Nonsignaling causal hierarchy of general multisource networks

Ming-Xing Luo *

Information Security and National Computing Grid Laboratory, Southwest Jiaotong University, Chengdu 610031, China

(Received 24 November 2019; revised manuscript received 18 May 2020; accepted 20 May 2020; published 9 June 2020)

Large-scale multisource networks have been employed to overcome the practical constraints that entangled systems are difficult to faithfully transmit over large distances or store over long times. However, a full characterization of the multipartite nonlocality of these networks remains out of reach, mainly due to the complexity of multipartite causal models. In this paper, we propose a general framework of Bayesian networks to reveal connections among different causal structures. The present model implies a special star-convex set of nonsignaling correlations from multisource networks that allows the construction of a polynomial-time algorithm for solving the compatibility problem of a given correlation distribution and a fixed causal network. It is then used to classify the nonlocality originating from the standard entanglement swapping of tripartite networks. Our model provides a unified device-independent information processing method for exploring the practical security against nonsignaling eavesdroppers in multisource quantum networks.

DOI: 10.1103/PhysRevA.101.062317

I. INTRODUCTION

Bell theorem states that by performing local measurements on an entangled system, remote observers can create nonlocal correlations, which are witnessed by violating a special inequality [1–3]. These correlations cannot be precisely predicted by any classical local model with the causal assumption that the measurement outcomes depend on shared local variables and freely chosen observables. Nevertheless, the nonsignaling condition allows local agents to build classical correlations going beyond to all quantum correlations [4]. Thus, it is important to further investigate what causal assumptions for a classical model are efficient to reproduce nonlocal correlations [5–9]. Interestingly, all bipartite quantum correlations are classically generated by relaxing either the local assumption or the realism causal assumption [10]. For multipartite scenarios, genuinely multipartite nonlocalities are introduced to characterize new quantum nonlocalities [11–16].

In comparison to the bipartite case, it is difficult to characterize most multipartite nonlocal correlations because of the exponential number of free parameters. Recently, the Bayesian network has been used to reveal the connections among different causal structures [17]. This model is efficient for depicting all the nonlocality classes in tripartite scenarios and exploring new nonlocal causal structures. Unfortunately, the potential applications of single entangled systems are limited because of practical constraints such as the transmission distance and storage time. Large-scale multisource networks are then proposed [18–22], shown schematically in Fig. 1. Differently from a Bell network consisting of one entanglement, all observers in multisource quantum networks are allowed to perform local joint measurements on different entangled systems. Remarkably, several spacelike separated observers without prior shared entanglement can create new nonlocal correlations with the help of others' local measurements. One typical example is standard entanglement swapping using Einstein-Podolsky-Rosen (EPR) states [2,23] to remotely generate a long-distance entangled singlet with local operations and classical communication beyond classical correlations [24]. Several Bell-type inequalities have recently been proposed to feature the extraordinary nonmultilocality of statistics obtained from local measurements on these quantum networks [25,26] or general quantum networks [27–30]. However, a unified model for causal relaxations, together with the nonmultilocality they lead to, is still an open problem [31].

Bell theory is the foundation for various fields, such as quantum information processing [32,33], unconditionally secure key distributions [34–36], randomness amplification [37–39], and quantum supremacy [40,41]. In most cases, the trustworthiness of quantum devices according to specification should be avoided in order to ensure adversary (noise)-tolerant realizations. Hence, the so-called device-independent protocols depend only on the statistics of measurement outputs. More importantly, precise Bell inequalities can be constructed to bound the leaked information for an eavesdropper or the secure key rate in quantum key distributions [34–36]. A natural problem is how to extend these results on single-source quantum networks to be suitable for multisource quantum networks.

