INTRODUCTION

There are many physical phenomena whose dynamics are asymmetric along the time axis, providing ‘arrows of time’ [2,4]. Among the least studied of these is the casual arrow of time: an action we take at time \(t = T\) always seems to have consequences on one side of the co-ordinate time axis (that we label by convention with increasing times \(t > T\)), and never seems to have consequences on the other side of that axis (labeled with decreasing times \(t < T\)) [3, 5–7]. This asymmetry cannot be explained away by assuming that actions affecting the past would lead to logical paradoxes, for non-paradoxical accounts of backwards-in-time causation can readily be found in philosophy [8], physics [9] and science fiction [10]. The origin of the time asymmetry of causality remains an open question. [11]

This work aims to contribute to this big question by asking a smaller question: is the perceived direction of the causal influence intrinsic to the system, or is it a relative property of how the observer interacts with the system? The answer is not immediately obvious, because causal relations are ascertained by taking actions on a system and observing the results, which requires the participation of both the observer and the system. We emphasize that what is at stake is not a metaphysical question but a practical one: does the direction of causality necessarily appear in an observer-independent description of a phenomenon? If not, then we can deduce that it is a property that depends in some way on the observer, because it is a relative property of the observer and the phenomenon, or because it is entirely contributed by the observer. For example, the statue of Aristotle in Thessaloniki has a rest-mass which may be considered an intrinsic property of the statue, because its value is the same regardless of how and by whom it is measured. Its velocity is an observer-relative property because it cannot be specified except in relation to the velocity of the approaching tourist. Finally, the history of the statue is neither a property of the statue in itself nor a property of its physical relation to us, but rather is contained in records of our perceptions of the statue. It is natural to wonder: is the perceived direction of causality intrinsic like rest-mass, relative like velocity, or is it part of the memory and perceptual apparatus of the observer?

Progress on this question can only be made within a conceptual framework that gives ‘causality’ a more precise meaning. One approach that has proven to be successful in a wide range of scientific disciplines is causal modeling [5, 12, 13]. In a causal model, there are essentially two modes by which an observer may ascertain causal relations. The first is called an observational scheme and represents a form of data collection that is ‘passive’ in some appropriate sense, usually understood to mean that the system is unaffected by the act of observation. The second mode of interaction is an interventionist scheme in which the observer probes or disrupts specific variables of the system and observes the effect on the other variables. The conceptual difference between these two modes of interaction may be likened to the difference between observing a frog in a pond and dissecting it on a table: generally, watching the frog in the pond is sufficient to suggest some hypotheses about how it functions, but cutting it up is really the only way to know for sure. Perhaps because causal modeling has its roots in engineering and the practical sciences, causal models strongly favour the view that causal relations, like the bones and musculature of the frog, are intrinsic to the system. Even if the direction of a cause cannot be ascertained from the observational scheme, it is presumed to have some actual orientation that will be revealed once the observer makes an intervention. Things become less clear when the systems under question are quantum. As we will see, the concept of an observational scheme be-
comes essentially ill-defined for quantum systems. In the face of this, a few authors have proposed to re-define an observational scheme in a way that remains true to the spirit of the classical definition, however, these attempts face conceptual difficulties. Many authors simply abandon the idea, choosing instead to interpret a quantum causal model purely within the setting of an interventionist scheme. Any formulation of a causal model purely in terms of an interventionist scheme denies us the means to separate the intrinsic from the observer-dependent aspects of causality. It is therefore necessary to propose a new definition of an observational scheme suitable for quantum systems. This requires us to abandon any notion of causality that presupposes a mechanism, and reinterpret causality as a kind of relation between counterfactuals. For example, we can regard the causal structure as a relation that enables us to deduce, from the frog’s movements in the pond, what would happen if we were to take it into the lab and probe each of its limbs separately. This view is compatible with a mechanical interpretation of causality, but does not require it.

In a longer companion work, a framework for causal modeling was developed based on this alternative interpretation of causality as relations among counterfactual experiments. The present work builds on that framework by using it to motivate a definition of a quantum observational scheme. It is shown that the probabilities obtained in the observational scheme plus the causal structure is sufficient to deduce what the probabilities would be for arbitrary interventions on any of the variables, if and only if the causal structure has the form of a ‘layered’ graph. The result is proven using the process matrix definition of a quantum causal model given by other authors in Ref. [1], but with certain restrictions imposed so as to make the model conform to the framework of Ref. [1]. I then ask what can be deduced from the quantum observational scheme about the direction of causal influences in a system, without prior knowledge of the causal structure. In stark contrast to the classical case, it is found that no information about the direction of causality is possible without performing an intervention. This implies that causal modeling of quantum systems can be done in a way that does not require a causal direction to be specified a priori, as is usually assumed. It is therefore compatible with the notion that the direction of causal influence is an observer-dependent phenomenon, in line with the views expressed in eg. Refs. [3, 7, 19].

FUNCTIONAL MODELS VERSUS CAUSAL MODELS

To understand causal models it is helpful to start with their conceptual predecessors: functional models.

