OPEN ACCESS

University of
BRISTOL

Peer reviewed version

University of Bristol - Explore Bristol Research
General rights

Non-classically causal correlations without backwards in time signalling

Tom Purves and Anthony J. Short

H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, U.K.

(Dated: January 11, 2019)

We investigate the relationship between no backwards in time signalling and classically causal correlations. We discover that unlike the case for two parties, the no backwards in time signaling paradigm for three parties is not enough to ensure that correlations can be reproduced with a classical causal ordering. We demonstrate this with an explicit example. We also generalise some existing results for linear two-time states to multi-party scenarios.

I. INTRODUCTION

Recently, there has been growing interest in indefinite causal order in quantum theory [1–7], for which no classical ordering of events can explain the results. This may arise for many reasons, for example due to quantum switches (in which the order of two operations is controlled by a quantum bit) [3, 8, 9], superpositions of causal order arising from a quantum theory of gravity [10–12], or post-selection [13–18].

An interesting situation to consider is one in which a number of parties carry out experiments within individual laboratories, each of which obeys standard quantum theory (with a fixed causal order), whilst the connections between laboratories are more exotic. The most general object representing such connections is a *process matrix* [2, 6, 19–22]. Process matrices can lead to correlations between laboratories which defy any classically causal explanation, in an analogous way to that in which Bell-inequality-violating correlations defy local explanation [23].

In a recent paper [18], a particular class of experiments was considered, in which each party performs a fixed measurement followed by a chosen transformation. In this scenario, no-backwards-in-time-signalling (NBTS) correlations are considered, for which each party individually sees results consistent with the normal flow of time within their laboratory (i.e. their measurement results do not depend on which transformation they later performed). This leads to NBTS-conditions which are analogous to the spatial no-signalling conditions for Bell-type scenarios. It was shown for two parties that a connection between the laboratories involving post-selection satisfies the NBTS-conditions if and only if it can be represented by a process matrix (or equivalently a linear two-time state). In this paper, we extend these results to multiple parties.

Surprisingly, the correlations that could be obtained in the two-party NBTS scenario were limited to the classically causal set, despite the existence of consistent non-classical possibilities. This raises the interesting question of whether process matrices always lead to classical correlations in the NBTS-scenario. If this were the case it could lead to interesting insights about the nature of quantum causality. In this paper, however, we show that this is not the case, by giving an explicit counterexample
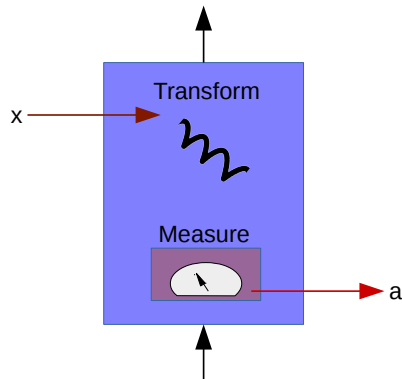

FIG. 1. A laboratory which obeys the NBTS paradigm. Experimenters in this lab always make measurements before transformations.

for three parties.

II. THE NBTS PARADIGM

We first clarify the NBTS scenario. Any protocol that fits within this framework contains a number of labeled laboratories, with each laboratory obeying a well defined procedure. The procedure (illustrated in figure 1) is always of the form:

1. A system enters laboratory.

2. A fixed measurement of the received system is made, and the classical result is recorded as an output.

3. A classical input is received. This can be thought of either as randomly generated, or provided by an independent external agent.

4. A transformation of the system is made, which may depend on the received input.

5. The system exists the lab.

If time in the labs is flowing normally we expect that outputs of the laboratories cannot depend on the inputs. Considering a single party, the conditional probability

distribution obeyed by her input x and output a therefore satisfies

$$p(a|x) = p(a) \tag{1}$$

which we call the NBTS condition [24].

Now consider three parties, Alice, Bob and Charlie, with each party obeying the NBTS procedure. Alice gets input x, and outputs a. Bob and Charlie get inputs y and z, and output b and c respectively. We expect the probability distribution for Alice's output, obtained by marginalizing $p(a,b,c|x,y,z)$ over Bob and Charlie's outputs, to also be independent of her input, because the NBTS scenario imposes that the generation of output happens strictly before her input. More explicitly, we expect

$$p_A(a|x,y,z) \equiv \sum_{b,c} p(a,b,c|x,y,z) \tag{2}$$

to be independent of x. Hence;

$$p_A(a|x,y,z) = p_A(a|x',y,z) \equiv p_A(a|y,z) \tag{3}$$

for all a, x, x', y, z. We also need to same to hold for Bob, as he should also conclude that there is no backwards in time signaling. Thus, Bobs marginal distribution should follow

$$p_B(b|x,y,z) = p_B(b|x,z), \tag{4}$$

and similarly for Charlie,

$$p_C(c|x,y,z) = p_C(c|x,y). \tag{5}$$

We may describe the permitted probability distributions as a convex polytope in probability-space [18, 20]. Here we consider the case where all inputs and outputs are binary for simplicity. For three parties, the situation is described by the 64 co-ordinates $p(a,b,c|x,y,z)$ for all possible values of the inputs and outputs. We demand that our coordinates are probabilities, and as such each coordinate varies between 0 and 1, and respect normalization,