Our goal in this paper is to investigate a causal model for general multisource networks. We develop a systematic way to characterize causal relaxations of Bell correlations for multisource networks using generalized Bayesian networks [31]. The compatibility problem of a given nonsignaling correlation and Bell-type directed acyclic graph is then formalized into a star-convex programming problem that can be solved by a

*mxluo@home.swjtu.edu.cn

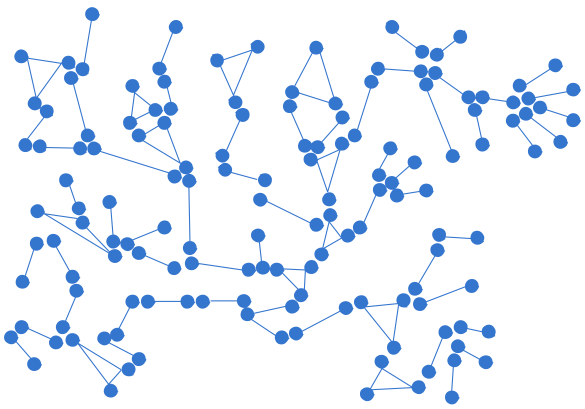

FIG. 1. Schematic multisource network. There are no fewer than two independent sources that distribute states to different spacelike separated agents.

polynomial-time algorithm. This result goes beyond the convex correlation polytope defined by single-source networks [9,33]. We then classify the causal structure of a tripartite network derived from entanglement swapping by presenting a full characterization of Bell localities. Interestingly, the new model is also useful to bring out a device-independent information processing model under the source-independence assumption. Specifically, for an eavesdropper who holds independent systems, the violation of Bell inequalities [29] provides the upper bound of the leaked information about the outcomes of justifiable agents in multisource networks consisting of generalized EPR and Greenberger-Horne-Zeilinger (GHZ) states [42]. This goes far beyond previous results on Bell networks [43] or chain-shaped and star-shaped networks [44].

The rest of this paper is organized as follows. In Sec. II, we propose a causal structure of generally multisource networks according to the causal implication. This model can be used to solve the compatibility problem, i.e., whether a given correlation distribution is compatible with a fixed causal relation of a given network. Section III provides the Bell classes of multipartite causal networks, especially for a network with three nodes and two hidden variables. Section IV is devoted to the monogamy relationship for device-independent information processing on multisource quantum networks. Section V proposes some examples, while Sec. VI concludes the paper.

II. CAUSAL STRUCTURES OF GENERAL MULTISOURCE NETWORKS

A. Causal structures of multisource networks

Causal structures of multisource networks are schematically represented by directed acyclic graphs (DAGs) [17,31], as shown in Fig. 2. Each node of a DAG represents a classical random variable, and each directed edge encodes a causal relation between two nodes. For each edge, the start vertex is called the parent and the arrival vertex is called the child. Given a set $\mathcal{V}=\{v_1,\ldots,v_n\}$, $\mathcal{V}$ forms a generalized Bayesian network with respect to the DAG $\mathcal{G}$ if the joint probability $P(v_1,\ldots,v_n)$ describing the statistics of $\mathcal{V}$ can be decomposed as

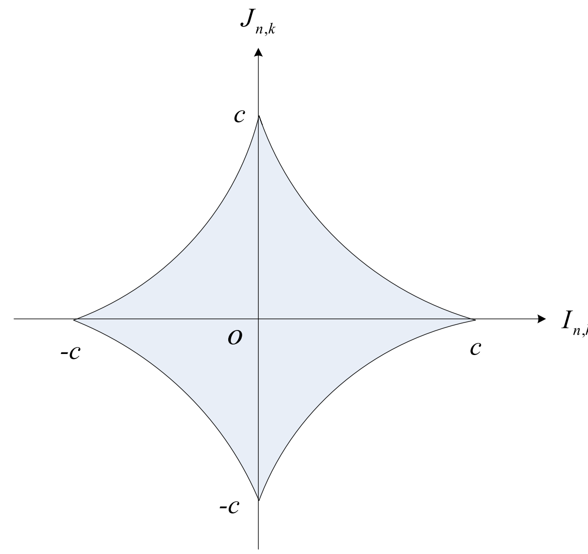

FIG. 2. Correlations defined by $|I_{n,k}|^{\frac{1}{k}}+|J_{n,k}|^{\frac{1}{k}}\leqslant c$ on the projected subspace spanned by $\{I_{n,k},J_{n,k}\}$. It consists of a star-convex set going beyond the convex set derived from single-source networks [1].

$$P(v_1,\ldots,v_n)=\prod_i p(v_i|S_{v_i}), \tag{1}$$

where S_{v_i} denotes the set of parent nodes of v_i in $\mathcal{G}$.

