$P_{pre}(x|\psi_i,U) = |\langle x|U|\psi_i\rangle|^2. \tag{46}$$

The postdiction probabilities are again found by Bayes' theorem:

$$P_{post}(\psi_i|x,U) = \frac{P_{pre}(x|\psi_i,U)P(\psi_i)}{P(x)}. \tag{47}$$

Since we are only told the set of possible states, we assume a flat prior over them:

$$P(\psi_i) = \frac{1}{n}, \tag{48}$$

and then the probability for the outcome is

$$P(x) = \sum_{i=1}^n \frac{1}{n}|\langle x|U|\psi_i\rangle|^2 = \mathrm{tr}\left(|x\rangle\langle x|\, U\left[\rho_A\right]\right) \tag{49}$$

where we have defined $\rho_A := \sum_i |\psi_i\rangle\langle\psi_i|/n$. And thus

$$P_{post}(\psi_i|x,U) = \frac{P_{pre}(x|\psi_i,U)}{n\,\mathrm{tr}\left(|x\rangle\langle x|\, U\left[\rho_A\right]\right)}. \tag{50}$$

So the prediction and postdiction probabilities are in general different if the space of possible initial configurations does not represent an orthonormal basis.

We can understand this asymmetry again in terms of implicit knowledge. Let B be a n-dimensional Hilbert space, and $\{b_i\}_{i=1}^n$ an orthonormal base for it. Choose an orthonormal basis $\{a_i\}_{i=1}^d$ on A and find a unitary U_P on $A \otimes B$ that maps:

$$U_P : |a_1\rangle \otimes |b_i\rangle \longmapsto |\psi_i\rangle \otimes |b_i\rangle. \tag{51}$$

Also define the unitary $U' : A \otimes B \to X \otimes B$ given by

$$U' = (U \otimes \mathbb{I}_B) \circ U_P. \tag{52}$$

We can relate the prediction and postdiction probabilities for these two games. For prediction we have:

$$P_{pre}(x|\psi_i,U) = P_{pre}(x|a_1b_i,U') \tag{53}$$

since, for an arbitrary basis $\{y_i\}_{i=1}^n$ of B:

$$\begin{aligned} P_{pre}(x|a_1b_i,U') &= \sum_{j=1}^n \langle xy_j|U'|a_1b_i\rangle \\ &= \sum_{j=1}^n |\langle xy_j|U\otimes\mathbb{I}_B|\psi_ib_i\rangle|^2 \\ &= |\langle x|U|\psi_i\rangle|^2 \cdot \sum_{j=1}^n |\langle y_j|b_i\rangle|^2 \\ P_{pre}(x|a_1b_i,U') &= P_{pre}(x|\psi_i,U), \end{aligned} \tag{54}$$

where we have used the definitions of U' and U_P to obtain the second equality.

For postdiction we have

$$P_{post}(\psi_i|x,U) = \frac{P_{post}(a_1 b_i|x,U')}{P_{post}(a_1|x,U')}, \tag{55}$$

which is entirely analogous with (43). The proof is just a matter of expressing nominator and denominator in terms of probabilities for the original task:

$$\begin{aligned} P_{post}(a_1 b_i|x,U') &= \frac{1}{n}\sum_{j=1}^{n} P_{post}(a_1 b_i|xy_j,U') \\ &= \frac{1}{n}\sum_{j=1}^{n} P_{pre}(xy_j|a_1 b_i,U') \\ &= \frac{1}{n} P_{pre}(x|a_1 b_i,U') \\ P_{post}(a_1 b_i|x,U') &= \frac{1}{n} P_{pre}(x|\psi_i,U) \end{aligned} \tag{56}$$

and so

$$\begin{aligned} P_{post}(a_1|x,U') &= \sum_{i=1}^{n} P_{post}(a_1 b_i|x,U') \\ &= \sum_{i=1}^{n} \frac{1}{n} P_{pre}(x|\psi_i,U) \\ P_{post}(a_1|x,U') &= \mathrm{tr}\left(|x\rangle\langle x|\, U[\rho_A]\right). \end{aligned} \tag{57}$$

Thus, cases where we have more general preparations can be understood in terms of preparations on orthonormal bases. From (55), in analogy with (43), that the preparation of non-orthonormal states contains some implicit information about an ancilla system that allows to prepare the non-orthogonal states. Again, we see that in this purified game, the only way to achieve postdiction probability 1 is to marginalise over b_i, *i.e.* not guess at all.

This method can be further generalised to the preparation of an arbitrary set of density operators by adding a purifying system that gets traced over in the prediction task, and is assigned a flat prior in the postdiction task. The asymmetry again can be understood as an asymmetry in the assumed knowledge in the tasks.

5 Relation between time-reversal and postdiction

We have explored the symmetry between guessing the future and guessing the past. We have seen that in the case of a closed system, the past is quantitatively as uncertain as the future, in the sense that the Born rule can also be used, without modification, to guess a past event based on information about a future event. In the case of a general quantum channel, this symmetry is broken: the prediction and postdiction probabilities are given by different formulas. However, the reason for this is that the implementation of a general quantum channel requires the preparation of a system in a given state, so that knowing a channel was implemented confers information about the input system, breaking the symmetry between prediction and postdiction. Quantum probabilities by themselves have no regard for the direction of time.

In this section, we examine another way in which the quantum probabilities do not distinguish between past and future by examining the more familiar notion of time symmetry: that of time-reversal symmetry. There are two notions of time-reversal, which we could call active and passive, or operational and descriptive. These notions of time-reversal are slightly different from, but closely related to, those that would be applied to a system of differential equations, where one maps solutions to solutions.

In the passive time-reversal of a system, one simply examines the changes in the system in the reverse order, starting from the future and moving towards the past. Applied to the inference tasks we have been considering, passive time-reversal amounts to switching from the prediction task to the postdiction task and *vice versa*. The previous sections have thus been an investigation of passive time-reversal symmetry; we proved that passive time-reversal symmetry is equivalent to inference-symmetry.

Active time-reversal consists instead in finding a process that undoes the change that was brought by a previous transformation. In the context of the inference tasks, active time-reversal consists in considering a new, time-reversed task.