Definition: A functional model is specified by a pair \( \{ G, V \} \) consisting of a directed acyclic graph (DAG) \( G \) with nodes corresponding to \( N \) variables \( V = \{ V_i : i = 1, 2, ..., N \} \); and a set of model parameters \( \{ \eta_i, P_i(\eta_i), f_i : i = 1, 2, ..., N \} \) where \( \eta_i \) are independent random variables with probabilities \( P_i(\eta_i) \), and \( f_i \) are deterministic functions such that \( V_i = f_i(\text{pa}V_i, \eta_i) \) where \( \text{pa}V_i \) denote the parents of \( V_i \) in \( G \).

In a functional model, probabilities for the variables \( V \) in an observational scheme are obtained by a simple procedure: one first samples a set of values from the \( P_i(\eta_i) \) and then uses the functions \( f_i(\text{pa}V_i, \eta_i) \) to progressively determine the values of each \( V_i \). The resulting distribution obeys a special constraint, called the Causal Markov Condition [5, 12, 20]:

\[
P_{\text{obs}}(V) = \prod_{i=1}^{N} P_{\text{obs}}(V_i | \text{pa}V_i)
\]

(Note that the form of this constraint depends on the sets \( \text{pa}V_i \) and hence is relative to the DAG \( G \)). Conversely, if some \( P(V) \) satisfies (1) relative to a DAG \( G \), then there exists a functional model which generates \( P(V) \) in this way. Since functional models treat the concept of ‘causal relation’ as synonymous with ‘functional dependence’, they also provide a direct way of formalizing the notion of intervention on any variable \( V_j \). Intuitively, an intervention is a physical action on a dependent variable \( V_j \) in a system that cuts off this variable from its causes within the system, and allows it to be prepared in an arbitrary state by the experimenter. In a functional model, this is done by introducing an experimentally controlled variable \( \xi_j \) and replacing \( f_j(\text{pa}V_j, \eta_j) \) with the new dependence \( f_j(\xi_j) \). The arrows connecting \( \text{pa}V_j \) to \( V_j \) are then deleted from the DAG. The probability distribution generated from this intervened model also obeys a special constraint:

\[
P_{\text{interv}}(V | \text{do}(V_j = \xi_j)) = \delta(V_j, \xi_j) \prod_{i \neq j} P_{\text{obs}}(V_i | \text{pa}V_i),
\]

where \( \delta(V_j, \xi_j) = 1 \) if \( V_j = \xi_j \) and 0 otherwise, and the notation \( \text{do}(V_j = \xi_j) \) reminds us that we are forcing \( V_j \) to the value \( \xi_j \) by intervention, instead of passively observing it. The collection of \( N \) such distributions, representing interventions on each \( V_i \), defines an interventionist scheme. To make the further step to a causal model, note that the form of both of the constraints (1), (2) is independent of the particular choices of model parameters \( \{ \eta_i, P_i(\eta_i), f_i \} \) from which they were derived; in fact, it only depends on the causal structure \( G \). Hence, only \( G \) and \( P_{\text{obs}} \) are needed to describe observations at the empirical level. In the companion work [1], this is elevated to an axiom, called ‘causal sufficiency’. It is a key property of causal models in our framework, because it allows us to elevate the concept of causal relation to a universal
status that transcends mere mechanism.

**Definition:** A causal model is a pair \( \{G, P_{\text{obs}}(V)\} \), where \( P_{\text{obs}}(V) \) satisfies the Causal Markov Condition \( 1 \) relative to \( G \), and the results of interventions are given by the formula \( 2 \).

The importance of this step cannot be overstated. In a functional model, causal structure was embodied in the deterministic mechanisms \( f_i(pa V_i, \eta_i) \), from which we derived the constraints \( 1, 2 \). In a causal model, as we interpret it here, this relationship is turned on its head: the constraints \( 1, 2 \) are taken as postulates that serve to define causal structure in terms of probabilities. This can be made more precise by noticing that the RHS of Eq. \( 2 \) is entirely determined by the probabilities from the observational scheme, plus causal structure. That is, ‘causal structure’ merely codifies the precise way in which data obtained in the observational scheme constrains the data that would be obtained for any conceivable intervention by an observer.

But what of the direction of causal influence? Curiously, the Causal Markov Condition \( 1 \) is not invariant under swapping the directions of all arrows in the DAG \( G \). This asymmetry is usually understood in terms of it’s simplest manifestation in the case of two nodes \( V_2, V_3 \) having a single parent \( V_1 \). In that case \( 1 \) is equivalent to the factorization of \( V_2, V_3 \) conditional on \( V_1 \), i.e. \( P(V_2, V_3 | V_1) = P(V_2 | V_1)P(V_3 | V_1) \), known as Reichenbach’s Principle of Common Causes. \( 21, 22 \). The asymmetry arises because reversing the direction of causal arrows in the DAG makes \( V_1 \) a common effect of \( V_2, V_3 \), and \( 1 \) does not require factorization of \( V_2, V_3 \) conditional on \( V_1 \) – on the contrary, conditioning on a common effect of two variables typically correlates them, a result known in statistics as Berkson’s effect \( 24 \). This makes it possible in many cases to deduce the direction of causality purely from an observational scheme. Traditionally, an observational scheme is formalized as a method of making measurements that does not disturb the system. This allows us to interpret \( P_{\text{obs}}(V) \) as representing information about the system that would be true even if nobody had observed the system, i.e. information about the intrinsic properties of the system. It follows that the direction of causality is most naturally interpreted as one of the system’s intrinsic properties.

