This is the accepted manuscript made available via CHORUS. The article has been published as:

Hierarchy in temporal quantum correlations
Huan-Yu Ku, Shin-Liang Chen, Neill Lambert, Yueh-Nan Chen, and Franco Nori
Phys. Rev. A 98, 022104 — Published 2 August 2018
DOI: 10.1103/PhysRevA.98.022104
Hierarchy in temporal quantum correlations

Huan-Yu Ku,1,2 Shin-Liang Chen,1,3 Neill Lambert,2 Yueh-Nan Chen,1,2,4∗ and Franco Nori2,5

1Department of Physics, National Cheng Kung University, 701 Tainan, Taiwan
2Theoretical Quantum Physics Laboratory, RIKEN Cluster for Pioneering Research, Wako-shi, Saitama 351-0198, Japan
3Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany
4Physics Division, National Center for Theoretical Sciences, 300 Hsinchu, Taiwan
5Department of Physics, The University of Michigan, Ann Arbor, 48109-1040 Michigan, USA

(Dated: July 2, 2018)

Einstein-Podolsky-Rosen (EPR) steering is an intermediate quantum correlation that lies in between entanglement and Bell non-locality. Its temporal analogue, temporal steering, has recently been shown to have applications in quantum information and open quantum systems. Here, we show that there exists a hierarchy among the three temporal quantum correlations: temporal inseparability, temporal steering, and macrorealism. Given that the temporal inseparability can be used to define a measure of quantum causality, similarly the quantification of temporal steering can be viewed as a weaker measure of direct cause and can be used to distinguish between direct cause and common cause in a quantum network.

I. INTRODUCTION

The concept of quantum steering was first articulated by Shrödinger [1] in response to the apparently non-local phenomenon of quantum correlations questioned by Einstein, Podolsky, and Rosen (EPR) [2]. Thanks to the celebrated inequality proposed by Bell [3], a great deal of theoretical and experimental investigation has been focused on quantum non-locality in the past few decades. Empowered by practical quantum information task requirements, spatial EPR steering was recently able to be studied in a more quantitative way [4–8]. Together with the concepts of Bell nonlocality and entanglement, EPR steering forms a hierarchy, and as such acts as an intermediate quantum correlation that lies in between the others [9,10], i.e., EPR steering is, in general, weaker than Bell nonlocality but stronger than quantum entanglement. Research on EPR steering in the past few years has seen the development of several interesting new avenues of study [9–23]. In addition to these theoretical developments, EPR steering has also been observed experimentally [24–26].

The notion of causality, cause and effect, is an intuitive concept. In quantum mechanics, however, applying the concept of causality is not always that straightforward. For example, quantum mechanics allows the superposition principle to be applied to causal relations, such that indefinite causal order may occur with proper design [27,28]. A measurement of a superposition of causal orders has been demonstrated very recently [29]. Another driving force for the research on quantum causality comes from Bell’s theorem, and its generalizations, that can be analyzed with a causal approach [30–34]. Potential applications of quantum casual relations in quantum information tasks have also been proposed [35,38].

In contrast to creating an indefinite causal order, some other experimental works related to quantum causal relations have also attracted attention, e.g. distinguishing different causal structures (common cause and direct cause) [39] and defining a measure of quantum causal effects (direct cause) [40].

Our goal in this work is to relate temporal steering to the notion of a quantum causal effect. To do so we first show that there also exists a hierarchy among the three temporal quantum correlations (temporal inseparability, temporal steering, and nonmacrorealism), which are provided when the condition of no-signalling in time (NSIT) [41–43] is obeyed. When NSIT in temporal steering is violated, non-vanishing temporal steering may occur under a dephasing process, which we prove to be the same as the distinguishability between two purely classical ensembles. Given that the temporal inseparability can be used to define a measure of quantum causal effects, we conclude that temporal steering can be viewed as a weaker measure of quantum causal effect and can be used to distinguish between direct cause and common cause in a quantum network.

II. MACROREALISM

Consider a system that evolves with time, and on which one can measure a physical quantity $Q$ at time $t_1$, $t_2$, or $t_3$ to obtain the corresponding values $Q(t_1)$, $Q(t_2)$, and $Q(t_3)$, respectively. In 1985, Leggett and Garg (LG) [44,45] proposed an inequality:

$$K \equiv C'_{12} + C'_{23} - C'_{13} \leq 1,$$

where $C'_{ij} \equiv \langle Q(t_i)Q(t_j) \rangle$ is the expectation value of the measurement outcomes at time $t_i$ and $t_j$ [46,47]. This inequality holds if the dynamics of the system is classical, in the realism sense, and the measurements are non-invasive. Violation of the inequality shows the incompatibility between quantum mechanics and macrorealism.