$$0 \leq p(a,b,c|x,y,z) \leq 1 \qquad \sum_{a,b,c} p(a,b,c|x,y,z) = 1. \tag{6}$$

We also require that the probability distribution obeys the NBTS conditions (3)-(5) on each party. This gives a convex set which we refer to as the NBTS-polytope (which plays an analogous role to the non-signalling polytpe in non-locality). We also define a classically causal polytope that sits within the NBTS polytope that contains probabilities that could be achieved classically (analogous to the local polytope), in which systems can be represented by classical random variables, and the laboratories are arranged either with a deterministic classical ordering, or a mixture of such orderings. Deterministic classical strategies will sit at vertices of such a

polytope, and hence it can be generated in vertex representation by considering all such classical strategies. For example, suppose Alice goes first. This means that her output can only ever be a constant bit, as it occurs before any input. We can consider a strategy where Alice is followed by Bob, and then Charlie. An example strategy of this from is,

$$p(a,b,c|x,y,z) = \begin{cases} 1 \text{ if} & a = \alpha, b = \beta x \oplus \delta \\ & c = \gamma xy \oplus \lambda x \oplus \nu y \oplus \mu \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $\alpha, \beta, \gamma, \delta, \lambda, \mu, \nu$ are bits and $\oplus$ denotes addition modulo 2. There are also strategies where there is some freedom to choose the subsequent ordering based on Alice's input. For example, suppose that Alice's output system encodes her input bit, and that when $x = 0$ the system is next given to Bob and then Charlie, whereas when $x = 1$ the system is given to Charlie then Bob. Such a strategy is described by the formula

$$p(a,b,c|0,y,z) = \begin{cases} 1 \text{ if} & a = \alpha, b = \beta, c = \delta y \oplus \gamma \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$p(a,b,c|1,y,z) = \begin{cases} 1 \text{ if} & a = \alpha, c = \lambda, b = \nu z \oplus \mu \\ 0 & \text{otherwise}. \end{cases} \tag{9}$$

where once again $\alpha, \beta, \gamma, \delta, \lambda, \mu, \nu$ are bits. We then construct the full polytope by considering all strategies of these forms, and constructing the convex hull. We find by direct computation that there are 560 such vertices, once redundancies are removed. A facet representation is computationally hard to determine, and for the scope of the current paper irrelevant thus has been omitted.

Finally, we also define a quantum set, which again is analogous to the convex set of the same name in the non-locality scenario. Strategies within the quantum set are allowed to use process matrices outside the laboratories and standard quantum theory inside the laboratories. Hence; the quantum set sits in the middle of the size hierarchy. In [18] it was shown that for the case of two parties the quantum set and the classically causal polytope are identical in the NBTS scenario. While there are plenty of exotic ways to wire together two parties, we find that any resultant physical correlations are no different to those generated by Alice going first or Bob going first, or a mixture of the two. We may then ask, is the NBTS paradigm enough to ensure such a classically causal description?

In section IV we will show that this is not the case, by giving an explicit quantum protocol for three parties that generates an extreme point of the NBTS polytope, that is strictly not contained within the classical polytope.

III. LINEAR TWO-TIME STATES

In [17] it was shown that any two-party process matrix can be simulated within quantum theory, given the

power of post-selection. In particular, process matrices correspond to a particular set of pre- and post-selected states called linear two-time states.

The formalism we use to describe pre- and post-selected states is developed from [13, 15, 16] and described in detail in [17]. Key elements of the formalism are that all states, measurement operators and channels are represented as vectors in Hilbert spaces and their dual spaces, while time evolution is represented in a symmetric way by the connections between these vectors. Different vectors can be combined via the $\bullet$ operation, which connects vectors within a Hilbert space and its dual to give a scalar (i.e. $_{\mathcal{A}}\langle\psi|\bullet|\phi\rangle^{\mathcal{A}} = \langle\psi|\phi\rangle$), whilst combining vectors in different Hilbert spaces as a tensor product. As in relativity, we perform contractions between vectors with the same label raised and lowered. We call a vector η (with labels $\eta_{A_o B_o C_o}^{A_i B_i C_i}$) a linear two-time state for three parties if it obeys the relation

$$p(a,b,c|x,y,z) = \left(J(a|x)_{A_i}^{A_o} \otimes K(b|y)_{B_i}^{B_o} \otimes L(c|z)_{C_i}^{C_o}\right) \bullet \eta \tag{10}$$

where $J(a|x)_{A_i}^{A_o}$ represents a standard quantum measurement channel[25] between Alice's input A_i and output A_o which depends on x and has output a, and similarly for the other two parties. We note (10) is a linear function of the state and measurements.

In appendix A we show explicitly that the linear two-time state and process matrix formalisms are equivalent for any number of parties; given a process matrix there will always be a linear two-time state that is isomorphic. In appendix B we also show that imposing the NBTS conditions on a pre- and post-selected quantum state for all choices of the parties' measurements and transformations enforces that the state can be represented by a linear two-time state. These results were shown in [17, 18] for two parties, but we generalize them to any number of parties.