5.1 Time-reversed task with unitary channel

Unitary maps are invertible, and thus for every evolution U of a closed quantum system, there exists a time-reversed evolution given by the adjoint $U^\dagger$.

Definition 4 (Time-reversed task)**.** *Consider a task in which a closed system is prepared in a basis $\{a_i\}_{i=1}^d$ and measured in a basis $\{x_i\}_{i=1}^d$ after undergoing the unitary evolution U. In the time-reversed task, the system is prepared in $\{x_i\}_{i=1}^d$ and measured in $\{a_i\}_{i=1}^d$ after undergoing the evolution $U^\dagger$.*

It follows immediately from the properties of the inner product that

$$P_{pre}(a|x,U^\dagger) = P_{pre}(x|a,U). \tag{58}$$

While this is result is trivial to derive, it is nevertheless profound, as it compounds with the inference symmetry

$$P_{pre}(x|a,U) = P_{post}(a|x,U), \tag{59}$$

to show how little regard the probabilities of closed quantum systems have about the direction of time.

Indeed, given two pure states a and x of the corresponding Hilbert spaces, the *same* quantity

$$|\langle x|U|a\rangle|^2 \tag{60}$$

is the solution for four conceptually distinct taks:

- $P_{pre}(x|a,U)$, for the prediction task with unitary evolution U,
- $P_{post}(a|x,U)$, for the postdiction task with unitary evolution U,
- $P_{pre}(a|x,U^\dagger)$, for the prediction task with time-reversed unitary evolution $U^\dagger$, and
- $P_{post}(x|a,U^\dagger)$, for the postdiction task with time-reversed unitary evolution $U^\dagger$.

The familiar Born rule, which is normally thought to apply only to the first case, actually applies to all four of these cases. In each one of them, x can be either the future or the past event, and can either be the known or unknown in the scenario.

$$\begin{tikzcd}
P_{pre}(x|a,U) \arrow[r, leftrightarrow, "\text{active}"] \arrow[d, leftrightarrow, "\text{passive}"'] \arrow[dr, phantom, "|\langle x|U|a\rangle|^2"] & P_{pre}(a|x,U^\dagger) \arrow[d, leftrightarrow, "\text{passive}"] \\
P_{post}(a|x,U) \arrow[r, leftrightarrow, "\text{active}"'] & P_{post}(x|a,U^\dagger)
\end{tikzcd}$$

Let us now consider the case of open systems and study the time-reversed version of the tasks examined in section 3. If we neglect the outcome of the measurement on B, we can write:

$$P_{pre}(a|xy,U^\dagger) = \sum_{i=1}^{d_B} P_{pre}(ab_i|xy,U^\dagger) \tag{61}$$

$$= \sum_{i=1}^{d_B} P_{pre}(xy|ab_i,U) \tag{62}$$

$$P_{pre}(a|xy,U^\dagger) = P_{post}(a|xy,U), \tag{63}$$

where we have used the time-reversal symmetry (58) in the second equality, and (11) in the third. This equation relates the probability of the prediction task to the probability of the time-reversed postdiction task. In both cases, we are calculating the probability of event a based on knowledge of event xy. However, in one case xy is to the past of a and in the other case the opposite is true. In a similar fashion, we can derive the following equations:

$$P_{pre}(ab|x,U^\dagger) = P_{post}(ab|x,U), \tag{64}$$

$$P_{post}(xy|a,U^\dagger) = P_{pre}(xy|a,U), \tag{65}$$

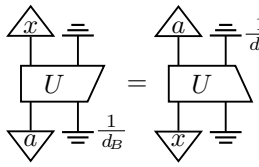
$$P_{post}(x|ab,U^\dagger) = P_{pre}(x|ab,U), \tag{66}$$

for when only one side is being ignored, as well as

$$P_{pre}(a|x,U^\dagger) = P_{post}(a|x,U), \tag{67}$$

$$P_{post}(x|a,U^\dagger) = P_{pre}(x|a,U), \tag{68}$$

for when data is being ignored on both sides. Recall the discussion in section 3 about the normalisation of the identity that is required to calculate the quantities on the right hand side. The same applies here, which reinforces the lesson: the direction of time is irrelevant to the normalisation of the identity. What matters is the direction of inference:

$P_{pre}(x|a,U) =$  $= P_{post}(x|a,U^\dagger).$ (69)

The equation above[1] refers to inference tasks in which we have information about the system A and want to guess something about the system X. It does not matter if A is in the past or the future of X. The answer is the same. There is no way of telling which way the arrow of time is pointing.

5.2 Time-reversed task with quantum channel

The natural candidate for the time-reversed task is the one with the adjoint $\Phi^\dagger$ of the quantum channel Φ. In fact, Chiribella *et. al.* proved that this is essentially the unique way of defining the active time-reversal of a quantum channel [61, 62]. We must distinguish two cases, as the adjoint of a CPTP map may or may not be a CPTP map.

The adjoint $\Phi^\dagger$ is trace-preserving if and only if Φ is unit preserving since, by the definition of the adjoint, for all $\rho \in \mathcal{L}(X)$:

$$\operatorname{tr}\Phi^\dagger[\rho] = \operatorname{tr}\mathbb{I}_A \circ \Phi^\dagger[\rho] = \operatorname{tr}\Phi[\mathbb{I}_A] \circ \rho. \tag{70}$$

Furthermore, we have seen that the bistochastic channels are exactly the inference symmetric channels. In analogy with the discussion above on the closed systems, for a bistochastic channel Φ, the quantity

$$\operatorname{tr}|x\rangle\langle x| \circ \Phi[|a\rangle\langle a|], \tag{71}$$

given by the generalised Born rule, yields the numerical value of four *a priori* conceptually distinct quantities:

- $P_{pre}(x|a,\Phi)$, for the prediction task with bistochastic channel Φ,
- $P_{post}(a|x,\Phi)$, for the postdiction task with bistochastic channel Φ,

[1] In the diagrammatic calculus, flipping a diagram along a horizontal axis denotes taking the hermitian conjugate.

- $P_{pre}(a|x, \Phi^\dagger)$, for the prediction task with time-reversed bistochastic channel $\Phi^\dagger$ and

- $P_{post}(x|a, \Phi^\dagger)$, for the postdiction task with time-reversed bistochastic channel $\Phi^\dagger$.

