

# Can a quantum state over time resemble a quantum state at a single time?

Dominic Horsman\*    Chris Heunen†    Matthew F. Pusey‡  
Jonathan Barrett§    Robert W. Spekkens‡

July 13, 2016

## Abstract

Standard quantum theory represents a composite system at a given time by a joint state, but it does not prescribe a joint state for a composite of systems at different times. If a more even-handed treatment of space and time is possible, then such a joint state should be definable, and one might expect it to satisfy the following five conditions: that it is a Hermitian operator on the tensor product of the single-time Hilbert spaces; that it represents probabilistic mixing appropriately; that it has the appropriate classical limit; that it has the appropriate single-time marginals; that composing over multiple time-steps is associative. We show that no construction satisfies all these requirements. If an even-handed treatment of space and time is possible, therefore, one or more axioms must be dropped. In particular, if Hermiticity is dropped, then we show that the construction is fixed uniquely up to an ordering convention.

## 1 Introduction

Quantum theory as usually formalised contains a fundamental asymmetry between space and time. States of a system are defined at a single time, and evolve over time under the action of a Hamiltonian. This goes against physical intuitions from relativity, giving rise to the so-called ‘problem of time’ in constructing quantum theories of gravity [1]. It is also in marked contrast with

---

\*Department of Physics, Durham, University, United Kingdom; supported by an FQXi Large Grant, the NQIT Quantum Hub, and EPSRC grant EP/L022303/1

†School of Informatics, University of Edinburgh, Scotland; supported by EPSRC Fellowship EP/L002388/1

‡Perimeter Institute for Theoretical Physics, Waterloo, Canada; Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

§Department of Computer Science, University of Oxford, United Kingdom; supported by an FQXi Large Grant, and the NQIT Quantum Hub

classical probability theory, which can assign joint probabilities to events independent of their spatio-temporal relationships [2]. If the apparent asymmetry between space and time in quantum theory is fundamental, then the understanding of time given by quantum theory must be different from that of both relativity and classical mechanics. If, on the other hand, it is not fundamental then it should be removable. One way of removing the asymmetry would be to construct quantum states for composites over time. States would then be defined across both space and time, without a separate formalism describing evolution over time.

The question of how to build a quantum state over time concerns the *composition* of quantum systems. The conventional quantum formalism specifies how to represent the joint state of different systems existing at a given time: the state is given by a density matrix acting on the tensor product of the Hilbert spaces associated with the component systems. This is the case regardless of whether the components are spatially separated systems or different degrees of freedom of a single system. But there is another sort of composite system besides these, namely one where the components exist at different times.

There have been a variety of proposals for expressing quantum theory in a manner that treats space and time in a more even-handed fashion. The sum-over-histories approach to quantum dynamics [3, 4] and the program of consistent/decoherent histories [5, 6, 7] are examples. More recently, there have been a family of proposals with a distinctively operational flavour, such as the multi-time formalism of Aharonov *et al.* [8], and various proposals that focus on causal structure [9, 10, 11, 12, 13].

In this paper, we investigate the possibility of supplementing standard quantum theory with a definition of joint states over time that closely matches the definition for composite systems at a given time. We start by considering a few options for defining the state of a composite over time as an operator on the tensor product of the Hilbert spaces associated to each temporally localized component. Specifically, we explore the proposal of Leifer and Spekkens [11] which is inspired by the view that quantum theory is a generalisation of classical probability theory, the proposal of Fitzsimons, Jones and Vedral [12], and a novel proposal based on discrete Wigner functions. By a close examination of these different options, we distill five criteria that the state of a composite over time should satisfy if it is to have the essential qualities of the state of a composite at a single time. Our main theorem states that there is no way to construct a state over time that satisfies all of these criteria. We further show that if one of the criteria is dropped – namely, the criterion that the state over time is a Hermitian operator – then the only construction satisfying the remaining four criteria corresponds to straightforward matrix multiplication of the initial state and the operators representing the quantum channels that evolve from one time-slice to the next.

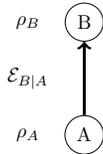


Figure 1: Two quantum systems  $A$  and  $B$ , and a channel between them. A potential quantum state over time  $\rho_{AB}$  should be a function of these variables.

## 2 States over time

We begin with the simplest situation in which to discuss the construction of a quantum state over time: a single system considered at two different times, as depicted in Fig. 1. Whereas Hilbert spaces are conventionally associated with systems that persist over time, we shall here associate Hilbert spaces to quantum systems that are considered to be localized to a given region of space-time (similar in spirit to ‘events’ in relativity). The system at the first time is denoted  $A$ , and that at the second time  $B$ . The most general possible evolution of the system corresponds to a *quantum channel*, that is, a trace-preserving completely positive map, denoted  $\mathcal{E}_{B|A}$ . Conventionally, the quantum state of  $A$  is represented by an operator acting on  $\mathcal{H}_A$ , denoted  $\rho_A$ , and the quantum state of  $B$  is represented by an operator acting on  $\mathcal{H}_B$ , denoted  $\rho_B$ . The state of  $B$  can be determined by the state of  $A$  and the channel, via  $\rho_B = \mathcal{E}_{B|A}(\rho_A)$ .

It is natural to consider the possibility of a quantum state for the composite system  $AB$ , represented by an operator  $\rho_{AB}$  acting on  $\mathcal{H}_A \otimes \mathcal{H}_B$ . The scenario of Fig. 1 is fully defined by the state  $\rho_A$  and the channel  $\mathcal{E}_{B|A}$  (as these determine the output state  $\rho_B$ ), so we assume that the state over time is a function of these:

$$\rho_{AB} = f(\rho_A, \mathcal{E}_{B|A}). \quad (1)$$

We restrict to the finite-dimensional case, so that  $\rho_{AB}$  can be represented by a square matrix, but we do not require that  $\rho_{AB}$  is a standard density matrix. The properties that should be required of the function  $f$ , and the possible forms  $f$  can take, are the main questions addressed in this paper.

### 2.1 Three possible definitions

It will be useful in the following to represent the channel  $\mathcal{E}_{B|A}$  by an operator acting on  $\mathcal{H}_A \otimes \mathcal{H}_B$ , denoted  $E_{B|A}$ . As it is also useful if the representation of the channel is basis-independent, we choose it to be the operator that is Jamiołkowski-isomorphic [14] to  $\mathcal{E}_{B|A}$ , defined as

$$E_{B|A} \equiv \sum_{ij} \mathcal{E}_{B|A}(|i\rangle\langle j|_A) \otimes |j\rangle\langle i|_A,$$

where  $\{|i\rangle\}$  is any orthonormal basis of  $\mathcal{H}_A$ . (Up to normalisation, this operator is the partial transpose of the operator that is Choi-isomorphic [15] to  $\mathcal{E}_{B|A}$ .)

