Realization of algorithmic identification of cause and effect in quantum correlations

Zhao-An Wang,[1,2] Yu Meng,[1,2,*] Zheng-Hao Liu,[1,2,†] Yi-Tao Wang,[1,2] Shang Yu,[3,1,4] Wei Liu,[1,2] Zhi-Peng Li,[1,2] Yuan-Ze Yang,[1,2] Nai-Jie Guo,[1,2] Xiao-Dong Zeng,[1,2] Jian-Shun Tang,[1,2,5,‡] Chuan-Feng Li,[1,2,5,§] and Guang-Can Guo[1,2,5]

[1]*CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, China*

[2]*CAS Center For Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China*

[3]*Quantum Optics and Laser Science, Blackett Laboratory, Imperial College London, Prince Consort Rd, London SW7 2AZ, United Kingdom*

[4]*Research Center for Quantum Sensing, Zhejiang Lab, Hangzhou 310000, People's Republic China*

[5]*Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*

(Received 24 May 2022; revised 20 April 2023; accepted 5 December 2023; published 3 January 2024)

Causal inference revealing causal dependencies between variables from empirical data has found applications in multiple subfields of scientific research. A quantum perspective of correlations holds the promise of overcoming the limitation of Reichenbach's principle and enabling causal inference with only observational data. However, it is still not clear how quantum causal inference can provide operational advantages in general cases. Here, we have devised a photonic setup and experimentally realized an algorithm capable of identifying any two-qubit statistical correlations generated by the two basic causal structures under an observational scenario, thus revealing a universal quantum advantage in causal inference over its classical counterpart. We further demonstrate the explainability and stability of our causal discovery method, which is widely sought in data processing algorithms. Employing a fully observational approach, our result paves the way for studying quantum causality in general settings.

DOI: 10.1103/PhysRevA.109.012406

I. INTRODUCTION

Causal identification has garnered considerable research interest in the field of big data science. While the human brain possesses a certain degree of causal thinking ability, many current artificial intelligences relying on black-box models excel at identifying correlations but encounter difficulties discerning causal relationships. Firstly, they prove challenging to interpret and may exhibit unexpected behaviors. For instance, in the context of COVID-19 diagnosis, image classifiers may divert attention towards confounding factors rather than focusing on the actual causative elements [1]; secondly, these models can become unstable and susceptible to confusion when confronted with minor alterations in input data. An example of this phenomenon can be observed in self-driving cars that operate effectively within the European continent but encounter difficulties in the distinct driving conditions of the British islands [2]. The capacity to differentiate between causes and effects thus presents the potential to enhance both the explainability and stability of acquired knowledge [3].

The challenge faced by machine learning–based methods [4] in comprehending causality stems from a well-known limitation known as Reichenbach's common cause principle (RCCP) [5]. According to RCCP, a latent variable that jointly influences two events can always produce the same correlation as a direct causal link between the two events. The consequence of RCCP for causal identification is the inability to differentiate between a common cause (CC) and a direct cause (DC) using observational data generated by classical entities. Given that quantum correlations have drastically different behaviors [6–8] from their classical counterparts, it is natural to consider the extent to which a quantum perspective on correlations and causations can overcome the limitations imposed by RCCP. The question has motivated a number of works exploring the quantum advantages in identifying causal relationships [9–19]. However, despite the extensive research in this area, none of the existing works have achieved unconditional discrimination between the fundamental causal structures, namely, distinguishing a DC from a CC solely based on observational quantum correlations.

In this study, we tackle the quantum counterpart of the two-point causal identification problem. Our aim is to determine whether the causal structure of a two-point qubit correlation is induced by the same particle going through a unitary quantum channel, which amounts to a DC, or two (possibly entangled) particles being successively measured, which constitutes a CC. We have realized an optical setup that allows the switching between the two different mechanisms with a photonic controlled-SWAP gate. Using the setup, we experimentally demonstrate that a few judiciously chosen measurements

*Present address: Center for Hybrid Quantum Networks (Hy-Q), The Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark; mengyu23@mail.ustc.edu.cn

†Present address: Center for Macroscopic Quantum States (bigQ), Department of Physics, Technical University of Denmark, Fysikvej, 2800 Kongens Lyngby, Denmark.

‡tjs@ustc.edu.cn

§cfli@ustc.edu.cn

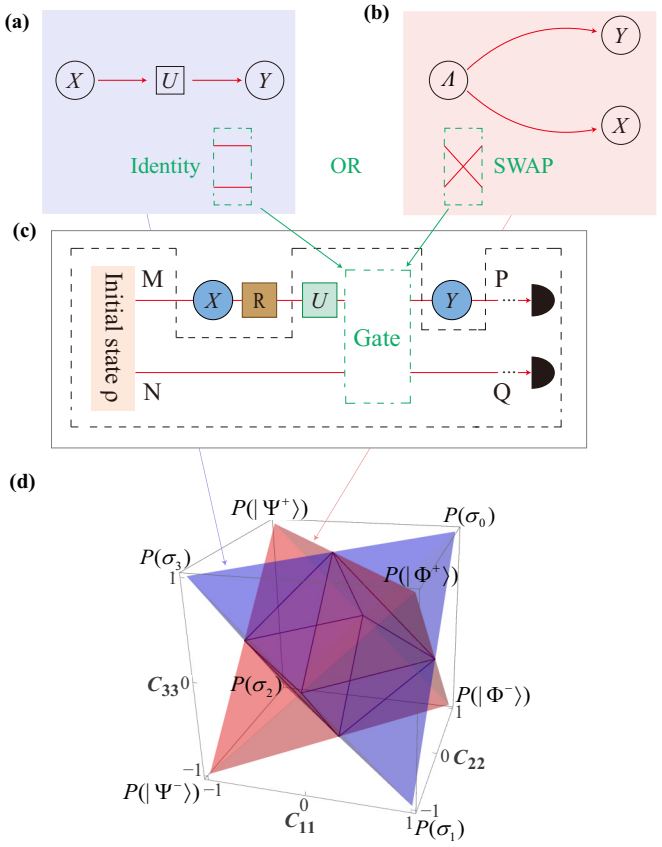

FIG. 1. Perspectives of two-point quantum causality. (a and b) Directed acyclic graph of DC and CC. (c) Quantum comb representation of causal structures. DC and CC can be switched by selecting the quantum gates in the green dashed box to be identity or SWAP. (d) Geometric description of two-qubit causal structures. The two-point correlations are characterized by pairs of identical Pauli measurements at the end of (c). The outcomes are represented by the three-dimensional coordinate of a point **P**. The blue and red tetrahedrons are the full set of points derived from DC and CC, respectively. The vertices of the DC and CC tetrahedrons are achieved using Pauli unitaries and Bell states.

suffice to distinguish any two-point qubit DC from CC in a geometric way. The initial formulation of the measurement strategy can be found in the work by Zhang *et al.* [20]; here, we further develop the strategy into a systematic algorithm that exhibits robustness against noise. Our findings unequivocally demonstrate the universal, explainable, and stable quantum advantage in two-point causal inference and thus could be highly interesting to quantum machine learning.