### QUANTUM INTERVENTIONS AND OBSERVATIONS

In the broadest sense, a quantum causal model is any model that formalizes causal structure as a means of encoding the consequences of possible interventions on quantum systems. Unlike in the classical case, quantum systems have only been known to us under highly controlled laboratory settings. It is therefore natural that the first attempts at causal modeling of these systems have focused on finding a quantum generalization of the interventionist scheme. These models can be summarized by an equation analogous to \( 2 \) having the form of a ‘generalized Born rule’ \( 6, 14, 22 \):

\[
P(V|\xi_1, \xi_2, \ldots, \xi_N) = \text{Tr} \left[ M_{v_1}^{\xi_1} \otimes \cdots \otimes M_{v_N}^{\xi_N} \cdot W \right].
\]

(3)

In this expression, each \( M_{v_i}^{\xi_i} \) is the Choi-Jamiolkowski matrix representation of a completely positive (CP) map \( M_{v_i}^{\xi_i}(\rho) : \mathcal{H}_{V_i}^{\text{in}} \rightarrow \mathcal{H}_{V_i}^{\text{out}} \), which represents the effect of setting the control variable to the value \( \xi_i \), and obtaining the outcome \( v_i \in \text{dom}(V_i) \) from the domain of possible values \( \text{dom}(V_i) \). The set of CP maps \( \{ M_{v_i}^{\xi_i} : v_i \in \text{dom}(V_i) \} \) defines a quantum instrument, which replaces the classical notion of intervention, and \( W \) is a positive semi-definite quantum process matrix that maps arbitrary choices of quantum instruments to probability distributions over the observed outcomes \( V \). Conceptually, this model treats causality in much the same way as in a functional model, with \( W \) standing in the role of the model parameters and where causal structure refers to the structure of the quantum process \( W \) from which constraints on the probabilities are derived. This view implicitly supports interpreting \( W \) (and hence the direction of causal influence) as a property intrinsic to the system being probed. At least, if one wishes to argue otherwise, it is necessary to supplement this description with an observational scheme that would allow us to define \( P_{\text{obs}}(V) \) as containing only the observer-independent information. The usual definition of an observational scheme in terms of non-disturbing measurements is problematic for quantum systems. On one hand, in order to perform inference, the measurements need to be informative about the system, but this implies that they disturb the state of the system. This feature of quantum systems has been formalized under the slogan “no information without disturbance” \( 29 \). Let us therefore re-evaluate the meaning of an observational scheme by looking carefully at the role it actually plays in a causal model. Two notable features emerge. First, from Eq. \( 1 \) we see that the scheme is passive in the sense that it requires no active choices from the observer: there are no ‘control variables’, there are only outcomes of a fixed set of measurements. Second, from Eq. \( 2 \), we see that the data from the observational scheme plus the causal graph \( G \) is sufficient to determine the results of an arbitrary intervention. This latter feature points to an interesting new way of thinking about an observational scheme: not as a scheme that reveals the system’s properties in absence of observation (as usually assumed), but as a scheme that indicates the system’s behaviour in the presence of any observation. The information obtained in such a scheme depends on the fact of observation, but not the particular features of the observation. It thus represents characteristics that may be called intrinsic to the system, not in the sense that they
obtain in the absence of observation, but in the sense that they obtain independently of which observation is made. Thus, what gives causal relations their ‘objective’ character in an observational scheme is not that we are taking a view from nowhere, but rather that we are adopting the viewpoint of no-one in particular [27]. Existing proposals for quantum observational schemes fail to meet both of the above requirements exemplified in Eqs. (1), (2). The informationally symmetric measurements proposed in Refs. [14, 28], as well as the active quantum observational scheme in Ref. [15] both allow the observer to actively choose which instruments to apply in each run, hence are not passive. Both definitions moreover have limited power to perform tomography of the quantum process $W$, and hence cannot meet the second requirement of enabling one to deduce the results of an arbitrary intervention. I now propose a definition of a quantum observational scheme that is passive in the sense described above, i.e. that does not depend on the active selection of control variables by an observer. Armed with this definition, we will subsequently deal with the problem of tomography.