Thus postdiction-symmetry is intimately linked with time-reversal invariance even in the context of open quantum systems.

In general, however, the adjoint of a CPTP map might fail to be trace non-increasing. So it might not only fail to represent a quantum channel, but may not even be a quantum map, *i.e.* part of an operation. In this case, an active time-reversal of the corresponding quantum channel does not exist, as has been recently shown in [61, 62].

5.3 Quantum channels towards the past

A quantum channel might not have an active time-reversed version. Nevertheless, we can learn something striking by looking at the time-reversed task of a *purification* of this quantum channel: CPTP maps can describe *postdictions* in situations in which part of the future data is left implicit and fixed. Consider the time-reversed version of the purified task in section 4.3. The system is prepared in some state corresponding to the basis $\{x_i y_j\}$ and measured on some basis $\{a_i b_j\}$ after undergoing the transformation described by ${U_\Phi}^\dagger$. Using the time-symmetry of unitary open systems (66), observe that

$$P_{pre}(x|a, \Phi) = P_{pre}(x|ab, U_\Phi) = P_{post}(x|ab, U_\Phi^\dagger). \tag{72}$$

That is, the quantity

$$\operatorname{tr} |x\rangle\langle x| \, \Phi[|a\rangle\langle a|] \tag{73}$$

is the solution to two different tasks:

- $P_{pre}(x|ab, U_\Phi)$ for the prediction task with unitary U_Φ, and

- $P_{post}(x|ab, U_\Phi^\dagger)$ for the time-reversed postdiction task with $U_\Phi^\dagger$.

Thus one can take the quantity (73) to relate to an inference towards the past, in a situation in which part of the future events are only implicitly described. This furthers the argument that the inference-asymmetry of the CPTP maps is not an asymmetry related to an intrinsic direction of time in the details of a quantum process, but in an asymmetry in the data about a system. The input of a CPTP map does not necessarily lie to the past of the output, it can also lie to its future.

This last insight will play a major role in dispelling a source of confusion regarding the time-orientation of quantum phenomena.

6 The arrow of inference

In this section, we discuss two closely related asymmetric aspects of the operational formalism, encapsulated by two maxims: "there exists a unique deterministic effect" and "no signalling from the future". These expressions reflect mathematical properties of the theory that are often taken as evidence of an asymmetry between the past and the future. Here we show that these properties reflect the asymmetry due to the arrow of inference, which is in principle independent of the arrow of time.

6.1 "There exists a unique deterministic effect"

As a mathematical statement, given the definitions, this maxim is correct. A quantum map is deemed *deterministic*, or *causal*, if it is trace preserving. An *effect* is a quantum map from a Hilbert space to the trivial Hilbert space. Thus a *deterministic effect* is an effect that is trace-preserving. It is easy to see that the only such effect is the one that maps a state to its trace, *i.e.* the discard operator. In this precise sense the discard is the only deterministic effect.

As a statement of the irreducible uncertainty of quantum phenomena, the maxim is also correct: the only way to be certain about a prediction in all cases is not to guess at all. In the context of prediction tasks, discarding means not trying to predict anything about the system and is represented by the identity operator as an effect (see section 3). Not guessing is the only way to always guess right, and this is the physical content of "there exists a unique deterministic effect."

The uniqueness of the deterministic effect is sometimes understood as the conservation of probability, a mere aspect of inference (see for example [63, 35]). However, the maxim is sometimes taken to signify something fundamental about the distinction between the past and the future [44, 64], and this is not correct. In fact, we can define by analogy what it means to discard something in the past. If discarding in the future means marginalising prediction probabilities, then discarding in the past means marginalising postdiction probabilities. Discarding in the past is also represented by the identity operator when computing probabilities—this time in the *input* side, as we have seen at the end of section 3. Indeed, looking at the formula (35) for postdiction with a general quantum channel, we see that the only way to have probability 1 is to discard the system, *i.e.* by not guessing at all.

Thus, the uniqueness of the discard operator is not a consequence of the time-orientation of quantum phenomena. It is instead a manifestation of thoroughly time-symmetric quantum indeterminism in the context of prediction.

6.2 "No signalling from the future"

This maxim means that the probabilities of the outcomes of a quantum operation do not depend on the nature of a later operation. To lay down some notation, we reproduce the standard proof of this property, which is an immediate consequence of the equations (22) and (26). In the next subsection we comment on the physical reasons for this and show that it too is in fact a property of the arrow of inference, not a property of time.

Suppose a system starts in a state ρ, and the operation $\mathcal{E}^{A\to D} = \{E_x\}$ is applied. The quantity

$$P_{pre}(x|\rho, \mathcal{E}) = \operatorname{tr} E_x[\rho] \tag{74}$$

is the probability of observing outcome x.

Suppose instead that the operation $\mathcal{E}$ is immediately followed by another operation $\mathcal{F}^{D\to Z} = \{F_y\}$. The probability for the outcomes x and y is given definition by (26):

$$P_{pre}(xy|\rho, \mathcal{F}\circ\mathcal{E}) = \operatorname{tr} F_y\left[E_x[\rho]\right] \tag{75}$$

An immediate application of probability theory and the completeness equation (22) yields

$$P_{pre}(x|\rho, \mathcal{F}\circ\mathcal{E}) = \sum_y \operatorname{tr} F_y\left[E_x[\rho]\right] = \operatorname{tr} E_x[\rho], \tag{76}$$

and we thus have the identity

$$P_{pre}(x|\rho, \mathcal{F}\circ\mathcal{E}) = P_{pre}(x|\rho, \mathcal{E}). \tag{77}$$

The probabilities of the outcomes of the first operation are independent on the nature of the second operation.[2] The case in which the operation $\mathcal{F}$ is performed and the case in which it is not are indistinguishable by looking only at what happens at $\mathcal{E}$. A future experimenter cannot affect the statistics in the present by manipulating a system in the future, hence "no signalling from the future."

What about the opposite direction? By linearity and (25) we have

$$P_{pre}(y|\rho, \mathcal{F}\circ\mathcal{E}) = P_{pre}(y|\mathcal{E}[\rho], \mathcal{F}), \tag{78}$$

so that the probability of y clearly depends on the first operation $\mathcal{E}$.