The operator  $E_{B|A}$  is independent of the basis used in its definition, and the state of the output of the channel can be expressed in terms of it as [11]:

$$\rho_B = \text{Tr}_A(E_{B|A}\rho_A).$$

We call  $E_{B|A}$  the *channel state*.

The joint state  $\rho_{AB}$  can be understood as being given by a binary operation  $\star$  that takes a pair of operators on  $\mathcal{H}_A \otimes \mathcal{H}_B$  to an operator on  $\mathcal{H}_A \otimes \mathcal{H}_B$ :

$$\star: (E_{B|A}, \rho_A \otimes \mathbf{1}_B) \mapsto \rho_{AB}. \quad (2)$$

We call this operation a *star product*. For notational convenience, we will in the following often suppress tensor products with identity, writing for example  $\rho_A$  instead of  $\rho_A \otimes \mathbf{1}_B$ , so that

$$\rho_{AB} = E_{B|A} \star \rho_A. \quad (3)$$

There have been a number of previous attempts to define a joint state  $\rho_{AB}$  along these lines. The proposal due to Leifer and Spekkens (LS) constructs joint states over time in the context of a programme of reframing quantum theory as a theory of Bayesian inference [11], which builds on [16] and [17]. Just as a density matrix on a composite of  $A$  and  $B$ ,  $\rho_{AB}$ , can be understood to be a quantum analogue of a joint probability distribution,  $P(XY)$ , LS observe that a channel state  $E_{B|A}$  can be understood as a quantum analogue of a conditional probability distribution,  $P(Y|X)$ , which describes a classical channel with input  $X$  and output  $Y$ . For example, the analogue of the normalization condition for conditional probabilities,  $\sum_Y P(Y|X) = 1$ , is  $\text{Tr}_B(E_{B|A}) = \mathbf{1}_A$ , and the analogue of the fact that the output of a classical channel satisfies  $P(Y) = \sum_X P(Y|X)P(X)$  is the fact that the channel state satisfies  $\rho_B = \text{Tr}_A(E_{B|A}\rho_A)$ . In the classical case, there is no obstacle to defining a joint probability distribution for the pair of systems constituting the input and the output of a channel via  $P(XY) = P(Y|X)P(X)$ . In the quantum case, however, the operators  $E_{B|A}$  and  $\rho_A \otimes \mathbf{1}_B$  do not in general commute, so it is not obvious what should replace this equation. LS considered the following option:

$$\rho_{AB}^{(\text{LS})} = E_{B|A} \star_{\text{LS}} \rho_A \equiv \rho_A^{\frac{1}{2}} E_{B|A} \rho_A^{\frac{1}{2}}, \quad (4)$$

where the right-hand side is a shorthand notation for  $(\rho_A^{\frac{1}{2}} \otimes \mathbf{1}_B) E_{B|A} (\rho_A^{\frac{1}{2}} \otimes \mathbf{1}_B)$ . The operator  $\rho_{AB}^{(\text{LS})}$  is Hermitian, and although it is not in general positive, it is locally positive, meaning that  $\langle ab | \rho_{AB}^{(\text{LS})} | ab \rangle \geq 0$  for all  $|a\rangle$  in  $\mathcal{H}_A$  and  $|b\rangle$  in  $\mathcal{H}_B$ . Significant problems with this proposal were already noted by LS in Sec. VII of [11].

An alternative proposal for a state over time is that of Fitzsimons, Jones, and Vedral (FJV) [12]. As it stands, the proposal is for multi-qubit systems, rather than systems of arbitrary dimension. In this work, it will suffice to consider the FJV construction for the special case in which  $A$  and  $B$  are qubits. In order to define the FJV joint state, imagine that a Pauli measurement  $\sigma_i$

is performed on  $A$ , with outcome  $\pm 1$  (here, the Pauli operators are denoted  $\sigma_0 = \mathbf{1}$ ,  $\sigma_1 = \sigma_x$ ,  $\sigma_2 = \sigma_y$ ,  $\sigma_3 = \sigma_z$ ). After this measurement, the state of  $A$  is imagined to be updated according to the projection postulate, before evolving according to the channel  $\mathcal{E}_{B|A}$ , and a Pauli measurement  $\sigma_j$  is then performed on  $B$ . Let  $\langle \sigma_i \otimes \sigma_j \rangle$  denote the expectation value of the product of the outcomes of the two measurements. The FJV joint state is given by

$$\rho_{AB}^{(\text{FJV})} \equiv \frac{1}{4} \left( \sum_{i,j=0}^3 \langle \sigma_i \otimes \sigma_j \rangle \sigma_i \otimes \sigma_j \right). \quad (5)$$

Note that although the FJV joint state is defined by reference to measurements that one imagines to be performed on  $A$  and on  $B$ , the joint state is intended as a description of the system  $AB$  in the case where no measurements are made on either, but rather  $A$  simply evolves into  $B$  according to the given channel. The matrix  $\rho_{AB}^{(\text{FJV})}$  is Hermitian, but not necessarily positive (nor even necessarily locally positive). It is clear that  $\rho_{AB}^{(\text{FJV})}$  depends only on  $E_{B|A}$  and  $\rho_A$ , hence can in principle be written in the form of Eq. (3), for a suitable definition of the star product. In fact, as we show in Appendix A, the FJV construction for two qubits corresponds to the Jordan product:

$$\rho_{AB}^{(\text{FJV})} = E_{B|A} \star_{\text{FJV}} \rho_A = \frac{1}{2} (\rho_A E_{B|A} + E_{B|A} \rho_A). \quad (6)$$

For a single qubit considered at an arbitrary number of time steps, the FJV construction is similar, involving imagined Pauli measurements performed at each time step, and coincides with an iterated application of the Jordan product.