One can consider a more general scenario to investigate temporal correlations. For instance, there can be

∗yuehnan@mail.ncku.edu.tw
two or more quantities being measured at each moment of time. For simplicity, we consider the scenario with two times \( t_1 = 0 \) and \( t_2 = t \), at which the quantity \( x \in \{x\}_{n_x}^{n_x} \) and the quantity \( y \in \{y\}_{n_y}^{n_y} \) are measured respectively during each round of the experiment. Accordingly, one obtains the outcome \( a \in \{a\}_{n_a}^{n_a} \) and the outcome \( b \in \{b\}_{n_b}^{n_b} \) (see Fig. 1). After many rounds of the experiment, one can obtain a set of probability distributions \( \{p(a,b|x,y)\}_{a,b,x,y} \). Then, a macrorealistic (MS) theory restricts the probability distributions to be of the following form:

\[
p(a,b|x,y) \equiv \sum \limits_{\lambda} p(a|x,\lambda) p(b|y,\lambda) p(\lambda) \quad \forall a, b, x, y. \tag{2}\n\]

The physical interpretation of the above equation is the following: The probability distribution \( p(a,b|x,y) \) between time \( t_1 \) and \( t_2 \) does not depend on the history of the experiment. Therefore, there exist hidden parameters \( \lambda \), which can be deterministic or stochastic \( \Lambda \), defining all physical properties and forming the probability distributions \( p(a|x,\lambda) \) and \( p(b|y,\lambda) \).

In quantum theory, a measurement outcome is typically not pre-determined due to intrinsic uncertainty. The probability distributions follow Born’s rule:

\[
p(a,b|x,y) \equiv \text{tr} \left[ E_{b|y} \cdot \mathcal{E}(\sqrt{E_{a|x}} p_0 \sqrt{E_{a|x}}) \right] \quad \forall a, b, x, y, \tag{3}\n\]

where \( p_0 \) is the initially prepared quantum state, \( \{E_{a|x}\}_a \) denotes the positive-operator valued measurement (POVM), \( E_{a|x} \geq 0 \), \( \sum_a E_{a|x} = \mathbb{1} \) of each \( x \), similarly \( \{E_{b|y}\}_b \) is POVM of each \( y \), and \( \mathcal{E} \) describes the dynamics of the system from \( t_1 = 0 \) to \( t_2 = t \). In the following, the sets of probability distributions which do not admit Eq. (4) will be called \textit{nonmacrorealistic}.

Similar to the spatial case, one can also write down the so-called \textit{temporal Bell inequalities} \(^{49}\) to be a set of constraints for the macroscopic probability distributions. For instance, setting \( n_x = n_y = n_a = n_b = 2 \) and shifting \( a, b \in \{1, 2\} \) to \( a, b \in \{\pm 1\} \), the temporal Clauser-Horne-Shimony-Holt (CHSH) kernel is written as

\[
\langle B \rangle_{T-\text{CHSH}} = C_{xy} + C_{x'y} + C_{xy'} - C_{x'y'}, \tag{4}\n\]

where

\[
C_{xy} = p(a = b|x,y) - p(a \neq b|x,y) \tag{5}\n\]

is the expectation value of \( a \cdot b \). For a qubit system, \( \langle B \rangle_{T-\text{CHSH}} \) is upper bounded by 2 and \( 2\sqrt{2} \) for the MS model and quantum mechanics, respectively.

To give a proper quantification of the degree of nonmacrorealistic dynamics, we follow the techniques used for standard Bell inequalities, i.e., optimizing all possible combinations of the measurement settings which give the maximal quantum violation of \( \langle B \rangle_{T-\text{CHSH}} \): 

\[
\langle B \rangle_{\text{max} \text{ T-CHSH}} = \max_{x,x',y,y'} \left\{ 0, \left( \frac{\langle B \rangle_{T-\text{CHSH}} - 2}{2\sqrt{2} - 2} \right) \right\}. \tag{6}\n\]

### III. TEMPORAL STEERING

Now, consider that one can perform quantum state tomography (QST) to obtain the quantum state at time \( t_2 = t \) instead of obtaining the probability distributions. After many rounds of the experiment, one can obtain a set of quantum states \( \{\sigma_{a|z}(t)\} \) corresponding to those states found after the measurement event \( a|x \) at time \( t_1 = 0 \). It is rather convenient to define the so-called \textit{temporal assemblage} as a set of subnormalized state \( \{\rho_{a|z}(t) \equiv p(a|x) \sigma_{a|z}(t) \} \). Through this, a temporal assemblage contains the information on both \( p(a|x) = \text{tr}[\rho_{a|z}(t)] \) and \( \sigma_{a|z}(t) = \rho_{a|z}(t) / \text{tr}[\rho_{a|z}(t)] \). If one believes the measurement at time \( t_1 = 0 \) is \textit{noninvasive}, i.e., knowing the outcome \( a \) in \textit{prior}, without disturbing the system and its subsequent dynamics, then the observed temporal assemblage should satisfy the hidden-state model \(^{17} \,^{18} \,^{50} \)