Thus the future operation does not affect the probabilities of past events, while the past operation affects the probability of future events. One seems compelled to conclude that this reveals a time-asymmetry in quantum theory. But this conclusion is too quick. In previous sections, we have encountered various situations in which asymmetric aspects of quantum theory should not be ascribed to a difference between past and future, but to the directionality of inference and an asymmetry of the data. We have also seen that the operational formulations assume that the arrow of time and the arrow of inference point the same way. We should be cautious.

6.3 "No signalling from the further unknown"

The calculation above in fact shows that the future operation does not affect the *prediction* probabilities of events in its past, while the past operation affects the *prediction* probabilities of events in its future. In other words, if somebody in the past of both operations is trying to guess what would be the outcome of the first operation, they can safely discard any information about the second operation that might be available at that time.[3] Put another way: an event further away from the data does not affect the prediction probabilities closer to the data.

A similar statement holds also when the arrow of inference points toward decreasing time, *i.e.* when doing postdiction. As we have already seen in section 5.2, one can understand a CPTP map as a shorthand to calculate a postdiction probability when a future event is implicitly considered fixed. The same is generally true of *any* operation: it can serve as a shorthand to aid the calculation of postdiction probabilities. Then (77), understood as a statement about postdiction probabilities, tells us that what happened in the further past does not affect our ability to infer what happens in the closer past. When predicting, an event does not affect prediction probabilities about an event in its past. When postdicting, an event does not affect postdiction probabilities of events in its future. In both cases, there is no signalling from the further unknown.

As an illustration, let us consider a purification of the two operations $\mathcal{E}^{A\to D}$ and $\mathcal{F}^{D\to Z}$ above. Let us assume for simplicity that there is no need to discard a part of the output system (by handling the more general case we would reach the same conclusions but with more typographical effort). Then there exists Hilbert spaces B, C, X and Y, unitaries $U_{\mathcal{E}} : A\otimes B \to X\otimes D$ and $U_{\mathcal{F}} : D\otimes C \to Y\otimes Z$, pure states $|b\rangle \in B$ and $|c\rangle \in C$ as well as orthonormal bases for X and Y labelled by x and y respectively,

[2] Note that the no-signalling from the future property (77) can also be seen as a motivation for the definition (26) for the probabilities of the outcomes of sequential operations in the first place. Indeed by linearity, these can be rewritten as:

$$\operatorname{tr} F_y\left[E_x[\rho]\right] = \operatorname{tr}\left[F_y\left(\frac{E_x[\rho]}{\operatorname{tr} E_x[\rho]}\right)\right]\cdot \operatorname{tr} E_x[\rho] \tag{79}$$

so that, by setting $\rho_x = E_x[\rho]/\operatorname{tr} E_x[\rho]$ one can write:

$$P_{pre}(xy|\rho, \mathcal{F}\circ\mathcal{E}) = P_{pre}(y|\rho_x, \mathcal{F})\cdot P(x|\rho, \mathcal{E}). \tag{80}$$

[3] To be sure, having knowledge of the outcome of the second operation *does* provide an advantage in guessing the outcome of the first one, since

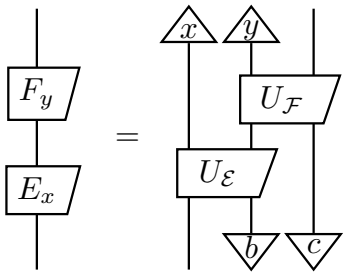
$$P(x|y, \rho, \mathcal{F}\circ\mathcal{E}) = \operatorname{tr} F_y[E_x[\rho]]/\operatorname{tr} F_y[\mathcal{E}[\rho]]. \tag{81}$$

This data however is unavailable to somebody sitting in the past of both operations.

such that for all $\rho \in \mathcal{L}(A)$ and $\sigma \in \mathcal{L}(D)$:

$$E_x[\rho] = \mathrm{tr}_X\, |x\rangle\langle x| \circ U_{\mathcal{E}}[\rho \otimes |b\rangle\langle b|], \tag{82}$$
$$F_y[\sigma] = \mathrm{tr}_Y\, |y\rangle\langle y| \circ U_{\mathcal{F}}[\sigma \otimes |c\rangle\langle c|], \tag{83}$$

meaning that:



. (84)

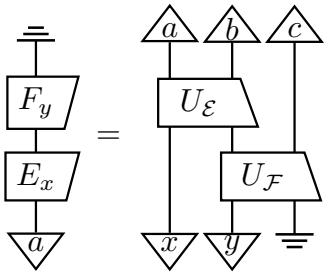
We can write down the prediction probabilities for the purified task in terms of the operations:

$$P_{pre}(xy|abc, U_{\mathcal{F}}U_{\mathcal{E}}) = \mathrm{tr}\, F_y[E_x[a]]. \tag{85}$$

Now we can use the property (66) of time-reversed unitary tasks to also write

$$P_{post}(xy|abc, U_{\mathcal{E}}^{\dagger}U_{\mathcal{F}}^{\dagger}) = \mathrm{tr}\, F_y[E_x[a]]. \tag{86}$$

Thus the quantities $\mathrm{tr}\, F_y[E_x[a]]$ are also postdiction probabilities for a different scenario, in which *future* events b and c are held fixed, and this knowledge is used to guess something happening to their past (x and y):



. (87)

In this case, the operation $\mathcal{F}$ contains information about something that happens *earlier* in the system, namely, the interaction $U_{\mathcal{F}}^{\dagger}$ between the subsystems Y and Z, and the outcome of the measurement of C. All of this information is irrelevant when postdicting only the preparation on X:

$$\sum_y P_{post}(xy|abc, U_{\mathcal{E}}^{\dagger}U_{\mathcal{F}}^{\dagger}) = P_{post}(x|abc, U_{\mathcal{E}}^{\dagger}). \tag{88}$$

Thus, when postdicting, the outcome and nature of the earlier operation is irrelevant to the probabilities for the outcome of the later operation.