IV. THE PROTOCOL

In this section, we construct a tripartite NBTS-scenario which cannot be explained with a classically causal strategy. We use the same wiring of the laboratories as that given by Baumeler et al. in [5] to show that classical process matrices can violate causal inequalities, but a new protocol that respects the NBTS-scenario and only requires a binary input for each party. Alice, Bob, and Charlie are sat in their closed labs. Outside of the labs, they have no idea of how they are connected together. They each perform the following protocol: At some point a qubit enters their lab. They first measure the qubit in the computational basis, then receive a classical input bit. If this input bit is 0, they pass the qubit out of the door of their labs, and if it is a 1, they flip the qubit (by applying the gate $X = |0\rangle\langle 1| + |1\rangle\langle 0|$) and then pass it out. The action of Alice, whose input and output spaces are labelled by A_i and A_o, and who received

input x and obtained output a in her measurement, can be represented by the vectors

$$M(a|x)_{A_i}^{A_o} = |a \oplus x\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_i}\langle a| \otimes |a\rangle^{\mathcal{A}_i{}^\dagger} \otimes {}_{\mathcal{A}_o^\dagger}\langle a \oplus x|. \tag{11}$$

We denote primitive Hilbert spaces with calligraphic letters, and full Hilbert spaces with Roman labeling, i.e $H^A = H^{\mathcal{A}} \otimes H_{\mathcal{A}^\dagger}$. The laboratories have been wired together in rather odd way. The labs are connected in a cyclic fashion, as shown in figure 2, with probability $1/2$ that there is an external z basis measurement between them all. There is also $1/2$ probability that they are all connected with a z basis measurement followed by an X gate. This can be represented by the two-time state

$$\eta = \frac{1}{2}(M_{A_o}^{B_i} \otimes M_{B_o}^{C_i} \otimes M_{C_o}^{A_i} + \bar{M}_{A_o}^{B_i} \otimes \bar{M}_{B_o}^{C_i} \otimes \bar{M}_{C_o}^{A_i}) \tag{12}$$

where the operators M and $\bar{M}$ are given as

$$M_A^B = M(0|0)_A^B + M(1|0)_A^B \tag{13}$$
$$\bar{M}_A^B = M(0|1)_A^B + M(1|1)_A^B. \tag{14}$$

One may think at first glance that this setup does not respect the NBTS paradigm; it seems like there is a direct path between Alice's output and her input. However, a careful consideration of the scenario shows that a maximally random state will always enter each of the parties laboratories, and as such the NBTS conditions are respected. Furthermore, we show in appendix C that η is a linear two-time state, and hence will satisfy the NBTS conditions regardless of the specific measurements and transformations used by Alice, Bob and Charlie.

In order to calculate the measurement probabilities for this scenario, we use (10) to get

$$p(a,b,c|x,y,z) = (M(a|x)_{A_i}^{A_o} \otimes M(b|y)_{B_i}^{B_o} \otimes M(c|z)_{C_i}^{C_o}) \bullet \eta. \tag{15}$$

Evaluating this expression we arrive at the compact probability relation

$$p(a,b,c|x,y,z) = \frac{1}{2}\left(\delta_{b,a\oplus x}\delta_{c,b\oplus y}\delta_{a,c\oplus z} + \delta_{b,a\oplus\bar{x}}\delta_{c,b\oplus\bar{y}}\delta_{a,c\oplus\bar{z}}\right) \tag{16}$$

where we use a bar above a bit to represent its negation (e.g. $\bar{x} = x \oplus 1$). This generates the probabilities in table I.

We note that it carries some interesting relations, namely we have that:

$$p(a,b,c|x,y,z) = p(\bar{a},\bar{b},\bar{c}|x,y,z) = p(a,b,c|\bar{x},\bar{y},\bar{z}) \tag{17}$$
$$= p(b,c,a|y,z,x) = p(a,\bar{b},c|x,y,\bar{z}). \tag{18}$$