II. QUANTUM CAUSALITY WITH GEOMETRY

We start by introducing the quantum versions of the two most fundamental causal structures, DC and CC. In classical causal theory, a DC implies that an earlier variable directly influences a subsequent variable, while a CC signifies a confounding factor co-influencing the two variables. In quantum theory, the notion of variables is replaced by quantum states. As shown in Figs. 1(a) and 1(b), a quantum channel connecting input and output states can serve as a quantum analog of DC, whereas a bipartite system consisting of two subsystems gives rise to a quantum CC. Throughout the present work, we focus on two-qubit entanglement and unitary channels.

To illustrate the detailed procedures involved in generating and measuring correlations, we employ a demonstration known as the quantum comb [21,22]. As demonstrated in Fig. 1(c), the black dashed box represents the central component of the "comb" and encompasses the initial state preparation and processing stages, symbolizing the unknown origin of correlation in classical causal inferences. The two shallow notches stand for the access to information during the quantum process. The red transverse lines from left to right correspond to the temporal propagation of the two-qubit system ρ.

We need to note that the causal structures inherent in quantum systems are not straightforwardly determined by the sequential order of their emergence. In fact, a memory system has the capability to create a temporal separation between two subsystems that are linked by a CC. This feature allows us to construct both the DC and CC structures of a two-qubit system within a single setup. In Fig. 1(c), we plotted a simple way to switch between DC and CC in the quantum comb. Specifically, when the gate is chosen to preserve the subsystems, the correlation between X and Y corresponds to a DC induced by a unitary channel, with its Kraus operator denoted as U. On the other hand, when the gate is selected to swap between the two subsystems, the correlation between X and Y arises from the entangled initial state ρ, representing a CC relationship (for convenience, we mark the two starting ports as M and N, and the two terminal ports as P and Q). If the quantum gates are chosen randomly, the causal structure of the quantum correlations extracted from X, Y is then random. As such, to infer the causal structure, the observer must rely on quantum measurements to extract further information, i.e., the two blue circles above the line representing quantum systems X and Y. The measurements can have multiple settings and be destructive as long as they do not introduce signaling toward the other measurement. Therefore, a procedure R [cf. Fig. 1(c)] is necessary to recover the postmeasurement state according to the Lüders rule [23]. This guarantees a preceding measurement will not change the marginal distribution of the next measurement and the causal inference scheme is strictly observational.

With the above preliminaries, we can recover the affirmation in Ref. [12] that observational quantum causal identification is sometimes possible even with fixed combinations of measurements. The idea can be clearly illustrated in a geometric way: denote $i, j \in \{1, 2, 3\}$ to be the measurement settings at X and Y, and the three indices mean the observables should be the three Pauli matrices σ_1, σ_2, and σ_3, respectively. Further, we define the following conditional probability:

$$C_{ij} := p(x = y|i, j) - p(x \neq y|i, j), \tag{1}$$

where x, y are the outcomes of the measurements at X and Y. The two-point correlations of the same-setting measurements can then be to a three-dimensional vector in the coordinate space:

$$\mathbf{P} = (C_{11} \;\; C_{22} \;\; C_{33})^{\mathrm{T}}. \tag{2}$$

Algorithm 1 The algorithm for discrimination of the two-point qubit causal structures. $\mathcal{T}_{\mathrm{DC}}$ ($\mathcal{T}_{\mathrm{CC}}$) is the DC (CC)tetrahedron in the main text, whose four vertices are $(+1,+1,+1)^{\mathrm{T}}$, $(+1,-1,-1)^{\mathrm{T}}$, $(-1,+1,-1)^{\mathrm{T}}$, and $(-1,-1,+1)^{\mathrm{T}}$ $[(-1,-1,-1)^{\mathrm{T}}$, $(-1,+1,+1)^{\mathrm{T}}$, $(+1,-1,+1)^{\mathrm{T}}$, and $(+1,+1,-1)^{\mathrm{T}}]$. The quantities δ, ε, and ε' are cutoff values and can be conveniently chosen. The function $\mathcal{D}$ means the Euclidean distance of its two arguments.

Input: A two-point Pauli correlation $\mathbf{P} = (C_{11}, C_{22}, C_{33})^T$, $\mathbf{P} \in \mathcal{T}_{\mathrm{DC}} \cup \mathcal{T}_{\mathrm{CC}}$. The sum of all entries of $\mathbf{P}$ is b
Output: The latent causal structure of the correlation, DC or CC
  if $\mathbf{P} \notin \mathcal{T}_{\mathrm{DC}} \cap \mathcal{T}_{\mathrm{CC}}$ **then** ▷ Distinguishable by Pauli correlations
    if $\mathbf{P}$ is in the DC tetrahedron **then return** DC
    else return CC
    end if
  else if $1 - b < \delta$ **then** ▷ Distinguishable by symmetrically modified Pauli correlations
    calculate V
    measure the new correlation C_{33}^{V}
    if $1 - C_{33}^{V} < \varepsilon$ **then return** DC
    else return CC
    end if
  else ▷ Distinguishable by asymmetrically modified Pauli correlations
    calculate $V = V_2 V_1$, $V' = V_2 \sigma_1 V_1$ ▷ V_1 and V_2 are obtained in two consecutive rounds
    measure the new correlation $\mathbf{P} = (C_{11}^{V'}, C_{22}^{V'}, C_{33}^{V'})^T$
    if $\mathcal{D}[\mathbf{P}_{V'}, \mathbf{P}(\sigma_3)] < \varepsilon'$ **then return** DC
    else return CC
    end if
  end if

We use $\mathbf{P}(U)$ and $\mathbf{P}(\rho)$ to differentiate the correlations generated from DC and CC. We plot the sets of all possible $\mathbf{P}(U)$ and $\mathbf{P}(\rho)$ in Fig. 1(d) as the blue and red polytopes, which we refer to as DC and CC tetrahedra. The tetrahedral shape follows from that any two-point qubit correlation can be expressed as the convex combination of the four extremal points at the corresponding tetrahedron's vertices [17].

III. ALGORITHMIC CAUSAL IDENTIFICATION

As the two tetrahedra in Fig. 1(d) are nonidentical, some vectors $\mathbf{P}$ observed in the disjoint areas will imply causation. However, the DC and CC tetrahedra also overlap in an octahedron. For those $\mathbf{P}$ falling there, both DC and CC causal explanations are possible, and the Pauli correlations alone are insufficient to discern the underlying causality. Here, we show how to remove this ambiguity by employing only a few additional measurements.