6.4 Why we *can* signal from the past

We have seen that the operation $\mathcal{E}$ affects the prediction probabilities for the outcome of $\mathcal{F}$. How is this related to the notion of *signalling*? Let's imagine two parties, one located where $\mathcal{E}$ takes place and one where $\mathcal{F}$ takes place. Let's call them Eve and Fred. How can Eve send a signal to Fred? The barebone scenario for signalling to the future is that Eve's operation is simply a state preparation of a qubit, and Fred's is a projective measurement. Eve, who knows the basis on which Fred will measure, *chooses* one of the two states so that she fully determines the outcome of Fred's operation. Like this, she can send one bit of information.

Notice, however, that Eve's ability to choose is crucial to this protocol. If she cannot pick what state to send to Fred, she cannot send a message to him. All she can do is try to predict what will be the outcome of Fred's measurement, once—and if—she knows what state she sent him. How does somebody prepare a system in a specific state? In practice, this is done by subjecting the system to a maximal test and discarding the systems yielding unwanted results [65]. Alternatively, one can apply a unitary transformation to the system, conditional on the outcome of the maximal test. Both of these procedures require an increase in the entropy of the universe, as the first involves picking and choosing [53, 66] and the second is an instance of erasure so Landauer's principle applies [67], a rigorous proof of which just appeared [68]. Thus, signalling is a concept beyond single quantum transition probabilities and has its origins in the thermodynamic arrow of time.

We are time oriented creatures, we know more about the past than about the future, we mostly try to guess the future. Quantum probabilities do not care if you are making guesses about the future or the past. They are about predicting what is unknown from what is known. The difference between what is known and what is unknown is at the origin of the time asymmetric maxims of the operational formulations. The first maxim arises from the fact that postdiction scenarios are rarer than prediction ones in practice. If we want to learn about the past, we find there are plenty of records about past events in the present. The existence of traces is not a property of the probabilities of individual quantum systems, so the fact that we rarely have to guess about the past the same way we have to guess about the future is no evidence of a time asymmetry of the physics of quanta.

The grip of the second maxim on the community is more subtle. It rests on the notion of signalling, which is itself tightly linked with ideas of causation and agency, concepts we have strong intuitions for and rely on daily. We humans make choices and these choices influence our future (not our past). The same goes for a lot of systems around us: when my laptop suddenly "decides" to break, it will affect my ability to finish a future paper (not a past one). Is this time-orientation a direct consequence of some time asymmetry in quantum phenomena? Hopefully, by now, the answer is clear. Is causation a fundamental

property of the world in some other way? This question has received surprisingly little attention from the physics community at large [1]. This question needs to be addressed carefully if one hopes of extending these operational formalisms to probe physics outside laboratories. The notion of causes always preceding effects is strongly related to notions of agency. And agency is a perspectival property, stemming from a partial description of systems and the presence of an entropy gradient [53, 66]. One should be cautious in extending notions of time oriented causation all the way to elementary physics.

7 Time orientation of other formalisms

Before concluding the paper, we add some comments regarding the time orientation in other formalisms.

In the de Broglie-Bohm interpretation [69, 70], the particle and the pilot-wave obey time-reversal invariant dynamics. The Everettian wavefunction [71, 72] also obeys time-reversal invariant dynamics. Branching towards the future is interpreted as a past low entropy condition: systems were uncorrelated in the past, hence the vanishing of von-Neumann relative entropy. Alternatively, it can be interpreted as a perspectival aspect of the Everettian relative state determined by an observation in the present.

The Copenhagen quantum state is correlated with past observations, not future ones, but the state is, of course, unobservable and its ontological status is debated. Its empirical content is given by the probabilities it allows to calculate. If we ask time-symmetric questions, the probabilities we obtain are time-symmetric [38]. The state is assumed to be correlated with past events because we are using past data to infer about the future. If we want to guess about the past, we could just as well use a quantum state correlated with a future event [39, 42, 73].

In QBism [47], quantum theory is interpreted as a means to aid decision making, allowing an agent to calculate the probabilities of the consequences of their interactions with the world. Because decision-making agents play a central role in bringing about the world according to the QBists [74], the ontology of the theory is fundamentally time oriented.

Lab measurements generally involve decoherence and amplification of a microscopic phenomenon to the macroscopic realm, both of which rely on the entropy gradient. Views of quantum theory that insist that the only real events are of this kind are therefore time-oriented, even though the probabilities themselves might be time-symmetric. These views also imply that there is no interpretation of the theory outside of the macroscopic approximation.

According to the relational interpretation of quantum mechanics (RQM) [31, 75], facts happen at every interaction between any two systems, but the facts are relative to the systems involved in the interaction. The quantum state only plays a computational role in this interpretation. Decoherence comes into play to 'stabilise' a fact [76], so that one might ignore its relational nature, which is manifest in interference effects. Decoherence requires information loss and an increase in entropy. Hence RQM is a time-symmetric formulation of quantum theory, but the dynamics of *relative* facts is time symmetric while the dynamics of *stable* facts is time oriented.

Conclusion

Quantum theory is not about predicting the future, it is about time-symmetric conditional probabilities relating events. The directionality internal to the theory is the arrow of inference, the difference between known and unknown.

There are *formulations* of quantum theory that break time reversal symmetry and use time oriented theoretical notions. These either refer to non-observable entities, or to assumptions about the time orientation of inference problems, or to the entropic time orientation of decoherence. Examples of such formulations are provided by the use of a quantum state determined by *past* interactions in the Copenhagen interpretation, and the use of the quantum operations described above. The time orientation of operations is due to them being high-level notions with a built-in assumption about time asymmetric capabilities of the experimentalists.

The time orientation of the formalisms we use is determined by the common boundary conditions we set for the physical processes we study: they come from the assumptions about the agent interacting with the system and the conditions she imposes on it. The agent is not directly modelled in the theory and is instead represented by the inferential boundary conditions and choice of operations, the exogenous variables of [77]. Time orientation is in this way external to the elementary quantum process being modelled. It can, in principle, be entirely accounted for at the level of statistical mechanics as a consequence of the existence of an entropy gradient, namely past low entropy [78, 2, 53, 66].