**Definition:** A quantum observational scheme assigns to each variable $V_i \in \text{dom}(V_i)$ a fixed quantum instrument $\tilde{M}_{V_i} := \{\tilde{M}_{v_i} : v_i \in \text{dom}(V_i)\}$ whose input and output Hilbert spaces have the same dimension, $\dim(\mathcal{H}_{V_i}^\text{in}) = \dim(\mathcal{H}_{V_i}^\text{out}) := d_i$. Furthermore, these instruments are passively informationally complete, which means they satisfy:

$$\tilde{M}_{v_i}(\rho) = \sqrt{F_{v_i}} \rho \sqrt{F_{v_i}} \quad \forall \tilde{M}_{v_i} \in \tilde{M}_{V_i} \quad (4)$$

where $F_{v_i} := \beta_{v_i} \Pi_{v_i}$ for rank-1 projectors $\Pi_{v_i}$, and the set $\{F_{v_i} : v_i \in \text{dom}(V_i)\}$ forms an informationally complete POVM for $\mathcal{H}_{V_i}$, i.e. this set spans the space of linear operators $\mathcal{L}(\mathcal{H}_{V_i})$, and satisfies $\sum_{v_i} F_{v_i} = 1^{V_i}$, where $1^{V_i}$ is the identity matrix on $\mathcal{H}_{V_i}$. Note that this implies $\text{dom}(V_i)$ must have a number of values $D_i \geq d_i^2$, and that the coefficients $\beta_{v_i}$ must satisfy $\sum_{v_i} \beta_{v_i} = d_i$.

It is important to mention that a passively informationally complete instrument is not the same as an informationally complete set of instruments sometimes mentioned in the literature [14, 29]. The latter refers to a set of instruments, corresponding to different possible settings of the control variables, whose defining property is that they can be used to perform tomography of $W$. By contrast, the defining feature of a passively informationally complete instrument is that it is ‘passive’ in the sense that it is a single instrument that cannot be changed from one run of the experiment to the next. As such, it has no associated control variable, and in general is not guaranteed to be able to perform tomography of $W$ – this is the problem to which we next turn our attention.

It is clear that reconstruction of $W$ from the observed probabilities will not be possible without making some assumptions about the form of $W$ itself, simply because any fixed choice of quantum instruments (even if they are passively informationally complete) are not sufficient to perform tomography of an arbitrary process [30]. Remarkably, full tomography of $W$ is possible if we merely restrict the causal structure of the process:

**Result 1:** Tomography of a general process matrix $W$ compatible with the causal structure $G$ is possible in a quantum observational scheme if and only if $G$ is a layered DAG (proof in Appendix A).

(Definition: A DAG $G$ on a vertex set $V$ is said to be layered if the nodes can be partitioned into $K$ subsets or ‘layers’ $V = L_1 \cup L_2 \cup \cdots \cup L_K$ such that no node is a causal descendant of another node in the same layer, and $j > i$ whenever there is a node in $L_j$ that is a descendant of a node in $L_i$.)

In order to uphold the key property of causal models – that the observational scheme should be sufficient to indicate what happens under an arbitrary intervention – it is therefore necessary and sufficient that the causal structure should be layered as described above. A further constraint that has been proposed in the literature is that $W$ should be an unbiased process [14, 15, 31], i.e. it should preserve the maximally mixed state. If we make this extra assumption, we obtain the following intriguing result:

**Result 2:** If the DAG $G$ is layered and the process $W$ is unbiased, then the probabilities $P_{\text{obs}}(V)$ obtained in the observational scheme are causally reversible, that is, they are compatible with a quantum process whose DAG $G^*$ is obtained by reversing the directions of all arrows in $G$ (proof in Appendix B).

One would therefore expect that a quantum generalization of the Causal Markov Condition to unbiased quantum processes would be causally reversible, unlike its classical counterpart. This expectation is confirmed in the companion paper [1], where such a generalization is constructed. From our present definition of a quantum observational scheme, it follows that the direction of causality cannot be identified by considering only the observer-invariant features of the system. The direction is only revealed (or, perhaps more provocatively, produced) when an intervention is performed, which implicates a particular observer. This suggests that a time-symmetric description of the fundamental laws of physics may yet be compatible with the apparent asymmetry of the causal arrow of time. Our work provides a complementary perspective to the work of Ref. [7], where it was shown that time-symmetric causality can be recovered by relaxing the constraints on quantum dynamics to allow for more general processes. Here, we have shown that the converse is also possible: time-symmetry of the causal arrow can be restored by restricting quantum dynamics to unbiased processes. Given that the present...
work points to the observer having a fundamental role in determining the direction of causality, the task of providing a formal description of how this comes about is an interesting topic for future work. In particular, it would be interesting to investigate whether the causal arrow is grounded in purely physical relations, or whether it depends on more complex properties of the observer, such as their ability to perceive and store information.