The starting point of our causal identification algorithm is that the cross sections of the DC and CC tetrahedra at $C_{33} = 1$ correspond to two perpendicular lines, and they are distinguishable in almost all regions. Motivated by this observation, we seek to find a new set of observables $\sigma_k \to V\sigma_k V^\dagger$, $k \in \{1,2,3\}$ determined by the unitary operator V, so the correlation vectors for the new observables become $\mathbf{P}_V(U) = \mathbf{P}(V^\dagger U V)$ and $\mathbf{P}_V(\rho) = \mathbf{P}[(V \otimes V)^\dagger \rho (V \otimes V)]$. The goal of choosing V is to make the third entry of $\mathbf{P}_V(U)$ equal to 1. To put it simply, the choice of V should cause the channel to preserve the eigenstates of $V\sigma_3 V^\dagger$.

The existence of such an operator V is intuitive: in the Bloch sphere representation, a qubit unitary channel behaves as a rotation, and the states on the rotational axis are invariant. Therefore, an appropriate choice of operator V should alter the orientation of the coordinate system so that its zenith aligns with the axis of rotation. Its explicit settings can be obtained from the Pauli correlations $\mathbf{P}$, which need to be measured in the first place. The key observation is that, under the new measurement settings, any two-qubit DC correlation $\mathbf{P}_V(U)$ will satisfy $C_{33}^{V}(U) = 1$ and $C_{11}^{V}(U) = C_{22}^{V}(U)$, where the sub- or superscripts mean the measured observables are the Pauli operators modified by V; in contrast, only a zero-measure set of $\mathbf{P}(\rho)$ lying on an "exceptional" plane, $\sum_{k=1}^{3} C_{kk} = 1$, makes $C_{33}^{V}(\rho) = 1$. Therefore, for those $\mathbf{P}$ not on the exceptional plane, the condition $C_{33}^{V} = 1$ alone can serve as an indicator for the DC causal structure. The detailed numerical determination of V, the proof of the aforementioned observation, can be found in Appendix B, and the pseudocode for the entire algorithm are provided in Algorithm1.

In practice, both the causal mechanism to be identified and the measurements themselves will be subject to noise. When applying the causal identification algorithm, we have to relax the condition for DC as $1 - C_{33}^{V} < \varepsilon$ to allow some tolerance, where $\varepsilon > 0$ is a cutoff value. The nonzero tolerance causes the indistinguishable CC set to have a nonzero measure. To handle this issue, whenever the initial correlation $\mathbf{P}$ is close to the exceptional plane ($1 - \sum_{k=1}^{3} C_{kk} < \delta$, where δ is a threshold value), we will modify the first measurement with V and the second measurement with $V' = \sigma_1 V$, take the resulting correlation as the new initial correlation, and run the algorithm once again. Crucially, as the additional σ_1 enforces $C_{33}^{V} = -1$, the new initial correlation will be opposite to the exceptional plane on the octahedron. The final correlation induced by any DC, $\mathbf{P}_{V'}(U)$, will thus always end close to $\mathbf{P}(\sigma_3) = (-1,-1,1)^{\mathrm{T}}$, far from the CC tetrahedron, and reflect causality with strong resilience of noise.

IV. EXPERIMENTAL DEMONSTRATION

We implemented the quantum causal identification algorithm in an optical platform. We encoded two qubits on the

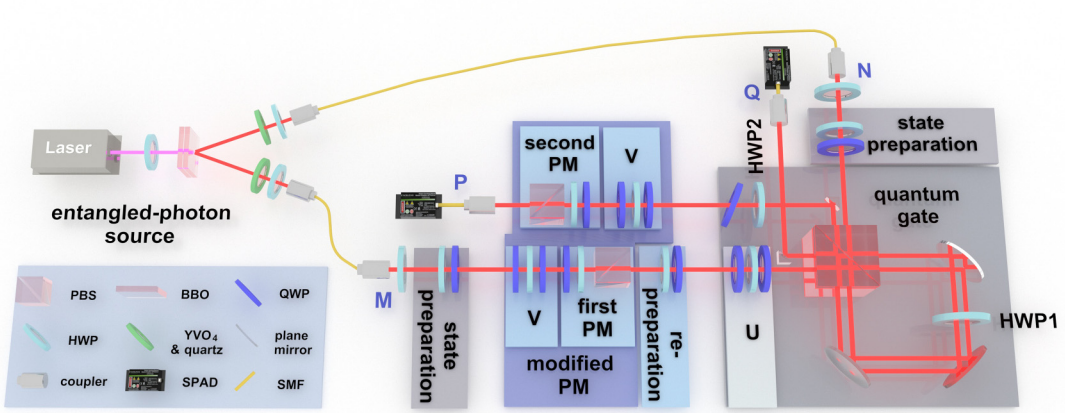

FIG. 2. Experimental setup. The checkpoints M, N, P, and Q match the notations in Fig. 1. The monochromatic panels were inside the quantum comb and the blue panels were accessible to the causal discovery algorithm. Entangled photons were generated by spontaneous parametric down-conversion. The initial state ρ, the unitary channel U, and the modifications of Pauli measurement (PM) basis V were realized by groups of wave plates, while the PMs themselves exploited an additional polarization beam splitter (PBS). The SWAP gate controlled by HWP1 and HWP2 can switch between DC and CC and the tilted QWP was for correcting the intragate polarization dispersion. HWP: half-wave plate; QWP: quarter-wave plate; SMF: single-mode fiber; SPAD: single-photon avalanche detectors.

polarization state of a photon pair generated from spontaneous parametric down-conversion, with the horizontal ($|H\rangle$) and vertical ($|V\rangle$) polarization states being the computational basis. The layout of our experimental setup is depicted in Fig. 2. It closely resembled the quantum comb in Fig. 1(c); we have labeled four control points—M, N, P, and Q—in both of the figures to clarify the correspondence. Using the photon source reported in Ref. [24], we obtained the maximally entangled state $(|HH\rangle + |VV\rangle)/\sqrt{2}$ with fidelity over 0.97. Subsequently, we utilized two groups of half-wave plates (HWPs) and quarter-wave plates (QWPs) to prepare the desired initial states, and an HWP between two QWPs to implement the unitary channel operation [25], so both of the operations would induce the same Pauli operations $\mathbf{P}$. A Sagnac-type ring interferometer at the center of the setup was exploited to switch the causal structure. Specifically, the polarizing beam splitter (PBS) at the entrance of the interferometer transmits (reflects) the photons with $|H\rangle$ ($|V\rangle$) polarization. When the orientation of HWP1 and HWP2 were set at $0°(45°)$, the photon that entered the setup from M will exit at P(Q) while keeping its polarization, so the ring would act as an identity (SWAP) gate and cause the photons at X and Y to be linked by DC (CC). Before and after the ring, we inserted a polarization measurement module containing a QWP, an HWP, and a PBS to implement a single-qubit projection onto the eigenstates of the Pauli operators. With two-photon coincidence counting, we can extract all the required conditional probabilities and determine the two-point correlations. The wave plates after the PBS in the first measurement reprepared the Pauli eigenstates. Due to the high degree of two-photon entanglement, the marginal state of one of the photons was maximally mixed, so the procedure did not introduce signaling and the causal discovery was strictly observational. Another group of wave plates introduced the V- and V' operations to modify the measurement directions.