Finally, one might also hope to construct a state over time in a manner that is analogous to the definition of a joint probability distribution, but where probabilities are replaced by the *quasi-probabilities* appearing in a discrete Wigner representation of the states and channels [18, 19]. A discrete Wigner (W) representation for a  $d$ -dimensional system  $A$  is defined by a set  $\Omega^A = \{K_i^A\}$  of *phase-point operators* which form a basis for the space of operators on  $\mathcal{H}_A$ , satisfying  $\text{Tr}(K_i^A K_j^A) = d\delta_{ij}$  and  $\sum_i K_i^A = d\mathbf{1}_A$ , hence  $\text{Tr}(K_i^A) = 1$ . A density matrix  $\rho_A$  can be written  $\rho_A = \sum_i r^A(i) K_i^A$ , where  $r^A(i)$  is a real-valued function on  $\Omega^A$ , with  $\sum_i r^A(i) = 1$ . Hence a system  $A$  can be described by the function  $r^A(i) \in [-1, 1]$ , which has the form of a quasi-probability distribution on a discrete phase space. Similarly, the operator  $E_{B|A}$  representing a channel can be written  $E_{B|A} = \frac{1}{d} \sum_{ij} r^{B|A}(j|i) K_i^A \otimes K_j^B$ , where  $r^{B|A}(j|i)$  is a real-valued function on  $\Omega^A \times \Omega^B$ . The extra factor of  $1/d$  is introduced so that  $r^{B|A}(j|i)$  satisfies  $\sum_j r^{B|A}(j|i) = 1$ ; these can therefore be thought of as conditional quasi-probabilities. The natural definition for the W representation of the state of the composite  $AB$  is then  $r^{AB}(ij) \equiv r^{B|A}(j|i)r^A(i)$ . This implies that the state over time is represented by the operator

$$\rho_{AB}^{(\text{W})} \equiv \sum_{ij} r^{B|A}(j|i) r^A(i) K_i^A \otimes K_j^B, \quad (7)$$

which, when expressed explicitly as a function of  $\rho_A$  and  $E_{B|A}$ , has the form

$$\rho_{AB}^{(W)} = E_{B|A} \star_W \rho_A \equiv \frac{1}{d^2} \sum_{ij} \text{Tr}_{AB}(E_{B|A} K_i^A \otimes K_j^B) \text{Tr}_A(\rho_A K_i^A) K_i^A \otimes K_j^B. \quad (8)$$

There are various ways, then, in which one might aim to define quantum states over time. This leads us to pose the question: can we isolate specific axioms or desiderata that a state over time ought to satisfy? Furthermore, given those axioms, is it possible to find a construction that satisfies them all?

### 3 Five criteria for a star product

We introduce five basic axioms for a star product, motivating each by a corresponding desideratum for the properties of a state over time.

**Hermiticity** We assume that, for a finite-dimensional system considered at  $k$  distinct time steps, the state over time,  $\rho_{A_1 \dots A_k}$ , is a Hermitian operator on the  $k$ -fold tensor product of copies of its Hilbert space,  $\mathcal{H}_{A_1} \otimes \dots \otimes \mathcal{H}_{A_k}$ . We denote the space of such operators by  $\mathbb{H}_{A_1 \dots A_k}$ , so that  $\rho_{A_1 \dots A_k} \in \mathbb{H}_{A_1 \dots A_k}$ . In particular, if  $A$  and  $B$  denote a system at two times, then  $\rho_{AB} \in \mathbb{H}_{AB}$ .

Given that  $\rho_A \otimes \mathbf{1}_B \in \mathbb{H}_{AB}$  and  $E_{B|A} \in \mathbb{H}_{AB}$ , it follows that we are assuming that the star product of Eq. (2) is a map of the form:

$$\star: \mathbb{H}_{AB} \times \mathbb{H}_{AB} \rightarrow \mathbb{H}_{AB}.$$

As mentioned, all three constructions, LS, FJV, and W, satisfy the assumption of Hermiticity. Below, we consider the consequences of dropping it.

**Preservation of probabilistic mixtures** The output state of a quantum channel for a given input state is a linear function of both the input state and the channel. This is necessary to ensure that a probabilistic mixture of input states maps under the channel to the corresponding mixture of output states, and that a probabilistic mixture of channels acting on a given input state yields the corresponding mixture of output states. By analogy, we would like the joint state over time for a probabilistic mixture of input states to be the corresponding mixture of the joint states for each input state, and similarly for probabilistic mixtures of channels.

Consider as an example the situation in Fig. 1, but where the quantum system  $A$  is conditioned on a classical variable  $x$ . For instance, it could be that  $x$  is the outcome of a fair coin toss, and a qubit  $A$  is prepared in a state that depends on this outcome:  $\rho_{A,x=h} \equiv |0\rangle\langle 0|$  for heads, and  $\rho_{A,x=t} \equiv |1\rangle\langle 1|$  for tails. If the channel between  $A$  and  $B$  is the identity channel, then  $E_{B|A}^{(\text{id})} \equiv 2|\phi^+\rangle\langle\phi^+|^{T_B}$  where  $|\phi^+\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle)$  and where  $T_B$  denotes the partial transpose on  $B$ . In this case, we would like the star product to satisfy

$$E_{B|A}^{(\text{id})} \star \left( \frac{1}{2} \rho_{A,x=h} + \frac{1}{2} \rho_{A,x=t} \right) = \frac{1}{2} \left( E_{B|A}^{(\text{id})} \star \rho_{A,x=h} \right) + \frac{1}{2} \left( E_{B|A}^{(\text{id})} \star \rho_{A,x=t} \right). \quad (9)$$

In general, we assume:

$$\text{Convex-bilinearity: } \begin{cases} (px + (1-p)y) \star z = p(x \star z) + (1-p)(y \star z), \\ x \star (py + (1-p)z) = p(x \star y) + (1-p)(x \star z). \end{cases}$$

While the FJV and W formulations are both straightforwardly convex-bilinear, there is a failure of this axiom in the case of the LS star product, which does not preserve convex combinations in  $\rho_A$ . For example, using the classical coin toss example, it is easy to show that the LS star product yields the mixture  $\frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$  for the right-hand side of (9), but the non-separable  $|\phi^+\rangle\langle\phi^+|^{TB}$  for the left-hand side.