\[
\rho_{a|z}(t) \equiv \sum \limits_{\lambda} p(\lambda)p(a|x,\lambda)\sigma_{\lambda} \quad \forall a, x, t \equiv \text{tr} \left[ E_{b|y}(t) \cdot \mathcal{E}(\sqrt{E_{a|x}} p_0 \sqrt{E_{a|x}}(t)) \right] \quad \forall a, b, x, y, t. \tag{7}\n\]

The physical interpretation of the temporal hidden-state model is the following: During each experimental round, there exists an ontic state \( \lambda \), which predetermines the outcome \( a \) when performing the measurement \( x \) at \( t_1 = 0 \), as well as pre-determining the quantum state \( \sigma_{\lambda} \) at time \( t_2 = t \).

A temporal assemblage which admits a quantum mechanical model can be written as:

\[
\rho_{a|z}(t) \equiv \mathcal{E}(\sqrt{E_{a|x}} p_0 \sqrt{E_{a|x}}(t)). \tag{8}\n\]
Given a temporal assemblage, one can know if it admits the hidden-state model Eq (7) by the feasibility problem of

\[
\{ \text{find } \sigma_\lambda \mid \rho_{a|x}(t) = \sum_\lambda p(\lambda)p(a|x, \lambda)\sigma_\lambda \}. \tag{9}
\]

We refer to those assemblages, which do not admit the hidden-state model, as temporal steerable, and the degree of the temporal steerability is quantified by the measure of temporal steerable weight \[18\] and temporal steering robustness (TSR) \[20\]. In the following, we will use TSR to quantify the degree of temporal steerability for a given temporal assemblage:

\[
\text{TSR} = \min_\alpha \alpha \quad \text{subject to}
\]

\[
\begin{aligned}
& \frac{1}{1 + \alpha} \rho_{a|x}(t) + \frac{\alpha}{1 + \alpha} \tau_{a|x}(t) = \sum_\lambda p(\lambda)p(a|x, \lambda)\sigma_\lambda \\
& \quad \quad \text{subject to } \sum_\lambda p(a|x, \lambda)\sigma_\lambda - \rho_{a|x}(t) \geq 0 \quad \forall a, x.
\end{aligned} \tag{10}
\]

where \(\tau_{a|x}(t)\) is a valid noisy temporal assemblage. This can be formulated as a semidefinite programming problem (SDP) \[10, 51–54\] as follows:

\[
\text{TSR} = \min \left( \text{tr} \sum_\lambda \sigma_\lambda - 1 \right), \quad \text{with } \sigma_\lambda \geq 0 \quad \forall \lambda
\]

\[
\text{subject to } \sum_\lambda p(a|x, \lambda)\sigma_\lambda - \rho_{a|x}(t) \geq 0 \quad \forall a, x.
\]

\[
\text{(11)}
\]

IV. PSEUDO DENSITY MATRIX AND TEMPORAL INSEPARABILITY

To complete the picture of a hierarchy of correlations, we give a brief introduction to the so-called pseudo density matrix introduced by Fitzsimons et al. \[40\]. A pseudo density matrix is a way to define the state of one (or more) system between two (or more) moments of time. By definition, the pseudo density matrix \(R\) of a qubit passing through a quantum channel is obtained by performing the QST before and after the evolution (see Fig. 1). Therefore, the pseudo density matrix is expressed as:

\[
R = \frac{1}{4} \sum_{i,j=0}^3 C_{ij} \cdot \sigma_i \otimes \sigma_j, \tag{12}
\]

where \(\{\sigma_i\}_{i=0,1,2,3} = \{1, \hat{X}, \hat{Y}, \hat{Z}\}\) is the set composed of the identity operator and the Pauli matrices. Here, \(C_{ij} = \text{tr} (R \cdot \sigma_i \otimes \sigma_j)\) are the expectation values of the result of these quantum measurements. A pseudo density matrix is hermitian and normalized, but not necessarily positive-semidefinite. In general, a pseudo density matrix can also describe the state between two systems at different time. One can see that \(R\) becomes a standard density matrix, which is positive-semidefinite, when the time-separation \(t_2 - t_1 = 0\). Therefore, the relation between two measurement events is called space-like correlated when \(R\) is positive-semidefinite.