We tested two families of temporally ordered quantum systems, whose causality is not unveiled by their Pauli correlations, to demonstrate the quantum causal discovery algorithm. The first family satisfies $C_{22} - C_{11} = 1, C_{33} = 0$; geometrically, the correlation is located on an edge of the indistinguishable octahedron where the DC and CC tetrahedra overlap. Although the two causal mechanisms can induce the same Pauli correlations, their induced correlations using the algorithmically determined operators are drastically different: in Fig. 3(a), we plot the trajectories of the Pauli correlation $\mathbf{P}$ and the correlations under the modified measurement settings, $\mathbf{P}_V(U)$ (blue) and $\mathbf{P}_V(\rho)$ (red). We find that whenever the causal structure of $\mathbf{P}$ is DC, $\mathbf{P}_V(U)$ will be close to the $C_{33}^V(U) = 1$ plane. The observation is quantitatively supported using the DC criterion $1 - C_{33}^V < \varepsilon$ as shown in Fig. 3(b), where we set $\delta = 0.15$ and the cutoff value $\varepsilon = 0.075$ successfully delineated all DC and CC far away from the exceptional plane. We also studied the cases of using mixed states as the source of CC to enrich the validity of this method, which is presented later.

Further, we corroborate that the algorithm can robustly handle the envisaged exception for Pauli correlations close to $\sum_{k=1}^{3} C_{kk} = 1$. As shown in the inset of Fig. 3(b), the DC criterion using P_V as input fails to identify the CC; however, the issue is fully resolved by running a second round of the algorithm. The resulted new correlations [cf. Fig. 3(a)], $\mathbf{P}_{V'}(U)$ (cyan) and $\mathbf{P}_{V'}(\rho)$ (pink) definitely fall into the distinguishable regions. We further studied the case of another family of Pauli correlations moving on the exceptional plane, with the result depicted in Fig. 3(c). Clearly, the separation of $\mathbf{P}_{V'}(U)$ and $\mathbf{P}_{V'}(\rho)$ is robust and insensitive to the exact form of the channel or state. To quantify the robustness of the algorithm we resort to another distance criterion for DC, $\mathcal{D}[\mathbf{P}_{V'}, \mathbf{P}(\sigma_3)] < \varepsilon'$. Here, the calligraphic $\mathcal{D}$ indicates the Euclidean distance of the two correlation vectors, and we conveniently select $\varepsilon' = 1/\sqrt{3}$. The results corresponding to the points running into and on the exceptional plane are given in Figs. 3(d) and 3(e). For all data points, the causal mechanisms were correctly identified beyond 70 standard deviations from the cutoff value. The results clearly

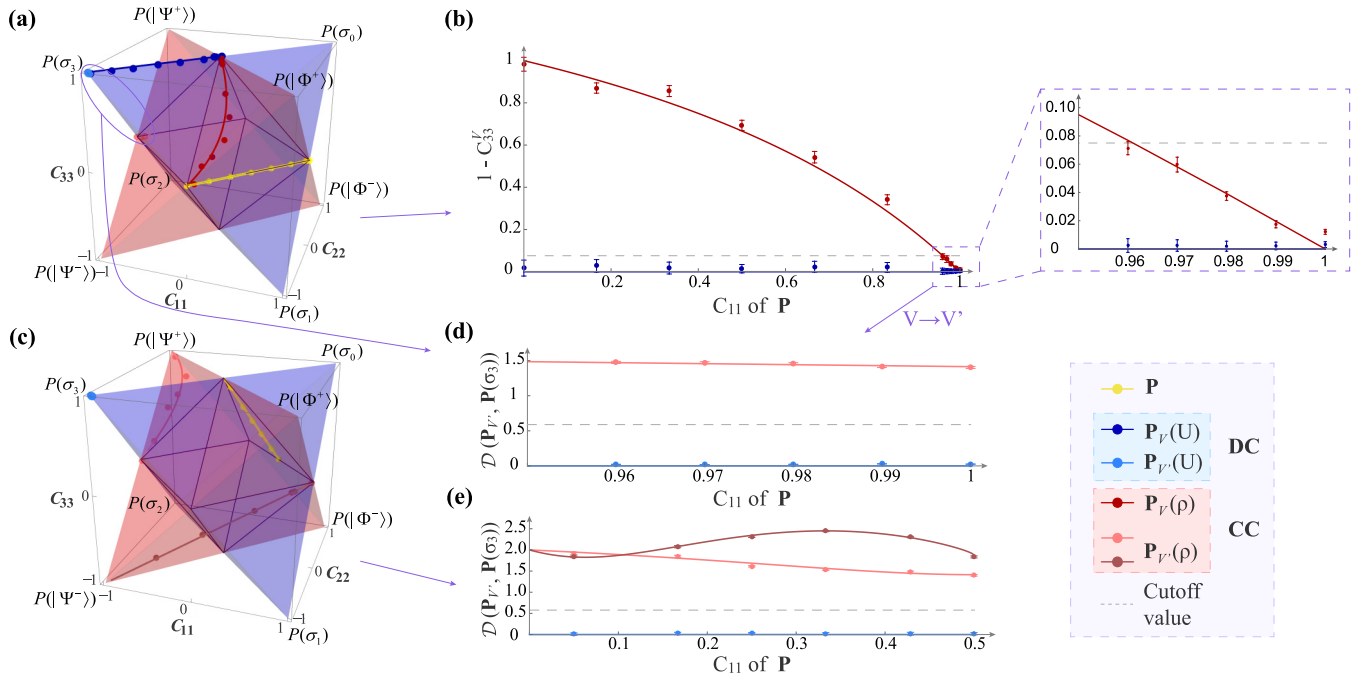

FIG. 3. Experimental results for pure initial states. (a) Visualization of $\mathbf{P}_V$ when $\mathbf{P}$ moves on one edge of the causally indistinguishable octahedron. The $\mathbf{P}_{V'}$ are evaluated only when the $\mathbf{P}$ are close to the exceptional plane. (b) Predicted (line) and measured (data points) DC criterion correspond to $\mathbf{P}_V$. The inset enlarges the indistinguishable part. (c) Visualization of $\mathbf{P}_{V'}$ when $\mathbf{P}$ moves on the exceptional plane. (d and e) Predicted and measured distance criterion for DC correspond to the points $\mathbf{P}_{V'}$ in (a) and (c), respectively. The pink and brown lines show the effect of different phases of Bell states in state preparation. Error bars are calculated via the Monte Carlo method, and except for the 1σ in the inset of (b), all the other error bars are 3σ.

demonstrated the stability of the quantum causal identification algorithm.