**Preservation of classical limit** In classical probability theory, defining joint probabilities for events at different times is unproblematic. This gives a constraint on the form of quantum temporal joint states, if it is assumed that classical joint probabilities must be reproduced when a quantum system is behaving entirely classically. For the situation of Fig. 1, this is the case if both the input quantum state and the operator corresponding to the channel are diagonal in the same basis. Consider for example, a channel that is completely dephasing in some basis, and an input state that is diagonal in the same basis. For some probabilities  $p(i)$  and conditional probabilities  $p(j|i)$ , the input state can be written  $\rho_A = \sum_i p(i) |i\rangle\langle i|_A$ , and the behaviour of the channel is given by

$$\mathcal{E}_{B|A}(\rho_A) = \sum_i \langle i|_A \rho_A |i\rangle_A \sum_j p(j|i) |j\rangle\langle j|_B. \quad (10)$$

For this channel,

$$E_{B|A} = \sum_{i,j} p(j|i) |i\rangle\langle i|_A \otimes |j\rangle\langle j|_B, \quad (11)$$

which is simply a matrix encoding of the classical conditional probabilities that define the channel. The classical joint probabilities would be given by  $p(i, j) = p(i)p(j|i)$ , which, when encoded in a matrix in the same way, are given by

$$\rho_{AB} = \sum_{i,j} p(j|i)p(i) |i\rangle\langle i|_A \otimes |j\rangle\langle j|_B = \rho_A E_{B|A}. \quad (12)$$

In order to reproduce classical probability theory, therefore, when states and channel states are simultaneously diagonalisable the joint state over time should be given by their matrix product. This is ensured if

$$[E_{B|A}, \rho_A] = 0 \implies E_{B|A} \star \rho_A = E_{B|A} \rho_A. \quad (13)$$

Hence we assume that a star product satisfies:

$$\text{Product on commuting pairs : } [x, y] = 0 \implies x \star y = xy.$$

Note that an implication of this condition is that composition with the identity satisfies

$$x \star \mathbf{1} = \mathbf{1} \star x = \mathbf{1}x = x, \quad (14)$$

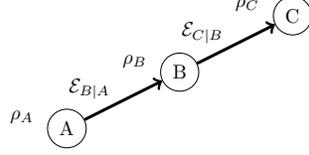


Figure 2: Three localized regions,  $A$ ,  $B$ , and  $C$ , the state on  $A$ ,  $\rho_A$ , the channel from  $A$  to  $B$ ,  $\mathcal{E}_{B|A}$ , and the channel from  $B$  to  $C$ ,  $\mathcal{E}_{C|B}$ .

since  $[x, \mathbf{1}] = 0$  for all operators  $x$ .

It is easy to see that the LS joint state satisfies  $\rho_{AB}^{LS} = E_{B|A}\rho_A$  when  $E_{B|A}$  and  $\rho_A$  commute. This is also the case for the FJV joint state for a qubit at two times, as is evident from the form of Eq. (6). However, the W construction does not satisfy this axiom. Consider, for example, an input state that is a pure state of a qubit, and a completely dephasing channel that preserves this state. Direct calculation shows that unless the input state is an eigenstate of one of the Pauli operators, the Wigner joint state  $\rho_{AB}^{(W)} \neq E_{B|A}\rho_A$ .

**Preservation of marginal states** An obvious desideratum for a joint state  $\rho_{AB}$  is that it returns the correct marginal states for  $A$  and  $B$ , i.e.,  $\text{Tr}_B \rho_{AB} = \rho_A$  and  $\text{Tr}_A \rho_{AB} = \rho_B$ . Given that  $\rho_{AB} = E_{B|A} \star \rho_A$  by definition, and that  $\rho_B = \text{Tr}_A(\rho_A E_{B|A})$ , as noted earlier, it follows that we require  $\text{Tr}_A(E_{B|A} \star \rho_A) = \text{Tr}_A(\rho_A E_{B|A})$ , hence also

$$\text{Tr}_{AB}(E_{B|A} \star \rho_A) = \text{Tr}_{AB}(\rho_A E_{B|A}).$$

This motivates the fourth axiom:

$$\text{Product when traced: } \text{Tr}(x \star y) = \text{Tr}(xy).$$

All constructions considered in this work satisfy this requirement.

**Compositionality** The final axiom for the star product concerns how it composes when we wish to describe a system at more than two distinct times. Consider the situation of Fig. 2, with three localized regions, labelled  $A$ ,  $B$  and  $C$ , joined by two channels,  $\mathcal{E}_{B|A}$  and  $\mathcal{E}_{C|B}$ .

There are two ways in which one can imagine forming the joint state  $\rho_{ABC}$  on the composite of the three regions. The first way proceeds as follows: form the joint state over the first two regions,  $\rho_{AB}$ ; define a novel kind of channel, denoted  $\mathcal{E}_{C|AB}$ , whose input is the composite of the first two regions and whose output is the third region; finally, take the star product of  $\rho_{AB}$  with  $E_{C|AB}$ , the operator corresponding to  $\mathcal{E}_{C|AB}$ :

$$\rho_{ABC} \equiv E_{C|AB} \star \rho_{AB}. \quad (15)$$

The second way proceeds as follows: define a novel kind of channel, denoted  $\mathcal{E}_{BC|A}$ , that has the first region as input and the composite of the second and

the third regions as output; next, take the star product of  $\rho_A$  with  $E_{BC|A}$ , the operator corresponding to this channel:

$$\rho_{ABC} \equiv E_{BC|A} \star \rho_A. \quad (16)$$

We now consider the consequences, for both approaches, of the fact that the only dependence that  $C$  has on  $A$  is mediated by  $B$ .

In the first approach, this implies that the channel  $\mathcal{E}_{C|AB}$  is simply  $\mathcal{E}_{C|B}$ , so that  $E_{C|AB} = E_{C|B} \otimes \mathbb{1}_A$ . Using this, together with the fact that  $\rho_{AB} = E_{B|A} \star \rho_A$ , Eq. (15) becomes

$$\rho_{ABC} \equiv E_{C|B} \star (E_{B|A} \star \rho_A), \quad (17)$$

where we have again adopted our convention of suppressing identity operators. In the second approach, this fact makes it natural to define the channel state for  $\mathcal{E}_{BC|A}$  via the star product, as  $E_{BC|A} \equiv E_{C|B} \star E_{B|A}$ , which one can understand as a quantum analogue of the Markov condition for classical causal networks [2]. We have here allowed a channel state to appear as the second argument of the star product. This is no more objectionable than allowing a state over time to appear as the second argument, as we did in Eq. (15), because for every channel state  $E_{B|A}$ , there is a state over time  $\rho_{AB}$  such that  $E_{B|A} = d_A \rho_{AB}$ , where  $d_A$  is the dimension of the Hilbert space of  $A$ . To see this, we need only note that for channel state  $E_{B|A}$  and input state  $\frac{1}{d_A} \mathbb{1}_A$ , the state over time is  $\rho_{AB} = E_{B|A} \star \frac{1}{d_A} \mathbb{1}_A$  and by our assumption of being product on commuting pairs, this reduces to  $\rho_{AB} = \frac{1}{d_A} E_{B|A}$ , so that (invoking convex-bilinearity) we can write  $E_{C|B} \star E_{B|A}$  as  $d_A (E_{C|B} \star \rho_{AB})$ .