We can use these relations to show that this probability table cannot be generated by any classically causal strategy, or affine combination of such strategies. Let p_i be some deterministic classical strategy that could contribute to the affine combination that makes up p, i.e.
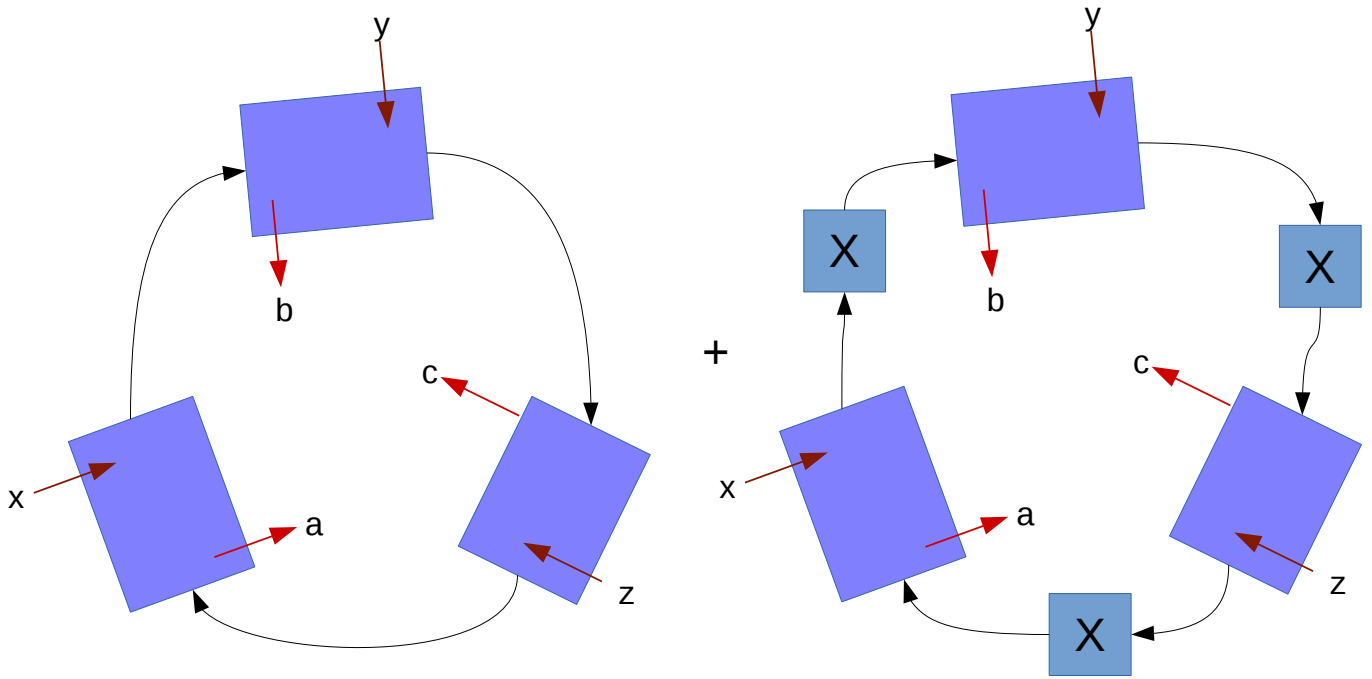
FIG. 2. Pictorial representation of the setup that η models. The black lines connecting laboratories represent channels with a z-basis measurement. The gates labeled X are flip gates. Notice that in the frame of each laboratory the output is produced before the input is received.

$p(a,b,c\|x,y,z)$		a,b,c							
		000	001	010	100	011	101	110	111
	000	1/2	0	0	0	0	0	0	1/2
	001	0	0	1/2	0	0	1/2	0	0
	010	0	0	0	1/2	1/2	0	0	0
x,y,z	100	0	1/2	0	0	0	0	1/2	0
	011	0	1/2	0	0	0	0	1/2	0
	101	0	0	0	1/2	1/2	0	0	0
	110	0	0	1/2	0	0	1/2	0	0
	111	1/2	0	0	0	0	0	0	1/2

TABLE I. Table of probabilities generated by the protocol.

$p = \sum_i \lambda_i p_i$, with $0 \leq \lambda \leq 1$. Then, consider a particular choice of a,b,c,x,y,z such that

$$p_i(a,b,c|x,y,z) > 0, \qquad (19)$$

Given that $p(a,b,c|x,y,z) \geq p_i(a,b,c|x,y,z)$, this is only possible if $p(a,b,c|x,y,z) = 0.5$ (the only non-zero value in the probability table), which also implies

$$p(a,\bar{b},c|x,y,\bar{z}) = 0.5 \text{ and } p(\bar{a},b,\bar{c}|x,y,\bar{z}) = 0.5. \quad (20)$$

Hence, using normalization

$$p(a,b,c|x,y,\bar{z}) = 0. \qquad (21)$$

The cyclic symmetry may be used to derive similar relations for the logical not of the other inputs, so that

$$p(a,b,c|x,\bar{y},z) = 0 \quad \text{and} \quad p(a,b,c|\bar{x},y,z) = 0. \quad (22)$$

Yet, the deterministic classically causal strategy p_i has some party that goes last when the inputs are x,y,z. The probability distribution must be the same when the last party's input is flipped (as it occurs after all outputs). Hence $p_i(a,b,c|x,y,z) > 0$ implies

$$p_i(a,b,c|\bar{x},y,z) + p_i(a,b,c|x,\bar{y},z) + p_i(a,b,c|x,y,\bar{z}) > 0, \qquad (23)$$

as one of the three terms must be positive. This further implies

$$p(a,b,c|\bar{x},y,z) + p(a,b,c|x,\bar{y},z) + p(a,b,c|x,y,\bar{z}) > 0, \qquad (24)$$

which creates a contradiction with (21) and (22). Hence we cannot write p as a mixture of classically causal strategies.

Any distribution in the polytope must satisfy the 20 NBTS equalities and normalization. The remaining constraints are all positivity inequalities, which are saturated when $p(a,b,c|x,y,z) = 0$. The probability distribution given in Table I contains 48 zeros, corresponding to saturated positivity inequalities. The combination of these are not all linearly independent, but instead yield 64 linearly independent equalities. As the space is 64 dimensional this leads to the point being extremal. The only way it would be possible to mix two points together to obtain this distribution would be for them both to violate the NBTS or normalisation equalities, or for one of them to violate a saturated positivity condition. However both of these methods utilize points outside the polytope. It

follows that the distribution produced by the protocol must be an extreme point of the NBTS set (and also an extreme point of the quantum set).