Some authors [44, 38] have built time-symmetric theories to replace quantum theory. However, the results of sections 2 and 5 show that quantum theory is already time-symmetric as it is. The transition probabilities calculated with quantum theory are blind to the direction of time. The probabilities of closed quantum systems are inference symmetric and time-reversal invariant. Thus, when accounting for all the relevant degrees of freedom, the predictions of quantum theory are thoroughly time-agnostic. The probabilities of open quantum systems are in general neither inference symmetric nor time-reversal invariant. We have shown that the asymmetry between prediction and postdiction in this case is only a con-

sequence of treating the two problems asymmetrically, by assuming more knowledge in one case than in the other. Both inference asymmetry and the failure of time-reversal invariance of quantum channels can be understood in the same terms. This asymmetry[4] is not intrinsic to the mechanical theory, but is, rather, an asymmetry of the questions we humans ask using the theory.

Acknowledgments

The authors thank Giulio Chiribella and his QICI group, Lucien Hardy, Laurent Freidel, Flaminia Giacomini, Francesca Vidotto, and especially Marios Christodoulou for critical and helpful discussions on early versions on this work. The authors also thank the three anonymous referees for thorough and helpful critical feedback.

This work was made possible through the support of the FQXi Grant FQXi-RFP-1818 and of the ID# 61466 grant from the John Templeton Foundation, as part of the "The Quantum Information Structure of Spacetime (QISS)" Project (qiss.fr). The work of P.D. is partially supported by the grant 2018-190485 (5881) of the Foundational Questions Institute and the Fetzer Franklin Fund.

The diagrammatic formulas were realised with TikZit (tikzit.github.io).

[4] This situation is reminiscent of the one in classical statistical mechanics, in which one can prove that entropy increases *both* towards the past *and* towards the future of the given initial state [1, 77]. In both situations we use time-symmetrical mechanical laws to extrapolate from a state of limited knowledge to a state at other times. This results in even more limited knowledge. The fact that entropy was lower in the past, is a fact external to the laws of mechanics. And the fact that we are mostly interested in extrapolating towards the future is in turn a consequence of the low entropy in the past.

References

[1] Huw Price. "Time's arrow & Archimedes' point: New directions for the physics of time". Oxford Paperbacks. Oxford University Press (1997). DOI: 10.1093/acprof:oso/9780195117981.001.0001.

[2] Carlo Rovelli. "Memory and entropy" (2020) eprint: arXiv:2003.06687.

[3] Michael A Nielsen and Isaac L Chuang. "Quantum computation and quantum information". Cambridge University Press (2000). DOI: 10.1017/CBO9780511976667.

[4] Giacomo Mauro D'Ariano, Giulio Chiribella, and Paolo Perinotti. "Quantum Theory from First Principles: An Informational Approach". Cambridge University PressCambridge (2017). DOI: 10.1017/9781107338340.

[5] Bob Coecke and Aleks Kissinger. "Picturing quantum processes". Cambridge University Press (2017). DOI: 10.1017/9781316219317.

[6] J. S. Bell. "On the Einstein Podolsky Rosen paradox". Physics Physique Fizika **1**, 195–200 (1964). DOI: 10.1103/physicsphysiquefizika.1.195.

[7] John S. Bell. "On the Problem of Hidden Variables in Quantum Mechanics". Reviews of Modern Physics **38**, 447–452 (1966). DOI: 10.1103/revmodphys.38.447.

[8] W K Wootters and W H Zurek. "A single quantum cannot be cloned". Nature **299**, 802–803 (1982). DOI: 10.1038/299802a0.

[9] Peter W. Shor. "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer". SIAM Journal on Computing **26**, 1484–1509 (1997). DOI: 10.1137/s0097539795293172. eprint: arXiv:quant-ph/9508027.

[10] Artur K Ekert. "Quantum cryptography based on Bell's theorem". Physical Review Letters **67**, 661–663 (1991). DOI: 10.1103/PhysRevLett.67.661.

[11] Lov K. Grover. "A fast quantum mechanical algorithm for database search". Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96 (1996). DOI: 10.1145/237814.237866. eprint: arXiv:quant-ph/9605043.

[12] Charles H. Bennett and Gilles Brassard. "Quantum cryptography: Public key distribution and coin tossing". Theoretical Computer Science **560**, 7–11 (2014). DOI: 10.1016/j.tcs.2014.05.025. eprint: arXiv:2003.06557.

[13] D Dieks. "Communication by EPR devices". Physics Letters A **92**, 271–272 (1982). DOI: 10.1016/0375-9601(82)90084-6.

[14] Charles H Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K Wootters. "Teleporting an unknown quantum state via dual classical and Einstein-

Podolsky-Rosen channels". Physical Review Letters **70**, 1895–1899 (1993). DOI: 10.1103/PhysRevLett.70.1895.

[15] A.S. Holevo. "The capacity of the quantum channel with general signal states". IEEE Transactions on Information Theory **44**, 269–273 (1998). DOI: 10.1109/18.651037. eprint: arXiv:quant-ph/9611023.

[16] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron. "Quantum computations without definite causal structure". Physical Review A **88**, 022318 (2013). DOI: 10.1103/physreva.88.022318. eprint: arXiv:0912.0195.

[17] Flaminia Giacomini, Esteban Castro-Ruiz, and Časlav Brukner. "Quantum mechanics and the covariance of physical laws in quantum reference frames". Nature Communications **10**, 494 (2019). DOI: 10.1038/s41467-018-08155-0. eprint: arXiv:1712.07207.

[18] Sougato Bose, Anupam Mazumdar, Gavin W. Morley, Hendrik Ulbricht, Marko Toroš, Mauro Paternostro, Andrew Geraci, Peter Barker, M. S. Kim, and Gerard Milburn. "A Spin Entanglement Witness for Quantum Gravity". Physical Review Letters **119**, 240401 (2017). DOI: 10.1103/physrevlett.119.240401. eprint: arXiv:1707.06050.

[19] Chiara Marletto and Vlatko Vedral. "Gravitationally-induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity". Physical Review Letters **119**, 240402 (2017). DOI: 10.1103/physrevlett.119.240402. eprint: arXiv:1707.06036.