Conversely, if \(R\) is not positive-semidefinite, it is definitely not constructed from a standard spatially separated system. In this case, the relation between two measurement events is called time-like correlated. In Ref. \[40\], the authors proposed a measure, called the \(f\)-function, to quantify the degree of such a temporal relation:

\[
f = \sum_i |\mu_i|, \tag{13}
\]

which is the summation over all the negative eigenvalues \(\{\mu_i\}\) of a given \(R\). In the rest of the discussions, all the pseudo density matrices are obtained by considering a single qubit at different times. Due to the mathematical similarity to a separable quantum state, in the following we will refer to the situation \(f \neq 0\) as temporally inseparable. It is worth noting that the “separability” here does not denote the separability with respect to two spatially separable systems, but indicates the pseudo density matrix can be written in the separable form \(R = \sum_\lambda p(\lambda)\omega_\lambda \otimes \theta_\lambda\), where \(p(\lambda)\) is the probability distribution, \(\omega_\lambda\) and \(\theta_\lambda\) are some valid quantum states acting on Hilbert spaces \(\mathcal{H}_\lambda\) at \(t_1 = 0\) and \(\mathcal{H}_\lambda\) at \(t_2 = t\), respectively \[52\]. Note that, in general, a temporal separable model implies \(f = 0\), but not vice versa.

V. A HIERARCHY OF TEMPORAL QUANTUM CORRELATIONS

Now we show a hierarchical relation between three temporal relations: nonmacrorealism, temporal steerability, and temporal inseparability. To this end, we show one can obtain the temporal assemblage \(\{\rho_{a|x}(t)\}_{a,x}\) by performing a set of POVMs \(\{E_{a|x}\}_{a,x}\) on the pseudo density matrix \(R\), in which \(\{E_{a|x}\}_{a,x}\) are the POVMs producing \(\{\rho_{a|x}(t)\}_{a,x}\). More precisely, we show

\[
\rho_{a|x}(t) = \text{tr}_A (E_{a|x} \otimes \mathbb{1} \cdot R), \tag{14}
\]

where

\[
\rho_{a|x}(t) = \mathcal{E} (\rho_{a|x}(0)) = \mathcal{E} (\sqrt{E_{a|x}} \rho_0 \sqrt{E_{a|x}}) \tag{15}
\]

and \(\rho_0 = \mathbb{1}/2\). The proof is given in Appendix A. Once Eq. (14) holds, the following formulation of an assemblage can be derived:

\[
\rho_{a|x}(t) = \text{tr}_A (E_{a|x} \otimes \mathbb{1} \cdot \sum_\lambda p(\lambda)\omega_\lambda \otimes \theta_\lambda) = \sum_\lambda p(\lambda)p_Q(a|x, \lambda)\theta_\lambda, \tag{16}
\]

where the set of probabilities \(p_Q(a|x, \lambda) := \text{Tr}(E_{a|x}\omega_\lambda)\) is constrained by the uncertainty relation \[50\]. In Eq. (7),
we assume that the pseudo density matrix $R$ is temporally separable. Since the set of probability distributions $p(a|x, \lambda)$ in a hidden-state model Eq. (2) is only constrained by the normalization property, a hidden-state model can reproduce an assemblage given by the above equation, but not vice versa. Therefore, we arrive at the hierarchical relation between temporal separability and temporal hidden-state model: a temporal assemblage constructed from a dynamical evolution admits a temporal hidden-state model if the corresponding pseudo-density matrix is temporally separable. Similarly,

$$p(a, b|x, y) = \text{tr} \left[ E_{b|y} \cdot \rho_{a|x}(t) \right] = \text{tr} \left[ E_{b|y} \cdot \sum_{\lambda} p(\lambda) p(a|x, \lambda) \sigma_{\lambda} \right] = \sum_{\lambda} p(\lambda) p(a|x, \lambda) p_Q(b|y, \lambda)$$

(17)

can be reproduced by macroscopic correlations [Eq. (2)], but not vice versa, i.e., there is a hierarchical relation between the temporal hidden-state model and macrorealism: a temporal correlation is macrorealistic if the corresponding temporal assemblage admits a temporal hidden-state model. The hierarchy can be described in a converse way: a non-macrorealistic dynamics leads to a temporal steerable assemblage, and a dynamics which leads to a temporal steerable assemblage gives an inseparable pseudo density matrix.

In the following, we propose a proposition which will be used in the example.

**Proposition.** When the initial state of the qubit is prepared in the maximally mixed state, the pseudo density matrix constructed under the amplitude-damping, the phase-damping, and the depolarizing channels are temporal-separable if $f = 0$, i.e.,

$$f = 0 \Rightarrow \exists \omega^A, \theta^A, \text{ such that } R = \sum_{\lambda} p(\lambda) \omega^A \otimes \theta^A. \quad (18)$$

The purpose of using the maximally mixed state as the initial condition is to produce assemblages which admit NSIT (cases which violate NSIT will be discussed later). The proof of the proposition is given in Appendix B.