[1] P. J. L., (2018), preprint at arXiv:1806.00805
[2] H. Zeh, The Physical Basis of The Direction of Time (Springer-Verlag Berlin Heidelberg, 2007).
[3] H. Price, Time’s Arrow & Archimedes’ Point: New Directions for the Physics of Time, Oxford Paperbacks: Philosophy (Oxford University Press, 1997).
[4] J. Halliwell, J. Pérez-Mercader, and W. Zurek, Physical Origins of Time Asymmetry (Cambridge University Press, 1996).
[5] J. Pearl, Causality (Cambridge University Press, 2009).
[6] B. C. Cnecoe and R. Lal, Phys. Rev. Lett. 108, 200403 (2012).
[7] O. Oreshkov and N. J. Cerf, Nature Physics 11, 853 (2015).
[8] J. J. Smith, in The Stanford Encyclopedia of Philosophy, edited by E. N. Zalta (Metaphysics Research Lab, Stanford University, 2018) summer 2018 ed.
[9] J. Earman, L. Wihchrich, and J. Manchak, in The Stanford Encyclopedia of Philosophy, edited by E. N. Zalta (Metaphysics Research Lab, Stanford University, 2016) winter 2016 ed.
[10] The reader can probably fill in their own examples. My favourite is Shane Carruth’s film, Primer.
[11] There are actually two aspects to this asymmetry: first, that actions always produce consequences in the same temporal direction for repeated actions on the same system; second, that the directions so obtained for any pair of distinct systems are always found to be in agreement.
[12] P. Spirtes, C. Glymour, and R. Scheines, Causation, Prediction, and Search, Lecture Notes in Statistics (Springer New York, 2012).
[13] J. Woodward, Making Things Happen: A Theory of Causal Explanation, Oxford scholarship online (Oxford University Press, 2003).
[14] K. Ried, M. Agnew, L. Vermeyden, D. Janz, R. W. Spekkens, and K. J. Resch, Nature Physics 11, 414420 (2015).
[15] J. Kühler and D. Braun (2018) preprint at arXiv:1803.00939
[16] F. Costa and S. Shrapnel, New Journal of Physics 18, 063032 (2016).
[17] J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens, Phys. Rev. X 7, 031021 (2017).
[18] A. Kissinger, M. Hoban, and B. Cnecoe, (2017) preprint at arXiv:1708.04118.
[19] P. A. Guérin and Časlav Brukner (2018) preprint at arXiv:1805.11249.
[20] D. Hausman and J. Woodward, Brit. J. Phil. Sci. 50, 521 (1999).
[21] E. Cavalcanti and R. Lal, J. Phys. A 47, 424018 (2014).
[22] F. Arntzenius, in The Stanford Encyclopedia of Philosophy, edited by E. N. Zalta (Metaphysics Research Lab, Stanford University, 2010) fall 2010 ed.
[23] H. Reichenbach, The Direction of Time (University of Los Angeles Press, 1956).
[24] J. Berkson, Int. J. Epidemiol. 43, 511 (2014).
[25] O. Oreshkov, F. Costa, and Č. Brukner, Nature Communications 3, 1092 (2012).
[26] P. Busch, “no information without disturbance”: Quantum limitations of measurement,” in Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony (Springer Netherlands, Dordrecht, 2009) pp. 229–256.
[27] A. Fine, Proceedings and Addresses of the American Philosophical Association 72, 7 (1998).
[28] K. Ried, Causal Models for a Quantum World (University of Waterloo, 2016) phD. Thesis.
[29] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Phys. Rev. Lett. 120, 040405 (2018).
[30] S. Miltz, F. A. Pollock, and K. Modi, Open Sys. & Inf. Dyn. 24, 1740016 (2017).
[31] F. Costa, M. Ringbauer, M. E. Goggin, A. G. White, and A. Fedrizzi. (2017) preprint at arXiv:1710.01776.
[32] G. Chiribella, G. D’Ariano, and P. Perinotti, “Quantum from principles,” in Quantum Theory: Informational Foundations and Foils, edited by G. Chiribella and R. Spekkens (Springer, Dordrecht, 2016) pp. 171–221.

Appendix A: Proof of Result 1

Claim: Tomography of a general process matrix $W$ which is compatible with a given causal structure represented by a DAG $G$ is possible in a quantum observational scheme if and only if $G$ is a layered DAG. To prove this claim, we interpret compatible factorizes over the DAG $G$ as defined in Ref. [10]. This will be explained below.

We first prove the ‘if’ part. Let the nodes of $G$ be partitioned into $M$ sets $L_1 \cup L_2 \cup \ldots \cup L_K$ corresponding to the layers of $G$. Let $H^I_{L_j}$ be the combined input Hilbert space of all nodes in the layer $L_j$, and similarly let $H^O_{L_j}$ be the combined output Hilbert space from that layer. Then any process matrix $W$ that factorizes over the DAG $G$ has the form [10]:

$$W = W^{I_1} \otimes W^{O_{I_2}} \otimes \ldots \otimes W^{O_{K-1}} \otimes I_K \otimes I^{O_{K}},$$

where $I^{O_{K}}$ is the identity operator on $H^O_{L_j}$. The probabilities for a quantum observational scheme are now obtained from the generalized Born rule [24], by substituting in the passively informationally complete instruments $M_{I_i}$, for each node. Let us define $\hat{M}_{I_i}$ as the CP map on $H^I_{L_j}$ corresponding to the vector of values $I := \{I_j\}$ obtained for all nodes in the layer $L_j$. Then the set of these maps for all sets of values $I_j$ defines a quantum instrument on the entire layer, $M_{L_j} := \{\hat{M}_{I_j} : I_j \in \text{dom}(L_j)\}$, which
is also passively informationally complete. We can now write the observed probabilities from the quantum observational scheme as:

\[
P_{\text{obs}}(l_1, l_2, \ldots, l_K) = \text{Tr} \left[ M_{l_1} \otimes \cdots \otimes M_{l_K} \cdot W \right].
\]  