Having shown the cases of the source of CC being pure, we now set our sights on mixed sources in which the causal structure can also be identified with the DC criterion discussed above. We tested cases where $\mathbf{P}$ are picked inside the octahedron, from two vertical lines (yellow dots and lines in Fig. 4). Following the procedure of quantum state purification, we prepared the initial states as $\rho = q\frac{\mathbf{I}}{4} + (1-q)\rho_{\text{pure}}$, where $0 \leqslant q \leqslant 1$. Since ρ mapping to $\mathbf{P}(\rho)$ has innumerable preparation methods with different q and ρ_{pure}, for simplicity, the ρ_{pure} we used are generated by $|\Phi^+\rangle$ accompanied by local operations and the corresponding q, and the experiment results are displayed in Fig. 4. The remarkable distance between $C_{33}^V(\rho)$ and 1 indicates a lucid distinguishing of DC and CC, implying the algorithm remains in force under various preparation methods.

The extension from pure to mixed initial states is geometrically intuitive: our algorithm is a unitary operation with respect to the initial states and $\mathbf{P}_V(\rho)$ shares the same linear

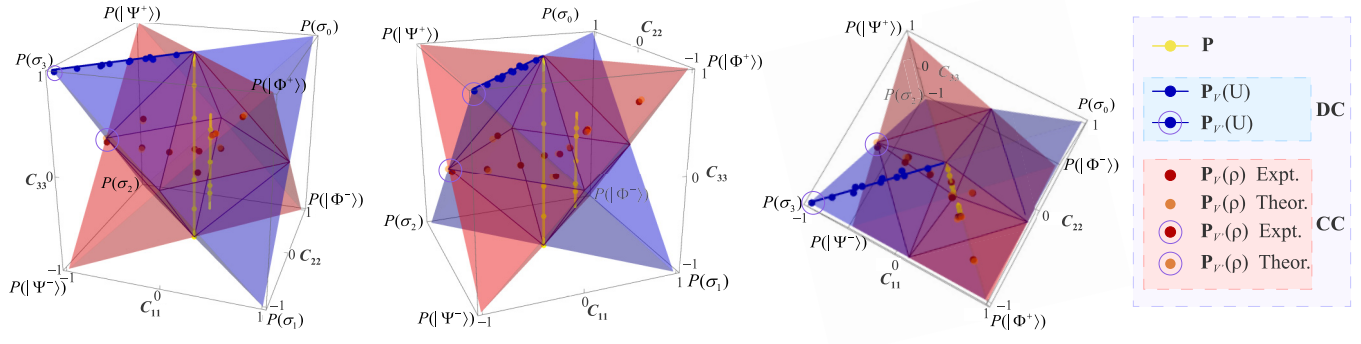

FIG. 4. Experimental results for mixed initial states in geometric description. In this plot, modification on two measurements are the same (usual cases) except for two points (exceptional cases) being circled out. Yellow dots and lines: $\mathbf{P}$ chosen on two vertical lines inside the octahedron, with mixed initial states ρ in the overlapping area, namely, the octahedron, and their trajectories; blue dots and lines: experiment results of $\mathbf{P}_V(U)$ and their theoretical trajectories; red dots, experiment results of $\mathbf{P}_V(\rho)$. The orange dots are the theoretical counterpart of the red dots. The results are plotted with two side views and one top view.

decomposition with $\rho = \sum p_k \rho_k$—the vector $\mathbf{P}_V(\rho)$ is the vector sum of the vectors $\mathbf{P}_V(\rho_k)$, where ρ_k is either pure or mixed. In the above exhibition, we have mentioned that it is the quantity $\sum_{k=1}^{3} C_{kk}$ that affects and bounds $C_{33}^V(\rho)$ (Appendixes B and C) and in turn affects the distinguishing outcome with the DC criterion $1 - C_{33}^V < \varepsilon$. Then, consider preparing the target ρ with pure states ρ_k in which $\mathbf{P}(\rho_k)$ are on a single plane $C_{11} + C_{22} + C_{33} = b$, which is feasible for coefficients of the vectors p_k adding to 1. We can conclude that $\mathbf{P}_V(\rho)$ behaves similarly whether ρ is pure or mixed, i.e., $C_{33}^V(\rho)$ is upper bounded by the same value. However, unlike the regular theoretical line presented in Fig. 3(a), $\mathbf{P}_V(\rho)$ in Fig. 4 are spread quite randomly. This is because the relation between $\mathbf{P}_V(\rho)$ and ρ is not unique when ρ is mixed. Finally, we remark that the procedure above is also valid for the exceptional part of our algorithm, and we present one example circled in Fig. 4.

V. DISCUSSION

We have conducted experimental investigations to identify the underlying causal structures of two-point quantum correlations in the framework of quantum comb representation and geometric description using a photonic platform. Our algorithm is built upon quantum correlation and unitary operations within an observational scheme, eliminating the need for interventions, which sets it apart from its classical counterpart and results in reduced resource requirements. Besides the framework we have used, physicists have also researched other formalisms such as pseudodensity matrix formalism [11,26] and process matrix formalism [27–29], to properly describe causal structures and seek less intervening and go on less resource-consuming inference methods [30]. Digging out more causal inference criteria that take the advantages of the quantum world could be a natural trend. Therefore, we hope that our comprehensive exploration of two-point quantum correlations from a causal perspective will contribute to the further advancement of quantum causal inference and pave the way for the construction of causal networks involving various quantum resources.

ACKNOWLEDGMENTS

This work was supported by the Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0301200), the National Natural Science Foundation of China (Grants No. 12174370, No. 12174376, No. 11821404, and No. 11904356), the Youth Innovation Promotion Association of Chinese Academy of Sciences (Grant No. 2017492), the Open Research Projects of Zhejiang Lab (Grant No. 2021MB0AB02), the Fok Ying Tong Education Foundation (Grant No. 171007), and the Fundamental Research Funds for the Central Universities (Grant No. WK2030000008).

APPENDIX A: RELOCATION BETWEEN THE INITIAL STATE ρ OR EVOLUTION U AND P

Inheriting the labeling in Ref. [20], we first look into cases with latent structure CC. Consider an arbitrary pure initial state $|\phi\rangle$ that can be decomposed as

$$|\phi\rangle = \sum_{i=1}^{4} w_i e^{i\theta_i} |b_i\rangle, \tag{A1}$$

where w_i and θ_i are real numbers satisfying $\sum_{i=1}^{4} w_i^2 = 1$, and $|b_i\rangle (i = 1, 2, 3, 4)$ denote four Bell states $|\Phi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^+\rangle$, $|\Psi^-\rangle$ in order. A straightforward calculation indicates that the phase term $e^{i\theta_i}$ does not contribute to $\mathbf{P}$ and there is a one-to-one map from $\mathbf{P}(|\phi\rangle)$ to a pure state with $\{w_i\}$, yielding $\mathbf{P}(|\phi\rangle) = \sum_{i=1}^{4} w_i^2 \mathbf{P}(|b_i\rangle) = (-\omega_1^2 - \omega_2^2 + \omega_3^2 + \omega_4^2, -\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2, -\omega_1^2 + \omega_2^2 + \omega_3^2 - \omega_4^2)^{\mathrm{T}}$, where $\mathbf{P}(|b_i\rangle)$ are the vertices of the CC tetrahedron ($\mathcal{T}_{\mathrm{CC}}$).