Using  $E_{BC|A} \equiv E_{C|B} \star E_{B|A}$ , Eq. (16) becomes

$$\rho_{ABC} \equiv (E_{C|B} \star E_{B|A}) \star \rho_A. \quad (18)$$

Our last condition on the star product is motivated by the idea that the two techniques just described for defining a joint state over three times should give the same result,

$$E_{C|B} \star (E_{B|A} \star \rho_A) = (E_{C|B} \star E_{B|A}) \star \rho_A. \quad (19)$$

This can be ensured by taking the star product to be associative (note that the considerations above merely motivate associativity but do not imply it). Our final condition, therefore, is

$$\textbf{Associativity} : x \star (y \star z) = (x \star y) \star z.$$

The W formulation defines an associative star product. The LS star product, however, is not associative, as noted by LS in Sec. VII of [11]. In particular, the second argument of the LS star product is required to be a positive operator—in order for the square root of the operator to be uniquely defined—and yet nonpositive operators arise as joint states over time. Indeed, the fact that the LS star product does not support composition across time was highlighted by LS

	Hermitian?	Convex-bilinear?	Product on commuting pairs?	Product when traced?	Associative?
LS	✓	×	✓	✓	×
FJV	✓	✓	✓	✓	×*
W	✓	✓	×	✓	✓

Table 1: Satisfaction of axioms by star products of the three constructions of states over time. (\*see text for details)

as one of its most significant shortcomings, motivating the search for alternative approaches.

The FJV construction is interesting with regard to associativity. One can verify that the FJV construction for a single qubit at three times gives a joint state:

$$\rho_{ABC}^{(\text{FJV})} = E_{C|B} \star_{\text{FJV}} (E_{B|A} \star_{\text{FJV}} \rho_A),$$

where, as noted previously,  $\star_{\text{FJV}}$  is the Jordan product. Although the Jordan product is non-associative in general, it is associative for triples of operators of the form  $E_{C|B} \otimes \mathbf{1}_A$ ,  $E_{B|A} \otimes \mathbf{1}_B$ , and  $\rho_A \otimes \mathbf{1}_B \otimes \mathbf{1}_C$ , that is, (suppressing identities as usual)

$$E_{C|B} \star_{\text{FJV}} (E_{B|A} \star_{\text{FJV}} \rho_A) = (E_{C|B} \star_{\text{FJV}} E_{B|A}) \star_{\text{FJV}} \rho_A. \quad (20)$$

The reason is that  $\rho_A \otimes \mathbf{1}_B \otimes \mathbf{1}_C$  commutes with  $E_{C|B} \otimes \mathbf{1}_A$ . These observations extend to a single qubit considered over an arbitrary number of time steps. Hence although the Jordan product is non-associative, it *is* associative when restricted to operators of the correct form to describe a quantum system passing through a sequence of channels. Hence a unique joint state is defined without privileging any particular way of grouping the different time steps. When one considers an arbitrary number of qubits, the situation becomes more subtle: an FJV star product can again be inferred, but in this case it is more complicated than the Jordan product of Eq. (6). A similar situation arises for systems whose dimension is a power of a prime. For other systems, the proposal is ambiguous. We omit the details as these issues are non-essential for our purposes.

## 4 Main theorem

We now have five criteria for states across time, expressed as axioms for a star product: Hermiticity, convex-bilinearity, product when commuting, product when traced, and associativity. The satisfaction or otherwise of these axioms, for each of LS, FJV, and the W construction are summarised in Table 1. None of these constructions satisfies all of the axioms.

**Theorem 1.** *Let  $\mathbb{H}_n$  be the set of  $n$ -by- $n$  Hermitian matrices. There is no function  $\star: \mathbb{H}_n \times \mathbb{H}_n \rightarrow \mathbb{H}_n$  satisfying convex-bilinearity, product on commuting pairs, product when traced, and associativity.*

In other words, there is no definition of a state over time that satisfies the five axioms given in this paper. We provide the proof of this result in Appendix B.

It is natural to consider what happens when one or more of the axioms is dropped.

Suppose that Hermiticity is not assumed, and that the star product is assumed to be a map  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$ , where  $\mathbb{M}_n$  is the complex vector space of all  $n$ -by- $n$  matrices. In this case, as we also show in Appendix B, the remaining four axioms force the following form of the star product.

**Theorem 2.** *Let  $\mathbb{M}_n$  be the complex vector space of  $n$ -by- $n$  matrices. The only functions  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$  satisfying convex-bilinearity, product on commuting pairs, product when traced, and associativity are  $x \star y = xy$  and  $x \star y = yx$ .*

In this case, there are only two possibilities for the star product, corresponding to straightforward (left and right) matrix multiplication. The possibility of taking the star product to be matrix multiplication was noted in [11] as one way to overcome the shortcomings of the LS star product. We comment on some of the consequences of such a choice in the discussion section.

If associativity is dropped, while the other four axioms (including Hermiticity) are maintained, then the FJV construction provides an interesting example of a means to construct states over time.

Finally, in Appendix C, we also consider the possibility that the result of the star product is a matrix in a larger (or smaller) state space, with

$$\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_m \quad m \neq n.$$

With suitable refinements for the axioms of *product on commuting pairs* and *associativity*, the conclusion of the theorem remains essentially the same.

## 5 Discussion

We have considered methods for constructing quantum states over time, focussing on simple cases of spatio-temporally localized quantum systems evolving through a sequence of quantum channels. We have identified a set of physically-motivated criteria for such a construction, and Theorem 1 shows that there is no construction that satisfies all of the criteria.

Is there a way that the apparent no-go result of Theorem 1 can be avoided? The most obvious way is to question whether the axioms that we have identified are themselves reasonable.

The axioms of *convex-bilinearity* and of *product when traced* seem to be essential if the joint state is to play a role that is analogous to the role played by the joint state of composite systems at a given time.

Reproducing classical joint probabilities when the initial state and the channel states all commute also seems essential, at least if the quantum states over time are to be seen as generalizations of classical probability distributions in the way that standard quantum density matrices are seen as generalizations of classical probability distributions. However, it is reasonable to judge a pairing of initial state and channel state to be classical only when they are jointly diagonalisable *in a basis that is a tensor product of a fixed basis for each system* rather than simply jointly diagonalisable. Therefore, the assumption that the joint state should be equivalent to a joint probability distribution for such pairs only warrants that the star product should reduce to the matrix product on certain pairs of commuting operators, rather than all pairs of commuting operators, as the condition asserts. Hence there may be some wiggle room in constructing joint states that reproduce classical joint probabilities in the desired sense but do not satisfy the axiom.

As noted in the previous section, if the state over time is not assumed to be Hermitian, then Theorem 2 shows that matrix multiplication is the only construction that satisfies the remaining axioms.