As a simple example, we consider a qubit experiencing a depolarizing channel, described by Eq. (B1c). In Fig. 2 we plot the dynamics of the $f$, the TSR, and $\langle B \rangle_{T-CHSH}^{\max}$.

We can see that the vanishing time, in which the corresponding classical model emerges, of each quantifier is different, demonstrating the hierarchical relation among the three temporal quantum relations.

VI. CLASSICAL STEERING

In the above discussions, the scenario we consider is under the condition of NSIT. That is, the obtained temporal assemblages $\{\rho_{a|x}(t)\}$ obey

$$\sum_{a} \rho_{a|x}(t) = \sum_{a} \rho_{a|x'}(t) \quad \forall x \neq x'. \quad (19)$$

Given that temporal hidden-state model and temporal Bell inequalities assume non-invasive measurements, observing non-macrorealism and temporal steering while satisfying NSIT gives a stricter example of both properties (in that it rules out certain types of examples of false signatures of both effects due to classical clumsiness). In this section, we give a simple example of such a false-signature, which appears when the obtained temporal assemblages are not restricted to NSIT.

First, we show that, by the following explicit example, instead of performing measurements on an initial quantum state, one can prepare a temporal assemblage which leads to temporal steerability by just preparing a set of “classical (subnormalized) states”:

$$\rho_{a|x}(0) = \text{diag}[\alpha_a|x, \beta_a|x], \quad (20)$$

where $\alpha_a|x$ and $\beta_a|x$ are non-negative real numbers with $a, x \in \{1, 2\}$. We refer to these states as “classical” since all of them have just diagonal terms. Therefore, each state can be created by mixing, say, the spin of electrons in just one direction (e.g., $z$-direction). Such a temporal assemblage is steerable but trivial, and this is the reason that this scenario is not considered in the previous discussion, and ruled out by assuming NSIT.

In Appendix C, we show that if the measurement settings at time $t_1 = 0$ are set to be two, the asymptotic
value of TSR (or temporal steerable weight) when time goes to infinity will be the same as the trace distance between the summation of the elements of the temporal assemblage in different measurement settings, i.e.,

\[
\text{TSR}(\{\rho_{a|x}(t \to \infty)\}) = D \left( \sum_a \rho_{a|x}, \sum_a \rho_{a|x'} \right),
\]

where \( D \) is the trace distance between two quantum states. One notes that the trace distance in the classical case represents the distinguishability between two probability distributions. Equation \( 21 \) means that the quantification of temporal steering arises from a classically “clumsy” experiment if the condition of NSIT is violated.

VII. INFERRING CAUSAL STRUCTURE WITH TEMPORAL STEERABILITY

Finally, motivated by Ref. [40] proposing the \( f \)-function as a measure of quantum causal effect, which discriminates between spatial and temporal correlations, we propose that the degree of temporal steeribility can also be another measure of a quantum causal effect.

First of all, let us define the scenario of quantum causality discussed here. Consider two quantum systems that interact with each other through a black box as shown in Fig. 3(a). The correlations between the two systems may be due to spatial correlations (common cause) in Fig. 3(b) or temporal correlations (direct cause) in Fig. 3(c). The problem we would like to address is that how to discriminate between these two scenarios without knowing the mechanism of the black box.

To illustrate that temporal steering can discriminate between common and direct cause, we propose to include an auxiliary qubit (qubit-3) coherently coupled to qubit-1. By examining the time-like steering of qubit-3 to qubit-2, one can infer whether the correlations are due to a common cause or a direct cause. This is because, if the correlations are from the common cause, there is no time-like steering of qubit-3 to qubit-2 (dashed line), while an oscillatory time-like steering (blue curve) exists if they are from a direct cause.

\[
H = \hbar J (\sigma_1^+ \sigma_2^- + \sigma_1^- \sigma_2^+),
\]

where \( J \) is the coupling strength and \( \sigma_i^\pm \) are the raising and lowering operators of qubit-\( i \). To obtain the temporal assemblage of qubit-2 at \( t_2 = t \), three measurements in mutually unbiased bases of \( X, Y, \) and \( Z \) are performed on qubit-3 at time \( t_1 = 0 \). Actually, this is the so-called spatio-temporal steering scenario \([53]\), which is a generalization of temporal steering. The TSR of qubit-2 \( \text{TSR}_{3\rightarrow 2}(\{\rho_{a\rightarrow x}^2(t)\}) \) is plotted in Fig. 3. We can see that in the case that two qubits share a common cause, \( \text{TSR}_{3\rightarrow 2}(\{\rho_{a\rightarrow x}^2(t)\}) \) is always zero, while in the case that two qubits are connected by a direct cause, \( \text{TSR}_{3\rightarrow 2}(\{\rho_{a\rightarrow x}^2(t)\}) \) oscillates with time. This simple example illustrates how, as one might expect, given the hierarchy of temporal correlation introduced earlier, that the temporal steeribility can be used to distinguish between the direct and common causal effect in a quantum network.