If the latent structure is DC, $\mathbf{P}$ also has a unique decomposition $\mathbf{P}(U) = \sum_{j=0}^{3} p_j \mathbf{P}(\sigma_j) = (p_0 + p_1 - p_2 - p_3, -p_0 + p_1 + p_2 - p_3, p_0 - p_1 + p_2 - p_3)^{\mathrm{T}}$, where the sum of nonnegative real numbers $\{p_j\}$ is unity, $\sigma_j (j = 0, 1, 2, 3)$ are Pauli matrices, and $\mathbf{P}(\sigma_i)$ are the vertices of the DC tetrahedron ($\mathcal{T}_{\mathrm{DC}}$). Given a $\mathbf{P}(U)$, U takes the form

$$U = e^{i\frac{\theta}{2}} \begin{pmatrix} e^{i\gamma_1} \cos(\varphi_0) & e^{i\gamma_2} \sin(\varphi_0) \\ -e^{-i\gamma_2} \sin(\varphi_0) & e^{-i\gamma_1} \cos(\varphi_0) \end{pmatrix}, \tag{A2}$$

where $\frac{\theta}{2} \in [0, \pi)$, $\cos(\varphi_0) = \sqrt{c_1}$, $\sin(\varphi_0) = \sqrt{d_1}$, $\gamma_1 = (-1)^{n_1} \arccos(\frac{c_2}{c_1})/2 + k_1\pi$ (if $c_1 = 0$, let $\gamma_1 = 0$), $\gamma_2 = (-1)^{n_2} \arccos(\frac{d_2}{d_1})/2 + k_2\pi$ (if $d_1 = 0$, let $\gamma_2 = 0$), $c_1 = p_0 + p_3$, $c_2 = p_0 - p_3$, and $d_1 = p_1 + p_2$, $d_2 = p_1 - p_2$, with $n_1, n_2, k_1, k_2 \in \{0, 1\}$. Although there are in total 16 different U corresponding to the $\mathbf{P}(U)$ up to a global phase, these candidates together with V derived from them only result in no more than four different $\mathbf{P}_V(U)$, among which desired $\mathbf{P}_V(U)$ with $C_{33}^V(U) = 1$ can always be found. In our experiment and the simulation to be carried out, we let $n_1 = n_2 = k_1 = k_2 = 0$.

In the following analysis, as we use a geometric method to distinguish causal structures, we need to focus on a specific plane $C_{11} + C_{22} + C_{33} = b (-1 \leqslant b \leqslant 1)$ that $\mathbf{P}$ is on. For a given $\mathbf{P}$, b equals either $\omega_1^2 + \omega_2^2 + \omega_3^2 - 3\omega_4^2 = 1 - 4\omega_4^2$ for $\mathbf{P}(|\phi\rangle)$ or $3p_0 - p_1 - p_2 - p_3 = 4p_0 - 1$ for $\mathbf{P}(U)$.

APPENDIX B: EFFECT AND THE CONSTRUCTION OF DESIRED MODIFICATION V FOR USUAL CASES

Consider a $\mathbf{P}$ with its entries summing up to b and two Pauli measurements modified by the same unitary operation V reading

$$V = \begin{pmatrix} e^{i\psi} \cos(\varphi) & e^{i\chi} \sin(\varphi) \\ -e^{-i\chi} \sin(\varphi) & e^{-i\psi} \cos(\varphi) \end{pmatrix}. \tag{B1}$$

For DC, exploiting $\mathbf{P}_V(U) = \mathbf{P}(V^\dagger U V)$, by careful calculation one derives the entries of $\mathbf{P}_V(U)$ satisfying $C_{11}^V(U) = 2(a_1^2 + b_2^2) - 1$, $C_{22}^V(U) = 2(a_1^2 + b_1^2) - 1$,

and $C_{33}^{V}(U) = 2(a_1^2 + a_2^2) - 1$, where

$$\begin{aligned} a_1 &= \cos(\varphi_0)\cos(\gamma_1), \\ a_2 &= \cos(\varphi_0)\cos(2\varphi)\sin(\gamma_1) - \sin(\varphi_0)\sin(2\varphi)\sin(\gamma_2 - \psi - \chi), \\ b_1 &= \sin(\varphi_0)\cos^2(\varphi)\cos(\gamma_2 - 2\psi) + \sin(\varphi_0)\sin^2(\varphi)\cos(\gamma_2 - 2\chi) - \cos(\varphi_0)\sin(2\varphi)\sin(\chi - \psi)\sin(\gamma_1), \\ b_2 &= -\sin(\varphi_0)\sin^2(\varphi)\sin(\gamma_2 - 2\chi) + \sin(\varphi_0)\cos^2(\varphi)\sin(\gamma_2 - 2\psi) + \cos(\varphi_0)\sin(2\varphi)\cos(\chi - \psi)\sin(\gamma_1), \end{aligned}$$

$$a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1. \tag{B2}$$

On one hand, noticing that the sum of these entries is $4a_1^2 - 1 = 4p_0 - 1 = b$, which is independent of V, the effect of V for DC is moving a $\mathbf{P}$ to a $\mathbf{P}_V(U)$ lying on a fix plane $C_{11}^{V}(U) + C_{22}^{V}(U) + C_{33}^{V}(U) = b$, which is exactly the same plane that $\mathbf{P}$ is on. On the other hand, $\mathbf{P}_V(U)$ is trapped in the $\mathcal{T}_{\mathrm{DC}}$, thus geometrically $\mathbf{P}_V(U)$ can be moved to the top of the $\mathcal{T}_{\mathrm{DC}}$ and out of the overlapping area, terminating at $(\frac{b-1}{2}, \frac{b-1}{2}, 1)^{\mathrm{T}}$. To do so, we can choose the parameters of U as

$$\psi + \chi = \gamma_2 - \frac{k_1\pi}{2}, \quad \varphi = \frac{k_2\pi - \omega}{2}, \tag{B3}$$

where $\sin(\omega) = \sin(\varphi_0)\sin(\gamma_2 - \psi - \chi)/r$, $\cos(\omega) = \cos(\varphi_0)\sin(\gamma_1)/r$, $r = [\cos^2(\varphi_0)\sin^2(\gamma_1) + \sin^2(\varphi_0)\sin^2(\gamma_2 - \psi - \chi)]^{1/2}$, $k_1 = 1, 2$, and $k_2 = 1, 2$. In our experiment and the simulation below, we always set $k_1 = k_2 = 1$ and $\psi = \chi$.