V. DISCUSSION AND CONCLUSIONS

We have presented an explicit protocol for three parties that cannot be realized with any classical causal order, yet can be implemented with a linear two time state (or process matrix), or within a more general theory that allows for indefinite causal orderings. It was already known that protocols of this form were able to violate causal inequalities [5], however we have confirmed that there also exist violations when the individual laboratories obey local no-backwards-in-time signalling principles. Moreover, the correlations we obtain are an extreme point of the NBTS polytope. This is in stark contrast to the two party NBTS scenario, where it has been shown that no such non-classically causal correlations exist. It is interesting to note that the state, measurements and transformations used in the protocol only involve Z-basis measurements and X-gates, and could therefore be obtained using classical stochastic processes on a classical linear two time state (classical process matrix). An interesting question is whether all quantum correlations for the NBTS scenario can be realised in such a way. In other words, do the NBTS conditions prevent any particular quantum advantage?

In this paper, we have focused on the case of indefinite relative timings between the three parties. In [18], cases in which the timings of two parties are definite are also considered, and it would be interesting to investigate these cases further for three or more parties.

Another interesting question is what, if any, additional constraints or scenarios can be considered beyond the NBTS case which are sufficient to ensure a classical causal explanation for any pre- and post-selected quantum procedure?

[1] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, Phys. Rev. Lett. **64**, 2965 (1990).
[2] O. Oreshkov, F. Costa, and Č. Brukner, Nature Communications **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].
[3] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Phys. Rev. A **88**, 022318 (2013).
[4] M. S. Leifer and R. W. Spekkens, Phys. Rev. A **88**, 052130 (2013).
[5] Ä. Baumeler, A. Feix, and S. Wolf, arXiv e-prints , arXiv:1403.7333 (2014), arXiv:1403.7333 [quant-ph].
[6] C. Branciard, M. Arajo, A. Feix, F. Costa, and Časlav Brukner, New Journal of Physics **18**, 013008 (2016).
[7] J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens, Physical Review X **7**, 031021 (2017).
[8] D. Ebler, S. Salek, and G. Chiribella, Physical Review Letters **120**, 120502 (2018), arXiv:1711.10165 [quant-ph].
[9] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, New Journal of Physics **17**, 102001 (2015), arXiv:1506.03776 [quant-ph].
[10] L. Hardy, Journal of Physics A Mathematical General **40**, 3081 (2007), gr-qc/0608043.
[11] L. Hardy, "Quantum gravity computers: On the theory of computation with indefinite causal structure," in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (Springer Netherlands, Dordrecht, 2009) pp. 379–401.
[12] M. Zych, F. Costa, I. Pikovski, and C. Brukner, ArXiv e-prints (2017), arXiv:1708.00248 [quant-ph].
[13] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Phys. Rev. **134**, B1410 (1964).
[14] Y. Aharonov and L. Vaidman, Journal of Physics A Mathematical General **24**, 2315 (1991).
[15] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, Phys. Rev. A **79**, 052110 (2009).
[16] R. Silva, Y. Guryanova, N. Brunner, N. Linden, A. J. Short, and S. Popescu, Physical Review A **89**, 012121 (2014), arXiv:1308.2089 [quant-ph].
[17] R. Silva, Y. Guryanova, A. J. Short, P. Skrzypczyk, N. Brunner, and S. Popescu, New Journal of Physics **19**, 103022 (2017).
[18] Y. Guryanova, R. Silva, A. J. Short, P. Skrzypczyk, N. Brunner, and S. Popescu, ArXiv e-prints (2017), arXiv:1708.00669 [quant-ph].
[19] O. Oreshkov and C. Giarmatzi, New Journal of Physics **18**, 093020 (2016).
[20] A. A. Abbott, C. Giarmatzi, F. Costa, and C. Branciard, Physical Review A **94**, 032131 (2016).
[21] Ä. Baumeler and S. Wolf, New Journal of Physics **18**, 035014 (2016).
[22] M. Araújo, P. A. Guérin, and Ä. Baumeler, Phys. Rev. A **96**, 052315 (2017), arXiv:1706.09854 [quant-ph].
[23] J. S. Bell, Physics Physique Fizika **1**, 195 (1964).
[24] For a single party, this is equivalent to a and x being uncorrelated.
[25] I.e. a trace non-increasing completely positive map.

Appendix A: Linear Two-Time States and Process Matrices give rise to the same probabilities

We give a general proof that process matrices and linear two time states exist that give the same probabilities. Furthermore, following [17] we observe a natural isomorphism between process matrices and the linear two-time states. we begin with the probability rule for process matrices from [19].