VIII. CONCLUDING REMARKS

It is worth to note that a hierarchy relation between temporal steeribility and macrorealism is also considered in Ref. [59]. However, in their work, neither the steerability witness nor the temporal CHSH inequality is optimized. The results of our work fill this gap. Open questions include: does the separable property of proposition Eq. \([13]\) hold for any quantum channel? Can a temporal assemblage be obtained directly from a pseudo density matrix under the requirement of the violation of
no-signaling in time? How will the hierarchical relation change if we consider another formulation of “a state over time”, e.g., the one in [64] or the one constructed by a discrete Wigner representation [61, 62] (see Ref. [63] for more comparisons between the three methods.)

ACKNOWLEDGMENTS

The authors acknowledge useful discussions with Adam Miranowicz and Roope Uola. H.-Y. K. acknowledges the support of the Graduate Student Study Abroad Program (Grants No. MOST 107-2917-I-006-002). S.-L. C. would like to acknowledge the support from Postdoctoral Research Abroad Program (Grants No. MOST 107-2917-I-564-007). This work is supported partially by the National Center for Theoretical Sciences and Ministry of Science and Technology, Taiwan, Grants No. MOST 103-2112-M-006-017-MY4. F. N. is supported in part by the: MURI Center for Dynamic Magneto-Optics via the Air Force Office of Scientific Research (AFOSR) (FA9550-14-1-0040), Army Research Office (ARO) (Grant No. 73315PH), Asian Office of Aerospace Research and Development (AORD) (Grant No. FA2386-18-1-0405), Japan Science and Technology Agency (JST) (the IMPACT program and CREST Grant No. JPMJCR1676), Japan Society for the Promotion of Science (JSPS) (JSPS-RFBR Grant No. 17-52-50023). N.L. and F. N. are supported by the Sir John Templeton Foundation and the RIKEN-AIST Challenge Research Fund.

Note added.— After completing this work, we became aware of [64], which independently proved a hierarchy between temporal steeribility and nonmacroscopicity.

Appendix A: Obtaining a set of temporal correlations and a temporal assemblage from a pseudo density matrix

Now, we show that one can obtain the temporal assemblage \( \{ \rho_{a|x}(t) \}_{a,x} \) by performing a set of positive-operator valued measurements (POVMs) \( \{ E_{a|x} \}_{a,x} \) with \( E_{a|x} \geq 0 \), satisfying \( \sum_x E_{a|x} = 1 \), on the pseudo density matrix \( R \), in which \( \{ E_{a|x} \}_{a,x} \) is the POVMs producing \( \{ \rho_{a|x}(t) \}_{a,x} \). More precisely, we show

\[
\rho_{a|x}(t) = \text{tr}_A( E_{a|x} \otimes \mathbb{1} R ) \quad \forall a, x, \quad (A1)
\]

where

\[
\rho_{a|x}(t) = \mathcal{E}(\rho_{a|x}(0)) = \mathcal{E}(\sqrt{E_{a|x}}\rho_0 \sqrt{E_{a|x}}) \quad (A2)
\]

and \( \rho_0 = \mathbb{1}/2 \).

Proof.— Without the loss of generality, we assume \( \{ E_{a|x} \}_{a=x1} \) be projectors for each \( x \), i.e.,

\[
E_{a|x} = \frac{1}{2}(\mathbb{1} + a \cdot \vec{x} \cdot \vec{\sigma}), \quad (A3)
\]

with \( a \cdot \vec{x} \) being the vector corresponding to projector \( E_{a|x} \) in the Bloch sphere and \( \vec{\sigma} = (\vec{X}, \vec{Y}, \vec{Z}) \) being the Pauli matrices. Besides, the post-measurement states for each measurement event \( a|x \) will be \( E_{a|x} \). The temporal assemblage would be

\[
\rho_{a|x}(t) = p(a|x) \mathcal{E} \left( E_{a|x} \otimes \frac{\mathbb{1}}{2} E_{a|x} \right)
\]

\[
= \frac{1}{2} \mathcal{E} \left( \frac{\mathbb{1}}{2} \right) + \frac{a}{8} \sum_{j=1}^{3} \text{tr} \left[ \sigma_j \mathcal{E}(E_{1|x}) - \mathcal{E}(E_{-1|x}) \right] \sigma_j
\]

\[
= \frac{1}{2} \mathcal{E} \left( \frac{\mathbb{1}}{2} \right) + \frac{1}{4} \sum_{k=0}^{3} \text{tr} \left[ \sigma_k \mathcal{E}(\frac{\mathbb{1}}{2}) \right] \sigma_k + \frac{a}{8} \sum_{j=1}^{3} \text{tr} \left[ \sigma_j \mathcal{E}(E_{1|x}) - \mathcal{E}(E_{-1|x}) \right] \sigma_j
\]