[20] Kok-Wei Bong, Aníbal Utreras-Alarcón, Farzad Ghafari, Yeong-Cherng Liang, Nora Tischler, Eric G. Cavalcanti, Geoff J. Pryde, and Howard M. Wiseman. "A strong no-go theorem on the Wigner's friend paradox". Nature PhysicsPages 1–7 (2020). DOI: 10.1038/s41567-020-0990-x.

[21] Lucien Hardy. "Quantum Theory From Five Reasonable Axioms" (2001) eprint: arXiv:quant-ph/0101012.

[22] Borivoje Dakic and Časlav Brukner. "Quantum theory and beyond: Is entanglement special?" (2009) eprint: arXiv:0911.0695.

[23] Lluís Masanes and Markus P. Müller. "A derivation of quantum theory from physical requirements". New Journal of Physics **13**, 063001 (2011). DOI: 10.1088/1367-2630/13/6/063001.

[24] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Informational derivation of Quantum Theory". Physical Review A **84**, 012311 (2011). DOI: 10.1103/physreva.84.012311. eprint: arXiv:1011.6451.

[25] Lucien Hardy. "Reconstructing quantum theory" (2013) eprint: arXiv:1303.1538.

[26] Philipp A. Höhn. "Toolbox for reconstructing quantum theory from rules on information acquisition". Quantum **1**, 38 (2017). DOI: 10.22331/q-2017-12-14-38. eprint: arXiv:1412.8323.

[27] Philipp A. Höhn and Christopher Wever. "Quantum theory from questions". Physical Review A **95**, 012102 (2017). DOI: 10.1103/PhysRevA.95.012102. eprint: arXiv:1511.01130.

[28] John H. Selby, Carlo Maria Scandolo, and Bob Coecke. "Reconstructing quantum theory from diagrammatic postulates". Quantum **5**, 445 (2021). DOI: 10.22331/q-2021-04-28-445. eprint: arXiv:1802.00367.

[29] Ding Jia. "Quantum from principles without assuming definite causal structure". Physical Review A **98**, 032112 (2018). DOI: 10.1103/physreva.98.032112. eprint: arXiv:1808.00898.

[30] Robert Oeckl. "A local and operational framework for the foundations of physics". Advances in Theoretical and Mathematical Physics **23**, 437–592 (2019). DOI: 10.4310/ATMP.2019.v23.n2.a4. eprint: arXiv:1610.09052.

[31] Carlo Rovelli. "Relational Quantum Mechanics". International Journal of Theoretical Physics **35**, 1637–1678 (1996). DOI: 10.1007/BF02302261. eprint: arXiv:quant-ph/9609002.

[32] Christopher A. Fuchs. "Quantum Mechanics as Quantum Information (and only a little more)". eprint: arXiv:quant-ph/0205039.

[33] "The Quantum Information Structure of Spacetime". http://www.qiss.fr/.

[34] Andrea Di Biagio. "Can we think time-symmetrically about causation?". http://pirsa.org/20120022 (2020). PIRSA:20120022.

[35] David Schmid, John H. Selby, and Robert W. Spekkens. "Unscrambling the omelette of causation and inference: The framework of causal-inferential theories" (2020) eprint: arXiv:2009.03297.

[36] Lucien Hardy. "Time Symmetry in Operational Theories" (2021) eprint: arXiv:2104.00071.

[37] Albert Einstein, Richard C Tolman, and Boris Podolsky. "Knowledge of Past and Future in Quantum Mechanics". Physical Review **37**, 780–781 (1931). DOI: 10.1103/PhysRev.37.780.

[38] Yakir Aharonov, Peter G Bergmann, and Joel L Lebowitz. "Time Symmetry in the Quantum Process of Measurement". Physical Review **134**, B1410–B1416 (1964). DOI: 10.1103/PhysRev.134.B1410.

[39] Satosi Watanabe. "Symmetry of physical laws. Part III. Prediction and retrodiction". Reviews of Modern Physics **27**, 179–186 (1955). DOI: 10.1103/RevModPhys.27.179.

[40] Stephen M. Barnett, David T. Pegg, and John Jeffers. "Bayes' theorem and quantum retrodiction". Journal of Modern Optics **47**, 1779–

1789 (2000). DOI: 10.1080/09500340008232431. eprint: arXiv:quant-ph/0106139.

[41] David T. Pegg, Stephen M. Barnett, and John Jeffers. "Quantum retrodiction in open systems". Physical Review A **66**, 022106 (2002). DOI: 10.1103/physreva.66.022106. eprint: arXiv:quant-ph/0208082.

[42] Fiona C. Speirits, Matthias Sonnleitner, and Stephen M. Barnett. "From retrodiction to Bayesian quantum imaging". Journal of Optics **19**, 044001 (2017). DOI: 10.1088/2040-8986/aa5ccf.

[43] Matthew Leifer and Matthew Pusey. "Is a time symmetric interpretation of quantum theory possible without retrocausality?". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **473**, 20160607 (2017). DOI: 10.1103/physreva.88.052130. eprint: arXiv:1607.07871.

[44] Ognyan Oreshkov and Nicolas J Cerf. "Operational formulation of time reversal in quantum theory". Nature PhysicsPages 853–858 (2015). DOI: 10.1038/nphys3414. eprint: arXiv:1507.07745.

[45] Robert Oeckl. "What is Quantum Theory?, QISS Virtual Seminar". https://www.youtube.com/watch?v=_Fwkh9cBMZw&t=5095s (2020).

[46] Edwin Thompson Jaynes. "Probability in Quantum Theory". In Wojciech Hubert Zurek, editor, Complexity, Entropy And The Physics Of Information Page 381. Addison-Wesley Pub. Co (1990).

[47] Christopher A. Fuchs and Blake C. Stacey. "QBism: Quantum theory as a hero's handbook" (2016) eprint: arXiv:1612.07308.

[48] M. S. Leifer and R. W. Spekkens. "Towards a Formulation of Quantum Theory as a Causally Neutral Theory of Bayesian Inference". Physical Review A **88**, 052130 (2013). DOI: 10.1103/physreva.88.052130. eprint: arXiv:1107.5849.

[49] Stephen M. Barnett. "Quantum Retrodiction". In Erika Andersson and Patrik Öhberg, editors, Quantum Information and Coherence Pages 1–30. Springer International Publishing (2014).