For CC, with a pure initial state $|\phi\rangle$ defined in Eq. (A1), the sum of entries of $\mathbf{P}_V(\rho)$ is calculated to be also equal to $1 - 4w_4^2 = b$, independent of V. Therefore, $\mathbf{P}_V(\rho)$ lies on plane $C_{11}^{V}(|\phi\rangle) + C_{22}^{V}(|\phi\rangle) + C_{33}^{V}(|\phi\rangle) = b$ which coincides with the plane $\mathbf{P}$ and $\mathbf{P}_V(U)$ lie on. Again, as $\mathbf{P}_V(\rho)$ is not able to leave the $\mathcal{T}_{\mathrm{CC}}$, the boundaries of $\mathbf{P}_V(\rho)$ are the intersection lines of $C_{11}^{V}(|\phi\rangle) + C_{22}^{V}(|\phi\rangle) + C_{33}^{V}(|\phi\rangle) = b$ and the surface of the $\mathcal{T}_{\mathrm{CC}}$ [one of which is $-C_{11}^{V}(|\phi\rangle) - C_{22}^{V}(|\phi\rangle) + C_{33}^{V}(|\phi\rangle) = 1$]. Consequently, $C_{33}^{V}(|\phi\rangle)$ is upper bounded by $\frac{1+b}{2}$ and $1 - C_{33}^{V}(|\phi\rangle) \leqslant \frac{1-b}{2}$.

The above analysis shows that the modification V never drags P out of the plane $C_{11} + C_{22} + C_{33} = b$ regardless of the latent structure, which is the core of our algorithm handling usual cases.

APPENDIX C: HANDLING EXCEPTIONS CASES WITH UNEQUAL MODIFICATIONS

The solution is decomposed into three steps: run the standard algorithm above to derive an operator V_1 resulting in $\mathbf{P}_{V_1}(U) = (\frac{b-1}{2}, \frac{b-1}{2}, 1)^{\mathrm{T}}$; use operators $V_- = \sigma_1$ and $V_+ = I$ to flip the sign of the last two entries of $\mathbf{P}_{V_1}(U)$ and move it $(\frac{b-1}{2}, \frac{1-b}{2}, -1)^{\mathrm{T}}$; and run the standard algorithm again deriving another operator V_2 which moves the point on plane $C_{11}^{V'}(U) + C_{22}^{V'}(U) + C_{33}^{V'}(U) = -1$ and terminates at $\mathbf{P}_{V'}(U) = (-1, -1, 1)^{\mathrm{T}} = \mathbf{P}(\sigma_3)$. Next, we investigate the behavior of $\mathbf{P}_{V'}(\rho)$ (CC) by calculation and numerical simulation, taking cases where $b = 1$ as examples [Figs. 3(c) and 3(e) in the main text]. $\mathbf{P}(|\phi\rangle)$ is explicitly written in terms of $w_i (i = 1, 2, 3, 4)$ as

$$\mathbf{P}(|\phi\rangle) = \begin{pmatrix} w_1^2 - w_2^2 + w_3^2 - w_4^2 \\ -w_1^2 + w_2^2 + w_3^2 - w_4^2 \\ w_1^2 + w_2^2 + w_3^2 - w_4^2 \end{pmatrix}. \tag{C1}$$

Combining $b = 1$ and the normalization condition of $|\phi\rangle$, w_4 has to be zero. When $b = 1$, V_1 and $V_\pm$ move $\mathbf{P}(U)$ to $(0, 0, -1)^{\mathrm{T}}$ and V_2 further moves it to $\mathbf{P}_{V'}(U) = (-1, -1, -1)^{\mathrm{T}}$, requiring the parameters of V_2 to satisfy $\phi = \frac{\pi}{4}$, and $\psi = \chi = \pm\frac{\pi}{8}$ or $\pm\frac{3\pi}{8}$. The different choices of ψ and χ correspond to different candidates of U. The calculation result of applying $V_\pm$ and V_2 reads

$$\mathbf{P}_{V'}(|\phi\rangle) = \begin{pmatrix} w_3^2 - w_1^2 - w_2^2 \\ w_1^2 - w_2^2 \\ w_1^2 - w_2^2 \end{pmatrix}, \tag{C2}$$

indicating that $V_\pm$ and V_2 moves every point on plane $C_{11} + C_{22} + C_{33} = 1$ to plane $C_{22}^{V'} = C_{33}^{V'}$. Since V_1 never drags $\mathbf{P}(|\phi\rangle)$ outside plane $C_{11} + C_{22} + C_{33} = 1$, there is no need to do further calculation for V_1. In the numerical simulation in Fig. 5, we show that different phases θ_i of $|\phi\rangle$ may lead to an enormous difference of $\mathbf{P}'_V(|\phi\rangle)$ (brown and pink dots) and $C_{33}^{V'}$ can be very close to 1, thus we cannot use $C_{33}^{V'}$ to distinguish DC and CC.

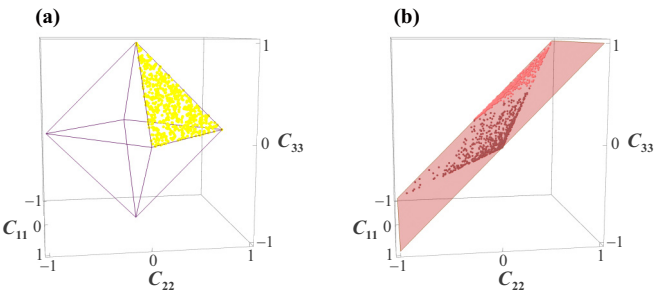

FIG. 5. Effect of unequal modification for CC cases by simulation. (a) Over 1000 $\mathbf{P}(|\phi\rangle)$ (yellow dots) are randomly spread on plane $C_{11} + C_{22} + C_{33} = 1$ in the overlapping area. (b) Brown and pink [matching the color in Fig. 3(c)] dots are the simulation results of $\mathbf{P}_{V'}(|\phi\rangle)$ when the phase terms in Eq. (A1) are $\{\theta_1, \theta_2, \theta_3, \theta_4\} = \{0, -\frac{\pi}{2}, \frac{\pi}{2}, 0\}$ and $\{0, \frac{\pi}{2}, -\frac{\pi}{2}, 0\}$, respectively. The light purple plane is $C_{22}^{V'} = C_{33}^{V'}$.

APPENDIX D: THE OTHER PROJECTIONS OF FIG. 3

See Fig. 6.