(6)

The definition of a passive informationally complete instrument implies that the Choi-Jamiolkowski matrices \( M_{l_j} \) have the form:

\[
M_{l_j} = \Pi_{l_j}^{O_j} \otimes F_{l_j}^{O_j},
\]

(7)

where

\[
\Pi_{l_j} := \bigotimes_{v_i \in L_j} \Pi_{v_i},
\]

\[
F_{l_j} := \beta_{l_j} \Pi_{l_j},
\]

\[
\beta_{l_j} := \prod_{v_i \in L_j} \beta_{v_i}.
\]

(8)

and the set of operators \( \mathcal{F}_{L_j} := \{ F_{l_j} : l_j \in L_j \} \) automatically forms an informationally-complete POVM for the whole layer. Inserting these operators and the factorized process from (5) into (6), we obtain:

\[
P_{\text{obs}}(l_1, l_2, \ldots, l_K) = \text{Tr} \left[ F_{l_1}^{I_1} \cdot W_{l_1}^{I_1} \right] 
\]

\[
\times \prod_{j=1}^{K} \text{Tr} \left[ \Pi_{l_j}^{O_j} \otimes F_{l_{j+1}}^{I_{j+1}} \cdot W_{l_j}^{O_j} I_{j+1} \right] = P(l_1)P(l_2|l_1), \ldots, P(l_K|l_{K-1}),
\]

(9)

where

\[
P(l_{j+1} | l_j) = \text{Tr} \left[ \Pi_{l_j}^{O_j} \otimes F_{l_{j+1}}^{I_{j+1}} \cdot W_{l_j}^{O_j} I_{j+1} \right]
\]

(10)

is the conditional probability to obtain \( l_{j+1} \) when measuring the POVM \( \mathcal{F}_{L_{j+1}} \) on the layer \( L_{j+1} \) given that the outcome \( l_j \) was obtained on the previous layer \( L_j \). Since \( \{ \Pi_{l_j}^{O_j} \otimes F_{l_{j+1}}^{I_{j+1}} \} \) form a set of rank-1 projectors that spans the space of linear operators \( \mathcal{L}(H_{L_j}^{O_j} \otimes H_{L_{j+1}}^{I_{j+1}}) \), the probabilities \( P(l_{j+1} | l_j) \) are sufficient to reconstruct an arbitrary sub-process \( W_{l_j}^{O_j} I_{j+1} \), and hence using Eq. (9) we can reconstruct the full process \( W \) from the observed probabilities \( P_{\text{obs}}(V) \).

For the ‘only if’ part of the proof, first note that for any DAG \( G \) the nodes can be partitioned into sets \( \{ S_j \} \) such that there is a directed path from a member of \( S_i \) to \( S_j \) whenever \( j > i \), and no directed paths between any members of the same set. As with layers, one can define the rank-1 projectors \( \Pi_{S_i} \) that together span the space of linear operators on the joint Hilbert space of all the nodes in the set \( S_j \). Now, if \( G \) is not a layered DAG, then there must exist three sets \( S_j, S_{k'}, S_{l'} \) with \( j' < k' < l' \) such that there is a path from \( S_{j'} \) to \( S_{l'} \) that does not intersect \( S_{k'} \). Without loss of generality, we can choose the specific labels \( j' = 2, k' = 3, l' = 4 \). Then the factorization condition from Ref. [16] for the DAG implies that \( W \) has the form:

\[
W = W_{<2} \otimes W_{O_2 I_3 O_3 I_4} \otimes W_{>4},
\]

(11)

where by assumption \( W_{O_2 I_3 O_3 I_4} \) cannot be further decomposed. Substituting this into (9) we obtain a factorization of \( P_{\text{obs}}(V) \) into a product of terms, which contains the term:

\[
P(s_4, s_5 | s_2) = \text{Tr} \left[ \Pi_{s_2}^{O_2} \otimes F_{s_3}^{I_3} \otimes F_{s_4}^{O_4} \otimes F_{s_4}^{I_4} \cdot W_{O_2 I_3 O_3 I_4} \right].
\]