$$
\begin{aligned}
P(a_1,...,a_N) &= \mathrm{tr}\left[W^{A_1^1 A_2^1 ... A_1^N A_2^N}\left(\bigotimes_{i=1}^{N} M_{a_i}^{A_1^i A_2^i}\right)\right]\\
&= \mathrm{tr}\left[\sum_{\mu_1...\mu_N} W^{A_1^1 A_2^1 ... A_1^N A_2^N}\left(\bigotimes_{i=1}^{N}\left[(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]})|\Phi^+\rangle\langle\Phi^+|(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]\dagger})\right]^T\right)\right]\\
&= \mathrm{tr}\left[\sum_{\mu_1...\mu_N} (W^{A_1^1 A_2^1 ... A_1^N A_2^N})^T\left(\bigotimes_{i=1}^{N}\left[(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]})|\Phi^+\rangle\langle\Phi^+|(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]\dagger})\right]\right)\right]\\
&= \mathrm{tr}\left[\sum_{\substack{k_1...k_{2N}\\ l_1...l_{2N}}}\sum_{\mu_1...\mu_N} w_{k_1...k_{2N},l_1...l_{2N}}|l_1...l_{2N}\rangle\langle k_1...k_{2N}|\left(\bigotimes_{i=1}^{N}\left[(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]})|\Phi^+\rangle\langle\Phi^+|(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]\dagger})\right]\right)\right]\\
&= \sum_{\substack{k_1...k_{2N}\\ l_1...l_{2N}}}\sum_{\mu_1...\mu_N} w_{k_1...k_{2N},l_1...l_{2N}}\left(\prod_{i=1}^{N}\mathrm{tr}\left[|l_{2i-1}l_{2i}\rangle\langle k_{2i-1}k_{2i}|(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]})\sum_{t_i,u_i}|t_i t_i\rangle\langle u_i u_i|(\mathbb{I}\otimes \hat{E}_{\mu_i,a_i}^{[i]\dagger})\right]\right)\\
&= \sum_{k_1...k_{2N}}\sum_{l_1...l_{2N}}\sum_{\mu_1...\mu_N} w_{k_1...k_{2N},l_1...l_{2N}}\left(\prod_{i=1}^{N}\langle k_{2i}|\hat{E}_{\mu_i,a_i}^{[i]}|k_{2i-1}\rangle\langle l_{2i-1}|\hat{E}_{\mu_i,a_i}^{[i]\dagger}|l_{2i}\rangle\right)\\
&= \left(\sum_{k_1...k_{2N}}\sum_{l_1...l_{2N}} w_{k_1...k_{2N},l_1...l_{2N}}\,_{\mathcal{A}_2^1}\langle k_2|\otimes|k_1\rangle^{\mathcal{A}_1^1}\otimes_{\mathcal{A}_1^{1\dagger}}\langle l_1|\otimes|l_2\rangle^{\mathcal{A}_2^{1\dagger}}\otimes ...\right.\\
&\qquad \left.... \otimes_{\mathcal{A}_2^N}\langle k_{2N}|\otimes|k_{2N-1}\rangle^{\mathcal{A}_1^N}\otimes_{\mathcal{A}_1^{N\dagger}}\langle l_{2N-1}|\otimes|l_{2N}\rangle^{\mathcal{A}_2^{N\dagger}}\right)\bullet\left(\sum_{\mu_1}\hat{E}_{\mu_1,a_1}^{[1]}\otimes \hat{E}_{\mu_1,a_1}^{[1]\dagger}\otimes ... \otimes \sum_{\mu_N}\hat{E}_{\mu_N,a_N}^{[N]}\otimes \hat{E}_{\mu_N,a_N}^{[N]\dagger}\right)\\
&= \eta_W \bullet (J_{a_1}^{[1]}\otimes ... \otimes J_{a_N}^{[N]}) \tag{A1}
\end{aligned}
$$

which is equivalent to the probability rule for linear two-time states as quoted in the main text in (10). Furthermore, note that η is 'positive' (in the sense required to be a valid two-time state [17]) if and only if W is a positive operator. The isomorphism between W and η is obtained by expressing the former as a vector in the basis chosen for the transpose and flipping bras and kets on the output spaces.

Appendix B: Connection between linearity of two-time states and the NBTS conditions

To see that a linear two-time state always satisfies the NBTS conditions, note that it is equivalent to a process matrix. These are defined such that the phenomena inside individual laboratories obeys standard quantum theory (without post-selection) and thus cannot lead to backwards in time signalling. Alternatively, one can construct a multiparty analogue of equation (30) in [18] and use a similar argument to that in theorem 2 of [17].

Our task then is to prove that satisfying the NBTS conditions implies that a pre- and post-selected quantum state can be represent by a linear two-time state, for which we use a multiparty generalisation of the argument in theorem 2 of [17]. We first consider the marginal state of the first party, given by

$$
\eta_{A_1} = (J^{[2]}\otimes ... \otimes J^{[N]})\bullet \eta \tag{B1}
$$

where $J^{[k]}$ is the channel for the k'th party summed over their output, which may depend on their input x_k, and could be an arbitrary channel. It was shown in [17] that the only single party marginal states that satisfy NBTS take the form of a product with the identity operator on the output

$$
\eta_{A_1} = (\rho^{A_1^1}\otimes \mathbb{I}_{A_2^1}) \tag{B2}
$$

corresponding to states with no post-selection, or trivial post selection. Hence,

$$(J^{[1]} \otimes J^{[2]} \otimes \ldots \otimes J^{[N]}) \bullet \eta = (J^{[1]})^{A_2^1}_{A_1^1} \bullet (\rho^{A_1^1} \otimes \mathbb{I}_{A_2^1})$$