We also noted that if one does not assume associativity, then the FJV construction satisfies the other four axioms. Furthermore, although the FJV star product is not associative for arbitrary triples of operators, we have shown that on the restricted domain of triples of operators corresponding to a sequence of channel states, it *is* associative.

A disadvantage of the FJV construction as stated, however, is that it is defined only for collections of qubits. It can be applied to systems whose dimension is a power of 2 by regarding them as composite systems made up of qubits, but then there is an arbitrariness in the choice of how to factorize the larger Hilbert space into component systems. There is also a natural generalization of the proposal from a qubit to a system of prime dimension, but how to generalize the proposal to systems of arbitrary dimension remains ambiguous.

There is an interesting relationship between the FJV construction and that corresponding to matrix multiplication, at least for a single qubit considered at two times: the joint state  $\rho_{AB}^{(\text{FJV})}$  is a Hermitian operator whose eigenspectrum coincides with the positive and real eigenspectrum of  $\rho_{AB}^{(\text{MM})} = E_{B|A}\rho_A$ . That is,  $\rho_{AB}^{(\text{FJV})}$  can be obtained from  $\rho_{AB}^{(\text{MM})}$  by removing the latter's imaginary eigenvalues.

In seeking to define a state for a composite over time, we have here considered to what extent the formal properties of such a state could resemble those of a state for a composite at a given time. It would also be interesting to approach the problem from a purely operational point of view, where the aim is to define a joint state by reference to predictions for experimental interventions on systems at different times. For instance, it would be interesting to consider the relationships between the various constructions considered here and the predictions for the outcomes of weak measurements performed across multiple time steps [20].

In all of the proposals for defining a state over time that we have considered, the set of operators that can describe a joint state of a system considered at two

times is *distinct* from the set of operators that can describe a state of two copies of the system at a given time. This contrasts with classical probability theory, where the possible joint probability distributions on a pair of systems is the same in the two cases. This suggests that even if a sensible notion of a state over time can be defined, quantum theory still manifests an asymmetry between space and time that is not present in classical probability theory (analogous, perhaps, to the fact that relativity theory still distinguishes space and time through the signature of the metric). Indeed, this conclusion is supported by the results of [21], which demonstrate that in quantum theory, unlike classical theories, the set of correlations that can be generated by two systems that are related as cause and effect is different from the set that can be generated by two systems that are both effects of a common cause. Further investigations into operationally-motivated definitions of states over time are likely to shed additional light on the precise nature of this asymmetry.

## References

- [1] Chris J. Isham. *Integrable Systems, Quantum Groups, and Quantum Field Theories*, chapter Canonical Quantum Gravity and the Problem of Time, pages 157–287. Springer Netherlands, Dordrecht, 1993.
- [2] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, NY, USA, 2000.
- [3] Richard P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Reviews of Modern Physics*, 20(2):367, 1948.
- [4] Rafael D. Sorkin. Forks in the road, on the way to quantum gravity. *International Journal of Theoretical Physics*, 36(12):2759–2781, December 1997.
- [5] Robert B Griffiths. Consistent histories and the interpretation of quantum mechanics. *Journal of Statistical Physics*, 36(1-2):219–272, 1984.
- [6] Roland Omnes. Logical reformulation of quantum mechanics. i. foundations. *Journal of Statistical Physics*, 53(3-4):893–932, 1988.
- [7] Murray Gell-Mann and James B Hartle. Quantum mechanics in the light of quantum cosmology. *Complexity, entropy and the physics of information*, 8, 1990.
- [8] Yakir Aharonov, Sandu Popescu, Jeff Tollaksen, and Lev Vaidman. Multiple-time states and multiple-time measurements in quantum mechanics. *Physical Review A*, 79(5):052110, 2009.
- [9] Lucien Hardy. The operator tensor formulation of quantum theory. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 370(1971):3385–3417, 2012.

- [10] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paulo Perinotti. Quantum circuit architecture. *Phys. Rev. Lett.*, 101:060401, 2008.
- [11] Matthew S. Leifer and Robert W. Spekkens. Towards a formulation of quantum theory as a causality neutral theory of Bayesian inference. *Physical Review A*, 88(5):052130, 2013.
- [12] Joseph Fyzsimons, Jonathan A. Jones, and Vladko Vedral. Quantum correlations which imply causation. *arXiv:1302.2731*, 2013.
- [13] Ognjan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature Communications*, 3:1092, 2012.
- [14] A. Jamiolkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.*, 3:275–278, 1972.
- [15] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10:285–290, 1975.
- [16] Matthew S Leifer. Quantum dynamics as an analog of conditional probability. *Physical Review A*, 74(4):042310, 2006.
- [17] Matthew S. Leifer and David Poulin. Quantum graphical models and belief propagation. *Annals of Physics*, 323(8):1899–1946, 2008.
- [18] William K. Wootters. A Wigner-function formulation of finite-state quantum mechanics. *Annals of Physics*, 176:1–21, 1987.
- [19] David Gross. Hudson’s theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, 2006.
- [20] Shmuel Marcovitch and Benni Reznik. Structural unification of space and time correlations in quantum theory. *arXiv:1103.2557*, 2011.
- [21] Katja Ried, Megan Agnew, Lydia Vermeyden, Dominik Janzing, Robert W. Spekkens, and Kevin J. Resch. A quantum advantage for inferring causal-structure. *Nature Physics*, 11(5):414–420, March 2015.
- [22] Chris Heunen and Clare Horsman. Matrix multiplication is determined by orthogonality and trace. *Linear Algebra and its Applications*, 439(12):4130–4134, 2013.
- [23] Bent Fuglede. A commutativity theorem for normal operators. *Proceedings of the National Academy of Sciences of the USA*, 36(1):35–40, 1950.

## A The FJV construction as a star product

In this appendix, we show that the Fitzsimons-Jones-Vedral construction (5) corresponds to the Jordan product  $\rho_{AB} = \frac{1}{2} (\rho_A E_{B|A} + E_{B|A} \rho_A)$  for a qubit at two times. Recall that the FJV construction defines a state of two systems as

$$\rho_{AB}^{(\text{FJV})} = \frac{1}{4} \left( \sum_{i=0}^3 \langle \sigma_i \otimes \sigma_j \rangle \sigma_i \otimes \sigma_j \right). \quad (21)$$

The expectations are defined operationally with respect to measurements of the Pauli operators with the projection postulate as the state update rule. Thus,  $\langle \sigma_i \otimes \sigma_j \rangle$  is read for a state over time as the expectation value of the product of outcomes of a projective measurement of  $\sigma_i$  followed by a measurement of  $\sigma_j$  (where, after the first measurement,  $A$  is in the eigenstate of  $\sigma_i$  corresponding to the outcome of that measurement).