Now \( W_{O_2 I_3 O_3 I_4} \) is a linear operator in the space \( \mathcal{L}(H_{S_2}^{O_2} \otimes H_{S_3}^{O_3} \otimes H_{S_4}^{O_4} \otimes H_{S_4}^{I_4}) \), for which a spanning set of operators must have at least \( (d_2 d_3 d_4)^2 \) elements, with \( d_j \) the dimension of the joint Hilbert space \( H_{S_j} \) associated to the set \( S_j \). However, the set of projectors \( \{ F_{s_2} \otimes F_{s_3} \otimes F_{s_4} \} \) for all sets of values of \( s_2, s_3, s_4 \) is only required to have \( (d_2 d_3 d_4)^2 \) elements in order to satisfy the definition of being passively informationally complete, and so will in general not be sufficient to reconstruct the sub-process \( W_{O_2 I_3 O_3 I_4} \) from the observed probabilities. This concludes the proof. (Note that this failure can be traced to the fact that the observations are ‘passive’ which here means that the projector \( \Pi_{S_3} \) on the joint output space of \( S_3 \) cannot be chosen independently of the POVM element \( F_{s_3}^{I_3} \) on the input space of \( S_3 \), viz. because the outcome of the latter fixes the post-measurement state of the former via the Lüders rule).

Appendix B: Proof of Result 2

Claim: If the DAG of \( W \) is layered and the process \( W \) is unbiased, then the probabilities \( P_{\text{obs}}(V) \) obtained in a quantum observational scheme are time reversible, that is, they are compatible with another quantum process \( \bar{W} \) whose DAG \( G^* \) is obtained by reversing the directions of all arrows in \( G \).

To be more precise, our goal is to show that if \( P_{\text{obs}}(V) \) factorizes as shown in Eq. (9), where each \( W_{O_j} I_{j+1} \) is a valid unbiased quantum process representing a CPT map from \( H_{L_j}^{O_j} \) to \( H_{L_{j+1}}^{I_{j+1}} \), then \( P_{\text{obs}}(V) \) must also factorize in the reverse order:

\[
P_{\text{obs}}(V) = P(l_1 | l_2), P(l_2 | l_3), \ldots, P(l_{K-1} | l_K), P(l_K),
\]

such that each term can be expressed as:

\[
P(l_j | l_{j+1}) = \text{Tr} \left[ \Pi_{l_{j+1}}^{O_{j+1}} \otimes F_{l_j}^{I_j} \cdot W_{O_{j+1} I_j} \right],
\]

where \( W_{O_{j+1} I_j} \) is a valid quantum process matrix representing a CPT map from \( H_{L_{j+1}}^{O_{j+1}} \) to \( H_{L_j}^{I_j} \). First, note that the factorization of \( \bar{W} \) can be obtained directly
from (9) by applying the standard formula for Bayesian inversion to each term:

\[ P(l_{j+1}|l_j) = \frac{P(l_{j+1})P(l_j|l_{j+1})}{P(l_j)} \]

Next, comparing (12) to (10), we see that the Claim is fulfilled only if we define

\[ W^{O_{j+1}I_j} := \{ W^{O,j+1}I_j \}_{I \leftrightarrow O} \frac{P(l_j)}{P(l_{j+1})} \frac{\beta_{j+1}}{\beta_j}, \quad (12) \]

where \( \{ \ldots \} \_{I \leftrightarrow O} \) means re-labeling the Hilbert spaces to switch their roles as ‘input’ and ‘output’ from their respective layers. It is important to note that a process is valid if, which is a CPT map for one choice of input/output labelling will not in general be a valid CPT map when the input/output labels are switched: this is just the mathematical manifestation of the ‘causal arrow of time’ for quantum processes. As a simple example, consider a quantum channel from the output of Alice’s laboratory \( (H_{OA}) \) to the input of Bob’s laboratory \( (H_{IB}) \) that discards Alice’s output and produces a fixed pure state \( \psi \) at Bob’s input. This is represented by the process matrix \( W^{O_{OA}IB} = I_{OA} \otimes |\psi\rangle\langle\psi|_{IB} \). Now consider the matrix obtained by switching the input and output:

\[ W^{IAOB} := \{ W^{O_{IA}IB} \}_{I \leftrightarrow O} I_{IA} \otimes |\psi\rangle\langle\psi|_{OB} \].

The Hilbert spaces are the same, but their roles are reversed: \( W^{IAOB} \) is now to be interpreted as a map from the output of Bob’s laboratory to the input of Alice’s laboratory. However, under this interpretation it does not represent a valid quantum process because

\[ Tr_{IA} [W^{O_{IA}IB}] = |\psi\rangle\langle\psi|_{OB} \neq I_{OB}, \]

\[ Tr [W^{IAOB}] = d_{IA} \neq d_{OB}. \quad (13) \]