[50] Dov Fields, Abdelali Sajia, and János A. Bergou. "Quantum retrodiction made fully symmetric" (2020) eprint: arXiv:2006.15692.

[51] Yuan Yuan, Zhibo Hou, Kang-Da Wu, Guo-Yong Xiang, Chuan-Feng Li, and Guang-Can Guo. "Experimental retrodiction of trajectories of single photons in double interferometers". Physical Review A **97**, 062115 (2018). DOI: 10.1103/PhysRevA.97.062115.

[52] Massimiliano Rossi, David Mason, Junxin Chen, and Albert Schliesser. "Observing and Verifying the Quantum Trajectory of a Mechanical Resonator". Physical Review Letters **123**, 163601 (2019). DOI: 10.1103/PhysRevLett.123.163601.

[53] Carlo Rovelli. "Agency in physics". eprint: arXiv:2007.05300.

[54] W. Forrest Stinespring. "Positive functions on C*-algebras". Proceedings of the American Mathematical Society **6**, 211–216 (1955). DOI: 10.2307/2032342.

[55] Masanao Ozawa. "Quantum measuring processes of continuous observables". Journal of Mathematical Physics **25** (1984). DOI: 10.1063/1.526000.

[56] L. J. Landau and R. F. Streater. "On Birkhoff's theorem for doubly stochastic completely positive maps of matrix algebras". Linear Algebra and its Applications **193**, 107–127 (1993). DOI: 10.1016/0024-3795(93)90274-r.

[57] Christian B. Mendl and Michael M. Wolf. "Unital Quantum Channels – Convex Structure and Revivals of Birkhoff's Theorem". Communications in Mathematical Physics **289**, 1057–1086 (2009). DOI: 10.1007/s00220-009-0824-2. eprint: arXiv:0806.2820.

[58] Giulio Chiribella and Carlo Maria Scandolo. "Microcanonical thermodynamics in general physical theories". New Journal of Physics **19**, 123043 (2017). DOI: 10.1088/1367-2630/aa91c7. eprint: arXiv:1608.04460.

[59] Alexander Streltsov, Hermann Kampermann, Sabine Wölk, Manuel Gessner, and Dagmar Bruß. "Maximal Coherence and the Resource Theory of Purity". New Journal of Physics **20**, 053058 (2018). DOI: 10.1088/1367-2630/aac484. eprint: arXiv:1612.07570.

[60] Peter Shor. "The structure of unital maps and the asymptotic quantum Birkhoff conjecture". https://youtu.be/bA9PzOAKycY?t=824 (2010).

[61] Giulio Chiribella, Erik Aurell, and Karol Życzkowski. "Symmetries of quantum evolutions". Physical Review Research **3**, 033028 (2021). DOI: 10.1103/PhysRevResearch.3.033028. eprint: arXiv:2101.04962.

[62] Giulio Chiribella and Zixuan Liu. "Quantum operations with indefinite time direction". eprint: arXiv:2012.03859.

[63] Jonathan Barrett. "Information processing in generalized probabilistic theories". Physical Review A **75**, 032304 (2007). DOI: 10.1103/physreva.75.032304. eprint: arXiv:quant-ph/0508211.

[64] Bob Coecke, Stefano Gogioso, and John H. Selby. "The time-reverse of any causal theory is eternal noise" (2017) eprint: arXiv:1711.05511.

[65] Asher Peres. "Quantum theory: concepts and methods". Volume 57. Springer Science & Business Media (2002). DOI: 10.1007/0-306-47120-5.

[66] G. J. Milburn and S. Shrapnel. "Physical grounds for causal perspectivalism" (2020) eprint: arXiv:2009.04121.

[67] Rolf Landauer. "Irreversibility and Heat Generation in the Computing Process". IBM Journal of Research and Development **5**, 261–269 (1961). DOI: 10.1147/rd.53.0183.

[68] Wayne C. Myrvold. "Shakin' All Over: Proving Landauer's Principle without neglect of fluctuations" (2020) eprint: arXiv:2007.11748.

[69] David Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II". Physical Review **85**, 180–193 (1952). DOI: 10.1103/physrev.85.180.

[70] Sheldon Goldstein. "Bohmian Mechanics". In Edward N. Zalta, editor, The Stanford Encyclopedia of Philosophy . Metaphysics Research Lab, Stanford Universitysummer 2017 edition (2017).

[71] Simon Saunders, Jonathan Barrett, Adrian Kent, and David Wallace, editors. "Many Worlds? Everett, Quantum Theory, and Reality". Oxford University PressOxford, England (2010). DOI: 10.1093/acprof:oso/9780199560561.001.0001.

[72] Lev Vaidman. "Many-Worlds Interpretation of Quantum Mechanics". In Edward N. Zalta, editor, The Stanford Encyclopedia of Philosophy . Metaphysics Research Lab, Stanford Universityfall 2018 edition (2018).

[73] Carlo Rovelli. "An argument against the realistic interpretation of the wave function". Foundations of Physics **46**, 1229–1237 (2016). DOI: 10.1007/s10701-016-0032-9. eprint: arXiv:1508.05533.

[74] Christopher A. Fuchs. "On Participatory Realism" (2016) eprint: arXiv:1601.04360. to appear in "Information & Interaction: Eddington, Wheeler, and the Limits of Knowledge", edited by Ian T. Durham and Dean Rickles.

[75] Carlo Rovelli. "Space is blue and birds fly through it". Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **376**, 20170312 (2018). DOI: 10.1098/rsta.2017.0312. eprint: arXiv:1712.02894.

[76] Andrea Di Biagio and Carlo Rovelli. "Stable Facts, Relative Facts". Foundations of Physics **51**, 30 (2021). DOI: 10.1007/s10701-021-00429-w. eprint: arXiv:2006.15543.

[77] Wayne C. Myrvold. "The Science of $\Theta\Delta^{cs}$". Foundations of Physics **50**, 1219–1251 (2020). DOI: 10.1007/s10701-020-00371-3. eprint: arXiv:2007.11729.

[78] Leonard Mlodinow and Todd A. Brun. "Relation between the psychological and thermodynamic arrows of time". Physical Review E **89**, 052102 (2014). DOI: 10.1103/physreve.89.052102.