Using the fact that the state at the output of the channel is given by  $\text{Tr}_A(E_{B|A}\rho_A)$ , we have

$$\begin{aligned} \langle \sigma_i \otimes \sigma_j \rangle &= \text{Tr} (E_{B|A}(\rho_A^{i+} \otimes \sigma_j)) - \text{Tr} (E_{B|A}(\rho_A^{i-} \otimes \sigma_j)) \\ &= \text{Tr} (E_{B|A}((\rho_A^{i+} - \rho_A^{i-}) \otimes \sigma_j)), \end{aligned} \quad (22)$$

where  $\rho_A^{i+}$  is the subnormalized state of  $A$  after projection onto the  $+1$  eigenspace of  $\sigma_i$ , and similarly for  $\rho_A^{i-}$ , i.e.,  $\rho_A^{i\pm} \equiv P^{i\pm} \rho_A P^{i\pm}$  where  $P^{i\pm} \equiv (\mathbf{1} \pm \sigma_i)/2$ .

Noting that

$$\begin{aligned} \rho_A^{i+} - \rho_A^{i-} &= \frac{1}{4} ((\mathbf{1} + \sigma_i)\rho_A(\mathbf{1} + \sigma_i) - (\mathbf{1} - \sigma_i)\rho_A(\mathbf{1} - \sigma_i)) \\ &= \frac{1}{2}(\sigma_i\rho_A + \rho_A\sigma_i), \end{aligned} \quad (23)$$

we obtain

$$\begin{aligned} \langle \sigma_i \otimes \sigma_j \rangle &= \frac{1}{2} \text{Tr} (E_{B|A}((\sigma_i\rho_A + \rho_A\sigma_i) \otimes \sigma_j)) \\ &= \frac{1}{2} \text{Tr} (\rho_A E_{B|A}(\sigma_i \otimes \sigma_j) + E_{B|A}\rho_A(\sigma_i \otimes \sigma_j)), \end{aligned} \quad (24)$$

where in the last equality we have used the cyclic property of the trace.

Eq. (21) implies that

$$\langle \sigma_i \otimes \sigma_j \rangle = \text{Tr} \left( \rho_{AB}^{(\text{FJV})} (\sigma_i \otimes \sigma_j) \right), \quad (25)$$

Combining Eqs. (24) and (25), and using the fact that the expectation values of products of Paulis uniquely define an operator, we conclude that  $\rho_{AB}^{(\text{FJV})} = \frac{1}{2}(\rho_A E_{B|A} + E_{B|A}\rho_A)$  as claimed.

## B Derivation of main theorems

Key to the derivation of Theorems 1 and 2 is the following result [22, Theorem 1]:

A function  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$  satisfies (A), (B), (I), (T) and (O) if and only if either  $x \star y = xy$  for all  $x, y \in \mathbb{M}_n$ , or  $x \star y = yx$  for all  $x, y \in \mathbb{M}_n$ .

Here, condition (A) is precisely our associativity condition and (T) is the product when traced axiom. Condition (I) is the identity composition, Eq. (14) above, and condition (O) is that

$$x \star y = 0 \text{ when } xy = yx = 0, \quad xx = x, \text{ and } yy = y, \text{ for } x, y \in \mathbb{M}_n.$$

Both the condition (I) and the condition (O) follow straightforwardly from the product on commuting pairs axiom. The one remaining condition, (B), is bilinearity over the complex numbers. On the face of it this is slightly stronger than our axiom of convex-bilinearity. Nonetheless, Theorems 1 and 2 follow from [22, Theorem 1], along with the following lemma.

**Lemma 3.** *Let  $\mathbb{H}_n$  be the set of  $n$ -by- $n$  Hermitian matrices, and  $\mathbb{M}_n$  the set of all  $n$ -by- $n$  complex matrices. Any convex-bilinear function  $\star: \mathbb{H}_n \times \mathbb{H}_n \rightarrow \mathbb{H}_n$  satisfying product on commuting pairs, product when traced, and associativity, extends uniquely to a complex bilinear function  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$ , which also satisfies product on commuting pairs, product when traced, and associativity.*

Theorem 2 follows from [22, Theorem 1], since if the star product is merely convex-bilinear and satisfies the other requirements, then it follows from the lemma that it must in fact be complex bilinear and still satisfy the other requirements. Theorem 1 follows since if there were a function on  $\mathbb{H}_n \times \mathbb{H}_n$  satisfying the five axioms, then by the lemma it would extend to a function on  $\mathbb{M}_n \times \mathbb{M}_n$  that satisfies (A), (B), (I), (T) and (O), hence could only be left or right matrix multiplication: but matrix multiplication does not preserve hermiticity.

We now provide the proof of lemma 3.

*Proof.* We first show that any convex-bilinear function  $\star: \mathbb{H}_n \times \mathbb{H}_n \rightarrow \mathbb{H}_n$  satisfying the other axioms is in fact bilinear over the reals. It follows from the product on commuting pairs axiom that  $x \star 0 = 0 \star x = 0$  where 0 here denotes the  $n$ -by- $n$  matrix all of whose components are 0. By considering convex combinations of the form  $x = rx' + (1-r)0$  (for  $0 \leq r \leq 1$ ), or  $x = (1/r)x' + (1-1/r)0$  (for  $r \geq 1$ ), it is then easy to show that  $(rx) \star y = r(x \star y)$  for any  $r \geq 0$ . Recall that a conic combination of  $x_i \in \mathbb{H}_n$  is  $\sum_i r_i x_i$  for nonnegative real weights  $r_i$ . Writing  $r = \sum_i r_i$  yields  $(\sum_i r_i x_i) \star y = r((1/r) \sum_i r_i x_i) \star y = \sum_i r_i (x_i \star y)$ , where the last equality follows from convex bilinearity. Hence the function  $\star$  is linear over conic combinations. Finally, consider a linear combination  $x = \sum_{i \in I} s_i x_i$  with  $s_i \in \mathbb{R}$ . Chop the index set  $I$  into  $I_- = \{i \in I \mid s_i < 0\}$  and  $I_+ = \{i \in I \mid s_i \geq 0\}$ . Then  $x + \sum_{i \in I_-} |s_i| x_i = \sum_{i \in I_+} s_i x_i$ , where each side of this equation is a conic combination of elements of  $\mathbb{H}_n$ . Taking the star product with  $y$  yields

$(x + \sum_{i \in I_-} |s_i| x_i) \star y = (\sum_{i \in I_+} s_i x_i) \star y$ , which after using conic linearity and rearranging gives  $x \star y = \sum_i s_i (x_i \star y)$ . Similar reasoning applied to the second argument yields bilinearity over the reals, as claimed.