In what follows, we focus on specific causal structures with two common features. One is that they have a set of unobservable nodes, the hidden variables λ_i, and two sets of observables, the inputs x_j and the outputs a_s, i.e., $\mathcal{V}_{\rm BN}=\{\lambda_i,a_j,x_s,\forall i,j,s\}$. The other is that each output a_i contains the input x_i and connected variables $\Lambda_i\subseteq\{\lambda_i,i\}$ as its parents, i.e., $\{x_i,\Lambda_i\}\subseteq S_{a_i}$. These DAGs are called networking Bell DAGs (NBDAGs). They reduce to special Bell DAGs (BDAGs) [17] when $m=1$.

Consider a Bayesian network in terms of the NBDAG. Assume that independent variables $\Lambda:=\lambda_1\ldots\lambda_m$ are shared by remote agents $\mathsf{A}_1,\ldots,\mathsf{A}_n$. Each agent A_i shares variables $\Lambda_i=\lambda_{j_1}\ldots\lambda_{j_{\ell_i}}$. The measure of Λ is given by $\mu(\Lambda)=\prod_{i=1}^m\mu_i(\lambda_i)$, where $(\Omega_i,\Sigma_i,\mu_i)$ denotes the measure space of λ_i, $i=1,\ldots,m$. Then Eq. (1) can be rewritten in terms of the generalized local hidden-variable (GLHV) model as [28,29,31]

$$P(a|x)=\int_{\Omega_1\times\ldots\times\Omega_m}\prod_{i=1}^m d\mu_i(\lambda_i)\prod_{j=1}^n P(a_j|x_j,\Lambda_j), \tag{2}$$

which satisfies the nonsignaling condition [33] as

$$P(a_i|x)=P(a_i|x_i) \tag{3}$$

for all a_i's and x_j's, where $a_i=a_1\ldots a_{i-1}a_{i+1}\ldots a_n$, $x_i=x_1\ldots x_{i-1}x_{i+1}\ldots x_n$. The causal model [17] is based on one hidden variable shared by all parties, while the present n-local model in Eq. (2) is based on independently hidden variables [28]. If all hidden variables λ_i can be correlated, the present model in Eq. (2) reduces to the single-variable model [17]. Other causal structures are obtained using the causal relaxations of these NBDAGs of single-source networks [6,10,45–50]. However, so far, there is no systematic investigation of n-partite causal structures of multisource networks [43,51].

B. The compatibility problem

Consider a multisource network $\mathcal{N}$ consisting of n agents $\mathsf{A}_1, \mathsf{A}_2, \ldots, \mathsf{A}_n$. Each agent A_i shares some independent sources Λ_i of $\lambda_1, \lambda_2, \ldots, \lambda_m$. x_i and a_i denote the respective input and output of agent A_i. Let $|x_i|$ and $|a_i|$ be the number of inputs and outputs of the ith agent, respectively, $i = 1, 2, \ldots, n$. Here, each multipartite correlation P is regarded as a vector with components $P_{a,x} := P(a|x)$ in the real space $\mathbb{R}^d$ with the dimension $d = \prod_{i=1}^n |a_i| \times |x_i|$, where $a = a_1a_2\ldots a_n$ and $x = x_1x_2\ldots x_n$.

Definition 1. $P(a|x)$ is compatible with an IONBDAG with given inputs $\{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n\}$ if it can be decomposed as

$$P(a|x) = \textstyle\sum_{\lambda_1,\ldots,\lambda_m} \prod_{i=1}^n R_{\Lambda_i}^{(i)}(a_i|\mathbf{I}_i) \prod_{j=1}^m \mu_i(\lambda_j), \quad (4)$$

where $|\mathbf{I}_i|$ denotes the number of parent inputs of the ith output, and μ_i is the probability distribution of the source λ_i. $R_{\Lambda_i}^{(i)}$ denotes the local deterministic response function of the ith output a_i of agent A_i given the parent inputs $\mathbf{I}_i$ for the local deterministic sources Λ_i, $i = 1, 2, \ldots, n$.