(4)

Then, by the definition of the pseudo density matrix \( R \)

\[
R = \frac{1}{4} \left[ \mathbb{1} \otimes \mathbb{1} + \sum_{k=1}^{3} \text{tr} \left[ \sigma_k \mathcal{E}(\frac{\mathbb{1}}{2}) \right] \mathbb{1} \otimes \sigma_k + \frac{1}{2} \sum_{i,j=1}^{3} \text{tr} \left[ \sigma_j \mathcal{E}(E_{1|i}) - \mathcal{E}(E_{-1|i}) \right] \sigma_i \otimes \sigma_j \right]. \quad (A5)
\]

Finally, the target quantity \( \text{tr}_A( E_{a|x} \otimes \mathbb{1} R ) \) in Eq. (A1) would be
Using the fact that \( \text{tr}(E_{a|x}\sigma_i) = a \cdot \delta_{x,i} \), the above equation will be the same as Eq. (A4). Since now we have the temporal assemblages, obtained from the pseudo density matrix, it is straightforward to obtain a set of temporal correlations \( p(a, b|x, y) \).

We should note that from Eq. (A1), the way one obtains the temporal assemblage by performing measurement on the pseudo density matrix is merely a mathematical relation between \( \rho_{a|x}(t) \), \( E_{a|x} \), and \( R \), instead of a physical system being measured. This is different from the case in the standard spatial scenario that one obtains an assemblage by performing a set of local measurements on a subsystem of a quantum state. On the other hand, as we mentioned before, the reason to use the maximally mixed state as initial state is to obey the condition of no-signalling in time (NSIT).

**Appendix B: Proof of Proposition**

To support the proposition, in the following we will show the partial transpose of pseudo density matrix \( R \) is always positive semidefinite, i.e. \( R^T \geq 0 \), by considering the three standard quantum channels - the amplitude-damping channel, the phase-damping channel, and the depolarizing channel - which are often used to describe the dynamics of a system. Then, using the positive-partial-transpose (PPT) criterion \[63\] and \[66\], it is easy to show that \( R \) is separable.

The dynamics of a qubit undergoing the amplitude-damping, the phase-damping, and the depolarizing channels, can be respectively described by the following three Lindblad-form master equations:

\[
\dot{\rho}_s = \frac{\gamma_A}{2} (2\sigma_+ - \sigma_-) \rho_s - \sigma_+ \rho_s(t) - \rho_s(t) \sigma_- ,
\]

(B1a)

\[
\dot{\rho}_p = \frac{\gamma_p}{4} (2\sigma_3 \rho_s(t) - \sigma_3^2 \rho_s(t) - \rho_s(t) \sigma_3^2) ,
\]

(B1b)

\[
\dot{\rho}_s = \frac{\gamma_p}{8} \sum_i (2\sigma_i \rho_s(t) \sigma_i - \sigma_i^2 \rho_s(t) - \rho_s(t) \sigma_i^2) ,
\]

(B1c)

where \( \rho_s \) is the standard density matrix of the qubit, \( \{\gamma_i \}_{i=A,P,D} \) denote the decay rates of the dynamics in the different channels, and \( \sigma_+ (\sigma_-) \) is the creation (annihilation) operator. Assisted by the definition of the pseudo density matrix, one can obtain the pseudo density matrix in each scenario:

\[
R_A = \begin{pmatrix}
\frac{1}{2} e^{-t\gamma_A} & 0 & 0 & 0 \\
0 & \frac{1}{2}(1 - e^{-t\gamma_A}) & \frac{1}{2} e^{-t\gamma_A/2} & 0 \\
0 & \frac{1}{2} e^{-t\gamma_A/2} & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2}
\end{pmatrix},
\]

(B2a)

\[
R_P = \begin{pmatrix}
\frac{1}{2} e^{-t\gamma_P} & 0 & 0 & 0 \\
0 & \frac{1}{2} e^{-t\gamma_P} & 0 & 0 \\
0 & 0 & \frac{1}{2} e^{-t\gamma_P} & 0 \\
0 & 0 & 0 & \frac{1}{2}
\end{pmatrix},
\]