We now prove that any product  $\star: \mathbb{H}_n \times \mathbb{H}_n \rightarrow \mathbb{H}_n$ , which is bilinear over the reals, and satisfies the remaining axioms, extends to a product on  $\mathbb{M}_n$ , which is bilinear over the complex numbers, and still satisfies the remaining axioms. Recall that a matrix  $x \in \mathbb{M}_n$  is Hermitian if  $x^\dagger = x$ , and anti-Hermitian if  $x^\dagger = -x$ ; there is a bijection  $x \mapsto ix$  between Hermitian and anti-Hermitian matrices. Any matrix  $x \in \mathbb{M}_n$  equals  $x = x_h + x_a$  for a unique Hermitian  $x_h \in \mathbb{H}_n$  and anti-Hermitian  $x_a$ , given by  $x_h = \frac{1}{2}(x + x^\dagger)$  and  $x_a = \frac{1}{2}(x - x^\dagger)$ . Define  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$  by

$$x \star y = x_h \star y_h - i(x_h \star iy_a) - i(ix_a \star y_h) - (ix_a \star iy_a).$$

Observe that this is well-defined, and that

$$\begin{aligned} (x \star y)_h &= x_h \star y_h - ix_a \star iy_a, \\ (x \star y)_a &= -i(x_h \star iy_a + ix_a \star y_h). \end{aligned}$$

Associativity of the extended function  $\mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$  now follows from associativity of the original function  $\mathbb{H}_n \times \mathbb{H}_n \rightarrow \mathbb{H}_n$  by a straightforward computation, and the same goes for the condition of being product when traced and for complex bilinearity. We write out the latter explicitly. Suppose that  $\lambda = \alpha + i\beta$  for  $\alpha, \beta \in \mathbb{R}$ . Observe that  $(\lambda x)_h = \alpha x_h + \beta ix_a$ , and  $i(\lambda x)_a = \alpha ix_a - \beta x_h$ . Then:

$$\begin{aligned} (\lambda x) \star y &= (\alpha x_h + \beta ix_a) \star y_h - i((\alpha x_h + \beta ix_a) \star iy_a) \\ &\quad - i((\alpha ix_a - \beta x_h) \star y_h) - (\alpha ix_a - \beta x_h) \star iy_a \\ &= (\alpha + \beta i)(x_h \star y_h) - i(\alpha + \beta i)(x_h \star iy_a) \\ &\quad - i(\alpha + \beta i)((ix_a) \star y_h) - (\alpha + \beta i)(ix_a \star iy_a) \\ &= \lambda(x \star y), \end{aligned}$$

and similarly  $x \star (\lambda y) = \lambda(x \star y)$ . Using  $(x+y)_h = x_h + y_h$  and  $(x+y)_a = x_a + y_a$  gives  $(x+y) \star z = x \star z + y \star z$  similarly easily.

Finally, consider the condition of being product on commuting pairs.

First, consider the case where  $x \in \mathbb{M}_n$  is *normal*, that is,  $xx^\dagger = x^\dagger x$ . By assumption  $xy = yx$ , and so  $x^\dagger y^\dagger = (yx)^\dagger = (xy)^\dagger = y^\dagger x^\dagger$ , too. By Fuglede's theorem [23], also  $xy^\dagger = y^\dagger x$ , and so  $x^\dagger y = yx^\dagger$ . Therefore

$$x_h y_h = 1/4(xy + x^\dagger y + xy^\dagger + x^\dagger y^\dagger) = 1/4(yx + yx^\dagger + y^\dagger x + y^\dagger x^\dagger) = y_h x_h,$$

and similarly  $x_h y_a = y_a x_h$ ,  $x_a y_h = y_h x_a$ , and  $x_a y_a = y_a x_a$ . It now follows that  $x_h \star y_h = x_h y_h$ ,  $x_h \star iy_a = ix_h y_a$ ,  $ix_a \star y_h = ix_a y_h$ , and  $ix_a \star iy_a = -x_a y_a$ . But this, together with the bilinearity of  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$ , implies:

$$x \star y = x_h \star y_h - i(x_h \star iy_a + ix_a \star y_h) - (ix_a \star iy_a) = x_h y_h + x_h y_a + x_a y_h + x_a y_a = xy.$$

Now we consider the case of a general  $x \in \mathbb{M}_n$  (not necessarily normal). Because both hermitian and anti-hermitian matrices are normal, an arbitrary

matrix  $x \in \mathbb{M}_n$  can be written as a sum of two normal matrices  $x = x_h + x_a$ . Combining the result for normal matrices with bilinearity of the extended function  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$ , we infer that  $x \star y = xy$  for any  $y \in \mathbb{M}_n$ .

This concludes the proof of Lemma 3.  $\square$

## C Larger state spaces

One could consider functions  $\star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_m$  for  $m$  possibly different from  $n$ . The condition of convex-bilinearity then still makes sense, but associativity and being product on commuting pairs do not. The most natural solution is to require that there exists a linear function  $p: \mathbb{M}_m \rightarrow \mathbb{M}_n$  and adapt the criteria accordingly:

- **associativity:**  $x \star p(y \star z) = p(x \star y) \star z$ ;
- **product on commuting pairs:** if  $xy = yx$  then  $p(x \star y) = xy$ ;
- **product when traced:**  $\text{Tr}(p(x \star y)) = \text{Tr}(xy)$ .

But then one easily sees that  $p \circ \star: \mathbb{M}_n \times \mathbb{M}_n \rightarrow \mathbb{M}_n$  satisfies the original criteria for a star product. So by Theorem 2, either  $p(x \star y) = xy$  for all  $x, y \in \mathbb{M}_n$  or  $p(x \star y) = yx$ . Hence  $p(x \star \mathbf{1}) = x$  and it follows that  $\text{rank}(p) \geq n^2$  and so  $m \geq n$ . Thus the function  $j: \mathbb{M}_n \rightarrow \mathbb{M}_m$  given by  $j(x) = x \star \mathbf{1}$  satisfies  $p(j(x)) = x$ . But then  $\text{Tr}(x \star y) = \text{Tr}(j(xy))$ , so that  $x \star y$  and  $j(xy)$  result in the same probabilistic predictions when measuring a state. In other words, there is nothing to be gained by moving from  $\mathbb{M}_n$  to  $\mathbb{M}_m$  for  $m > n$ .