$$= \mathbb{I}_{A_1^1} \bullet \rho^{A_1^1} \tag{B3}$$

where we have used the fact that for any trace-preserving channel $\mathbb{I}_B \bullet J_A^B = \mathbb{I}_A$. Note that (B3) is independent of $J^{[1]}$ and hence does not depend on x_1. Using similar arguments for the other parties, $(J^{[1]} \otimes \ldots \otimes J^{[N]}) \bullet \eta$ must be a constant independent of the channels $J^{[1]}$ to $J^{[N]}$ (and hence independent of all of the inputs $x_1, \ldots, x_N$). By appropriately normalising η (which doesn't cause any physical effects in the two time state formalism) we can obtain

$$(J^{[1]} \otimes \ldots \otimes J^{[N]}) \bullet \eta = 1 \tag{B4}$$

for all choices of channels $J^{[1]}$ to $J^{[N]}$. From the general rule for probabilities in the two time state formalism,

$$p(a_1, \ldots, a_N) = \frac{(J_{a_1}^{[1]} \otimes \ldots \otimes J_{a_N}^{[N]}) \bullet \eta}{(J^{[1]} \otimes \ldots \otimes J^{[N]}) \bullet \eta}$$

$$= (J_{a_1}^{[1]} \otimes \ldots \otimes J_{a_N}^{[N]}) \bullet \eta \tag{B5}$$

which is the probability rule given in (10), and thus η is a linear two-time state.

Appendix C: Verification that η is a Linear Two-Time State

Verifying that η is linear two-time state amounts to showing that the relation

$$(J \otimes K \otimes L) \bullet \eta = 1 \tag{C1}$$

holds for all trace-preserving channels J, K, and L. Given that η contains a z-basis measurement in every link between parties, it is sufficient to consider only classical stochastic channels for the parties (see Appendix D). For classical bits, there are a set of four channels that are extreme points of the convex set of classical channels- the identity, the flip or X gate, and the throw away and replace with 0 and 1 channels. If (C1) holds for these it will hold for any convex combination of classical channels by linearity. Without loss of generality, consider Alice as the party of interest.

Let J be the identity channel firstly. Then consider

$$(I_{A_i}^{A_o} \otimes K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{A_o}^{B_i} \otimes M_{B_o}^{C_i} \otimes M_{C_o}^{A_i} + \bar{M}_{A_o}^{B_i} \otimes \bar{M}_{B_o}^{C_i} \otimes \bar{M}_{C_o}^{A_i}). \tag{C2}$$

By using the identity channel on Alice to connect and contract the indices this is equal to

$$(K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{A_o}^{B_i} \bullet M_{C_o}^{A_o} \otimes M_{B_o}^{C_i} + \bar{M}_{A_o}^{B_i} \bullet \bar{M}_{C_o}^{A_o} \otimes \bar{M}_{B_o}^{C_i}). \tag{C3}$$

We have the relations operationally that $M_A^B \bullet M_B^C = M_A^C$ and $\bar{M}_A^B \bullet \bar{M}_B^C = M_A^C$ which gives

$$(K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{C_o}^{B_i} \otimes M_{B_o}^{C_i} + M_{C_o}^{B_i} \otimes \bar{M}_{B_o}^{C_i}) \tag{C4}$$

$$=(K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet (M_{C_o}^{B_i} \otimes \frac{1}{2}(M_{B_o}^{C_i} + \bar{M}_{B_o}^{C_i})). \tag{C5}$$

We note that $\frac{1}{2}\left(M_{B_o}^{C_i} + \bar{M}_{B_o}^{C_i}\right)$ is equivalent to the operation throw away and replace with the maximally mixed state, i.e. has the operational form $\mathbb{I}_{B_o} \otimes \frac{1}{2}\mathbb{I}^{C_i}$. Since all the remaining channels are by definition trace preserving they satisfy $\mathbb{I}_B \bullet J_A^B = \mathbb{I}_A$, so we can move the identity through the state giving

$$\frac{1}{2}(K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet (M_{C_o}^{B_i} \otimes \mathbb{I}_{B_o} \otimes \mathbb{I}^{C_i})$$

$$=\frac{1}{2}(\mathbb{I}_{B_i} \otimes L_{C_i}^{C_o}) \bullet (M_{C_o}^{B_i} \otimes \mathbb{I}^{C_i})$$

$$=\frac{1}{2}L_{C_i}^{C_o} \bullet (\mathbb{I}_{C_o} \otimes \mathbb{I}^{C_i})$$

$$=\frac{1}{2}\mathbb{I}_{C_i} \bullet \mathbb{I}^{C_i}$$

$$=1. \tag{C6}$$

We should also confirm this when Alice applies the flip channel;

$$X_A^B = |0\rangle^\mathcal{B} \otimes {}_\mathcal{A}\langle 1| \otimes |1\rangle^{\mathcal{A}^\dagger} \otimes {}_{\mathcal{B}^\dagger}\langle 0| + |1\rangle^\mathcal{B} \otimes {}_\mathcal{A}\langle 0| \otimes |0\rangle^{\mathcal{A}^\dagger} \otimes {}_{\mathcal{B}^\dagger}\langle 1| \tag{C7}$$

which gives

$$(X_{A_i}^{A_o} \otimes K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{A_o}^{B_i} \otimes M_{B_o}^{C_i} \otimes M_{C_o}^{A_i} + \bar{M}_{A_o}^{B_i} \otimes \bar{M}_{B_o}^{C_i} \otimes \bar{M}_{C_o}^{A_i}). \tag{C8}$$