(B2b)

\[
R_D = \begin{pmatrix}
\frac{1}{4}(1 + e^{-t\gamma_D}) & 0 & 0 & 0 \\
0 & \frac{1}{4}(1 + e^{-t\gamma_D}) & \frac{1}{2} e^{-t\gamma_D} & 0 \\
0 & \frac{1}{2} e^{-t\gamma_D} & \frac{1}{4}(1 - e^{-t\gamma_D}) & 0 \\
0 & 0 & 0 & \frac{1}{4}(1 + e^{-t\gamma_D})
\end{pmatrix}.
\]

(B2c)

It can be shown that the partial transpose of each above pseudo density matrix is always positive semidefinite, i.e., \( R^T \geq 0 \) for \( t \in (0, \infty) \). The fact that \( f = 0 \) implies \( R \geq 0 \), indicating \( R \) can be treated as a valid density matrix.
describing a qubit-qubit system. By using the positive-partial-transpose (PPT) criterion [63] and [64]: a density matrix $\varrho_{AB}$ describing a qubit-qubit (or a qubit-qutrit) system is separable if and only if its partial transpose $\varrho_{AB}^{T_A}$ is positive-semidefinite. In summary, we prove the proposition by the following steps:

$$R \geq 0 \land R^{T_A} \geq 0 \land \text{PPT criterion} \Rightarrow R \text{ is separable.}$$  \hspace{1cm} (B3)

**Appendix C: Proof of Eq. 10 in the main text**

Following the property of temporal steerable weight (TSW) [18], one realizes that

$$\sigma_{a|x}^T - \sum_{\lambda} D_\lambda(a|x)\sigma_\lambda \geq 0,$$  \hspace{1cm} (C1)

where $D_\lambda(a|x)$ are the extremal deterministic values, $\lambda$ represents a local hidden variable, $x$ is the measurement basis, and $a$ is the measurement outcome. Since $\sum_a D_\lambda(a|x) = 1$, one has the following

$$\sum_a \sigma_{a|x}^T - \sum_{\lambda} \sigma_\lambda \geq 0,$$  \hspace{1cm} (C2)

If we are limited to two measurement inputs and preparing the assemblages with a classical way (without the off-diagonal terms), the summation of the temporal assemblages $\sigma_{a|x}^T$ can be written as

$$\sum_a \sigma_{a|1}^T = \begin{pmatrix} \alpha & 0 \\ 0 & 1 - \alpha \end{pmatrix},$$  \hspace{1cm} (C3)

$$\sum_a \sigma_{a|2}^T = \begin{pmatrix} \beta & 0 \\ 0 & 1 - \beta \end{pmatrix}.$$  \hspace{1cm} (C4)

Let us assume $\alpha > \beta$. The summation of the local hidden assemblage $\tilde{\sigma}_\lambda = \sum_a \sigma_\lambda$ that can best mimic the temporal assemblages and fulfill the requirement of Eq. (B2) is thus written as

$$\tilde{\sigma}_\lambda = \begin{pmatrix} \beta & 0 \\ 0 & 1 - \alpha \end{pmatrix}.$$  \hspace{1cm} (C5)

To prove that Eq. (C5) is the optimal solution, one can add a non-negative number $\epsilon$ into the diagonal terms of the matrix in Eq. (C5). It is easy to see that $\text{Tr}(\tilde{\sigma}_\lambda)$ is maximum when $\epsilon = 0$. Therefore, the TSW is equal to the trace distance between the two states $\sum_a \sigma_{a|1}^T$ and $\sum_a \sigma_{a|2}^T$, i.e.

$$\text{TSW} = 1 - \text{Tr}(\tilde{\sigma}_\lambda) = \alpha - \beta.$$  \hspace{1cm} (C6)

A similar argument can also be applied to the temporal steering robustness (TSR) [20] with the following requirement

$$\sum_{\lambda} \sigma_\lambda - \sum_a \sigma_{a|x}^T \geq 0.$$  \hspace{1cm} (C7)

This leads one to write the summation of the local hidden assemblage as

$$\tilde{\sigma}_\lambda = \begin{pmatrix} \alpha & 0 \\ 0 & 1 - \beta \end{pmatrix},$$  \hspace{1cm} (C8)

and the corresponding TSR is written as

$$\text{TSR} = \text{Tr}(\tilde{\sigma}_\lambda) - 1 = \alpha - \beta.$$  \hspace{1cm} (C9)

These conclude our proof that, in the classical scenario (no off-diagonal elements), the temporal steering is equal to the trace distance between the summation of the elements of the temporal assemblage in different measurement settings.


[55] These two Hilbert spaces are basically the same. However, due the the following discussion, e.g. take partial trace of $R$, it is necessary to use distinguishable subscripts..