Qualitatively, the first property means the map forces the output of Bob’s lab to be \( |\psi\rangle\langle\psi|_{OB} \) (regardless of Bob’s efforts to produce something else) and hence represents a case of deterministic post-selection of Bob’s output, which is not allowed by quantum theory. (Indeed, the assumption that there can be no deterministic post-selection is often called causality in the literature.) The second property shows that the state produced at Alice’s input is \( I_{IA} \), which is not normalized. Moreover, generally, the conditions for any process matrix to represent a valid CPT map from \( H^{O_j}_{L_{j+1}} \) to \( H^{I_{j+1}}_{L_j} \) are given by (5):

\[ W^{O_{j+1}I_j} \geq 0, \]

\[ Tr_{I_{j+1}} [W^{O_{j+1}I_j}] = I_{O_j}, \]

\[ Tr [W^{O_{j+1}I_j}] = d_{O_j}. \quad (14) \]

Notice that these conditions are not symmetric under a re-labelling of \( O \leftrightarrow I \), and hence it matters which Hilbert space is interpreted as the ‘output’ of the preceding layer (hence the input to \( W \)) and which is the ‘input’ to the next layer (hence the output from \( W \)). It is therefore important to verify whether the matrix \( \tilde{W}^{O_{j+1}I_j} \) defined in (12) represents a valid quantum process. As we now show, this is guaranteed if the original process \( W^{O_{j+1}I_j} \) is unbiased, which means it satisfies the additional constraint:

\[ Tr_{O_j} [W^{O_{j+1}I_j}] = \frac{d_{j+1}}{d_j} I_{I_{j+1}}. \quad (15) \]

The conditions for \( W^{O_{j+1}I_j} \) to be a valid CPT map from \( H^{O_{j+1}}_{L_{j+1}} \) to \( H^{I_{j+1}}_{L_j} \) are given by:

\[ W^{O_{j+1}I_j} \geq 0, \]

\[ Tr_{I_j} [W^{O_{j+1}I_j}] = \Pi_{O_{j+1}}, \]

\[ Tr [W^{O_{j+1}I_j}] = d_{O_{j+1}}. \quad (16) \]

To prove that these conditions are met, we first simplify the expression (12) using the fact that \( W^{O_{j+1}I_j} \) is unbiased. Since the initial term \( W^{I_{j+1}}_{I_j} \) of (5) can be thought of as a process from a trivial Hilbert space (with dimension \( d = 1 \)) to the space \( H^{I_{j+1}}_{L_j} \), the condition (15) implies that \( W^{I_{j+1}}_{I_j} = \frac{1}{d_j} I_{I_{j+1}} \). Next, we note that if one marginalizes over the outcome \( v_i \) of a passively informationally complete instrument, the result is a CPT map that is automatically unbiased:

\[ \sum_{v_i} \tilde{M}_{v_i} (I) = \sum_{v_i} \beta_{v_i} \Pi_{v_i} (I) \Pi_{v_i} \]

\[ = \sum_{v_i} \beta_{v_i} \Pi_{v_i} = I. \quad (17) \]

The above two results imply that for a unbiased process \( W \) with a layered DAG, the marginal probability for the outcome set \( I_j \) in a given layer \( L_j \) (i.e., after summing over the outcomes obtained in all other layers) is equal to the probability of obtaining \( I_j \) when measuring the maximally mixed state, that is:

\[ P(I_j) = Tr \left[ \frac{1}{d_j} I_{I_j} \right] = \frac{\beta_j}{d_j}. \quad (18) \]

Substituting this into (12) we obtain the simpler expression:

\[ W^{O_{j+1}I_j} := \{ W^{O_{j+1}I_j} \}_{I \leftrightarrow O} \frac{d_{j+1}}{d_j}. \quad (19) \]

Since \( \frac{d_{j+1}}{d_j} > 0 \), the first condition of (16) is immediately met. For the second condition, with the help of (15), we find:

\[ Tr_{I_j} [W^{O_{j+1}I_j}] = Tr_{I_j} \left( \{ W^{O_{j+1}I_j} \}_{I \leftrightarrow O} \right) \frac{d_{j+1}}{d_j} 
\]

\[ = \left\{ Tr_{I_j} [W^{O_{j+1}I_j}] \right\}_{I \leftrightarrow O} \frac{d_{j+1}}{d_j} 
\]

\[ = I_{O_{j+1}}, \quad (20) \]
\[ \text{and} \]
\[ \text{Tr} \left[ \bar{W}_{O_{j+1}I_j} \right] = \left\{ \text{Tr} \left[ W^{O_jI_{j+1}} \right] \right\}_{I_{j+1}O} \frac{d_{j+1}}{d_j} \]
\[ = (d_j) \frac{d_{j+1}}{d_j} \]
\[ = d_{j+1} \quad (21) \]

and hence all conditions (10) are met. We conclude that \( \bar{W}_{O_{j+1}I_j} \) is a valid process representing a CPT map from \( \mathcal{H}^{O_{j+1}I_j} \) to \( \mathcal{H}^{I_{j+1}L_j} \). Since the complete reverse process formed by

\[ \bar{W} := I^{O_1} \otimes W^{O_2I_1} \otimes ... \otimes W^{O_{K-1}I_K} \otimes I^K, \quad (22) \]

clearly factorizes over the reversed DAG \( G^* \), this completes the proof.