Now we have the relations that $X_A^B \bullet M_B^C = \bar{M}_A^C$ and $X_A^B \bullet \bar{M}_B^C = M_A^C$. Then we get

$$\begin{aligned}
&(K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{B_o}^{C_i} \otimes \bar{M}_{A_i}^{B_i} \bullet M_{C_o}^{A_i} + \bar{M}_{B_o}^{C_i} \otimes M_{A_i}^{B_i} \bullet \bar{M}_{C_o}^{A_i}) \\
&= (K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{B_o}^{C_i} \otimes \bar{M}_{C_o}^{B_i} + \bar{M}_{B_o}^{C_i} \otimes \bar{M}_{C_o}^{B_i}) \\
&= (K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet (\frac{1}{2}(M_{B_o}^{C_i} + \bar{M}_{B_o}^{C_i}) \otimes \bar{M}_{C_o}^{B_i}) \\
&= \frac{1}{2}(K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet (\mathbb{I}_{B_o} \otimes \mathbb{I}^{C_i} \otimes \bar{M}_{C_o}^{B_i}) \\
&= 1.
\end{aligned} \tag{C9}$$

By playing the same trick with all other channel types we can recover similar results. For instance the throw away and replace with zero channel has the operational form

$$\mathbb{I}_{A_i} \otimes |0\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_o^\dagger}\langle 0|. \tag{C10}$$

the method here can be visualised easily - if we have a circular structure of future preserving channels, we can always move identity 'backwards' around the circle, until it 'hits' the prepared state and gives 1. In particular

$$\begin{aligned}
&(\mathbb{I}_{A_i} \otimes |0\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_o^\dagger}\langle 0| \otimes K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{A_o}^{B_i} \otimes M_{B_o}^{C_i} \otimes M_{C_o}^{A_i} + \bar{M}_{A_o}^{B_i} \otimes \bar{M}_{B_o}^{C_i} \otimes \bar{M}_{C_o}^{A_i}) \\
&= (|0\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_o^\dagger}\langle 0| \otimes K_{B_i}^{B_o} \otimes L_{C_i}^{C_o}) \bullet \frac{1}{2}(M_{A_o}^{B_i} \otimes M_{B_o}^{C_i} \otimes \mathbb{I}_{C_o} + \bar{M}_{A_o}^{B_i} \otimes \bar{M}_{B_o}^{C_i} \otimes \mathbb{I}_{C_o}) \\
&= (|0\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_o^\dagger}\langle 0| \otimes K_{B_i}^{B_o} \otimes \mathbb{I}_{C_i}) \bullet \frac{1}{2}(M_{A_o}^{B_i} \otimes M_{B_o}^{C_i} + \bar{M}_{A_o}^{B_i} \otimes \bar{M}_{B_o}^{C_i}) \\
&= (|0\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_o^\dagger}\langle 0| \otimes K_{B_i}^{B_o}) \bullet \frac{1}{2}(M_{A_o}^{B_i} \otimes \mathbb{I}_{B_o} + \bar{M}_{A_o}^{B_i} \otimes \mathbb{I}_{B_o}) \\
&= (|0\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_o^\dagger}\langle 0| \otimes \mathbb{I}_{B_i}) \bullet \frac{1}{2}(M_{A_o}^{B_i} + \bar{M}_{A_o}^{B_i}) \\
&= (|0\rangle^{\mathcal{A}_o} \otimes {}_{\mathcal{A}_o^\dagger}\langle 0|) \bullet \mathbb{I}_{A_o}) \\
&= 1.
\end{aligned} \tag{C11}$$

Hence, the state is a linear two time state for all four extreme classical trace-preserving channels, and thus η is a linear two time state.

Appendix D: Quantum channel between measurements

We observe the fact here that any quantum channel that is sandwiched between two z-measurements is effectively stochastic, and superposition is destroyed. Consider a quantum channel with Kraus operators A_k, 'sandwiched' between two z basis measurements;

$$M_z[\mathcal{E}[M_z(\rho)]] = \sum_{i,j,k} |j\rangle\langle j|A_k|i\rangle\langle i|\rho|i\rangle\langle i|A_k^\dagger|j\rangle\langle j| = \sum_{i,j} q_{j|i}|j\rangle\langle i|\rho|i\rangle\langle j| \tag{D1}$$

where $q_{j|i}$ is equal to

$$q_{j|i} = \sum_k \langle j|A_k|i\rangle\langle i|A_k^\dagger|j\rangle. \tag{D2}$$

$q_{j|i}$ is clearly positive and real, being the magnitude of a complex number. We can identify D1 with a classical stochastic map if the numbers $q_{j|i}$ can be identified with probabilities, which they can be since

$$\sum_i q_{j|i} = \sum_{k,i} \langle j|A_k|i\rangle\langle i|A_k^\dagger|j\rangle = \sum_k \langle j|A_k A_k^\dagger|j\rangle = \langle j|I|j\rangle = 1. \tag{D3}$$

Thus, any quantum channel for qubits that gets sandwiched between two z-basis measurements can be simulated by a stochastic channel.