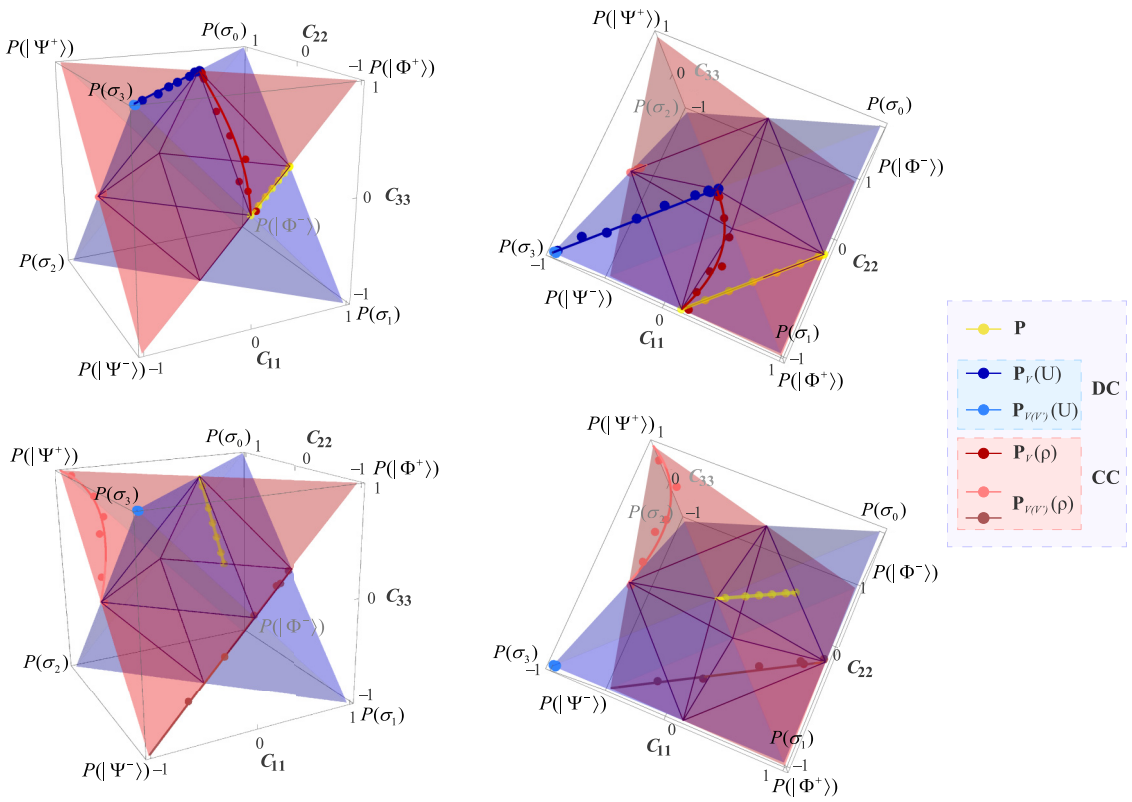

FIG. 6. The other projections of Fig. 3. The other side view (left) and the top view (right) of Fig. 3(a) (upper) and 3(c) (lower) in the main text, respectively.

[1] A. J. DeGrave, J. D. Janizek, and S.-I. Lee, AI for radiographic COVID-19 detection selects shortcuts over signal, Nat. Mach. Intell. **3**, 610 (2021).

[2] N. Savage, Why artificial intelligence needs to understand consequences, Nature (London) (2023).

[3] P. Cui and S. Athey, Stable learning establishes some common ground between causal inference and machine learning, Nat. Mach. Intell. **4**, 110 (2022).

[4] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms* (MIT, Cambridge, 2017).

[5] H. Reichenbach and M. Reichenbach, *The Direction of Time* (University of California Press, Oakland, 1991).

[6] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).

[7] S. Kochen and E. P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. **17**, 59 (1967).

[8] A. J. Leggett and A. Garg, Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? Phys. Rev. Lett. **54**, 857 (1985).

[9] E. G. Cavalcanti and R. Lal, On modifications of Reichenbach's principle of common cause in light of Bell's theorem, J. Phys. A: Math. Theor. **47**, 424018 (2014).

[10] J. Henson, R. Lal, and M. F. Pusey, Theory-independent limits on correlations from generalized Bayesian networks, New J. Phys. **16**, 113043 (2014).

[11] J. F. Fitzsimons, J. A. Jones, and V. Vedral, Quantum correlations which imply causation, Sci. Rep. **5**, 18281 (2015).

[12] K. Ried, M. Agnew, L. Vermeyden, D. Janzing, R. W. Spekkens, and K. J. Resch, A quantum advantage for inferring causal structure, Nat. Phys. **11**, 414 (2015).

[13] R. Chaves, C. Majenz, and D. Gross, Information–theoretic implications of quantum causal structures, Nat. Commun. **6**, 5766 (2015).

[14] F. Costa and S. Shrapnel, Quantum causal modelling, New J. Phys. **18**, 063032 (2016).

[15] John-Mark A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens, Quantum common cause and quantum causal models, Phys. Rev. X **7**, 031021 (2017).

[16] R. Chaves, G. Carvacho, I. Agresti, V. D. Giulio, L. Aolita, S. Giacomini, and F. Sciarrino, Quantum violation of an instrumental test, Nat. Phys. **14**, 291 (2018).

[17] M. Hu and Y. Hou, Discrimination between quantum common causes and quantum causality, Phys. Rev. A **97**, 062125 (2018).

[18] J. Barrett, R. Lorenz, and O. Oreshkov, Quantum causal models, arXiv:1906.10726.

[19] S. Milz, J. Bavaresco, and G. Chiribella, Resource theory of causal connection, Quantum **6**, 788 (2022).
[20] C. Zhang, Y. Hou, and D. Song, Quantum observation scheme universally identifying causalities from correlations, Phys. Rev. A **101**, 062103 (2020).
[21] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum circuit architecture, Phys. Rev. Lett. **101**, 060401 (2008).
[22] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).
[23] G. Lüders, Über die zustandsänderung durch den messprozeß, Ann. Phys. (Berlin) **443**, 322 (1950).
[24] C. Zhang, Y.-F. Huang, Z. Wang, B.-H. Liu, C.-F. Li, and G.-C. Guo, Experimental Greenberger-Horne-Zeilinger-type six-photon quantum nonlocality, Phys. Rev. Lett. **115**, 260402 (2015).
[25] B.-G. Englert, C. Kurtsiefer, and H. Weinfurter, Universal unitary gate for single-photon two-qubit states, Phys. Rev. A **63**, 032303 (2001).
[26] C. Marletto, V. Vedral, S. Virzì, A. Avella, F. Piacentini, M. Gramegna, I. P. Degiovanni, and M. Genovese, Temporal teleportation with pseudo-density operators: How dynamics emerges from temporal entanglement, Sci. Adv. **7**, eabe4742 (2021).
[27] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012).
[28] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).
[29] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, Rep. Math. Phys. **3**, 275 (1972).
[30] X. Liu, Y. Qiu, O. Dahlsten, and V. Vedral, Quantum causal inference with extremely light touch, arXiv:2303.10544.