Each $R_{\Lambda_i}^{(i)}$ in Eq. (4) can be represented by

$$R_{\Lambda_i}^{(i)}(a_i|\mathbf{I}_i) := \delta_{a_i, f_{\Lambda_i}^{(i)}}(\mathbf{I}_i), \quad (5)$$

where δ denotes the Kronecker delta and $f_{\Lambda_i}^{(i)}$ is the local deterministic assignment of $\mathbf{I}_i$ into a_i depending on the sources Λ_i. The Λth global deterministic response function is given by the product

$$R_\Lambda := R_{\Lambda_1}^{(1)} \times R_{\Lambda_2}^{(2)} \times \ldots \times R_{\Lambda_n}^{(n)}. \quad (6)$$

Note that R_Λ can also be represented by a vector in $\mathbb{R}^d$, with components $R_{\Lambda,a,x} := R_\Lambda(a|x) = (R_{\Lambda_1}^{(1)}(a_1|\mathbf{I}_1), R_{\Lambda_2}^{(2)}(a_2|\mathbf{I}_2), \ldots, R_{\Lambda_n}^{(n)}(a_n|\mathbf{I}_n))$.

If the sets Λ_i are nonintersecting, P consists of a polytope in the real space $\mathbb{R}^d$, which is defined by the convex hull of a finite number of external points, where each external point is given by the vector R_{Λ_i} for different sources λ_i. Hence, the problem of determining whether a given correlation P is compatible with a Bayesian network with respect to the inputs $\{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n\}$ (in the causal polytope) is equivalent to solving a linear programming problem that can be solved using the standard convex-optimization tools [17,52].

However, Λ_i's intersect for general networks with multiple sources. In this case, P actually consists of a star-convex set in the real space $\mathbb{R}^d$ going beyond the convex set. In fact, for a general network with more than two independent agents who have no prior shared sources, it is easy to prove that all the nonsignaling correlations P satisfy the inequality

$$\mathcal{R}_{\rm ns} := |I_{n,k}|^{\frac{1}{k}} + |J_{n,k}|^{\frac{1}{k}} \leqslant 2, \quad (7)$$

where $I_{n,k}$ and $J_{n,k}$ are two quantities defined by $I_{n,k} = \frac{1}{2^k}\sum_{x_i, i\in\mathcal{I}} \langle a_{x_1}a_{x_2}\ldots a_{x_n}\rangle_{\overline{\mathcal{I}}}^0$ and $J_{n,k} = \frac{1}{2^k}\sum_{x_i, i\in\mathcal{I}} (-1)^{\sum_{j\in\mathcal{I}} x_j} \langle a_{x_1}a_{x_2}\ldots a_{x_n}\rangle_{\overline{\mathcal{I}}}^0$, $\mathcal{I} = \{i_1, i_2, \ldots, i_k\}$ denotes all the indexes of the independent agents A_{i_j}, $\overline{\mathcal{I}} = 1, 2, \ldots, n \setminus \mathcal{I}$, $\langle a_{x_1} a_{x_2}\ldots a_{x_n}\rangle_{\overline{\mathcal{I}}}^0 = \sum_a (-1)^{\sum_{i=1}^n a_i} P(a|x_{\mathcal{I}}; x_s = 0, s\in\overline{\mathcal{I}})$, and $\langle a_{x_1}a_{x_2}\ldots a_{x_n}\rangle_{\overline{\mathcal{I}}}^1 = \sum_a (-1)^{\sum_{i=1}^n a_i} P(a|x_{\mathcal{I}}; x_s = 1, s\in\overline{\mathcal{I}})$. Inequality (7) defines a star-convex set with the center point at the origin. This set contains the subset defined by $\mathcal{R}_c \leqslant 1$ in terms of the classical hidden-variable model and the subset defined by $\mathcal{R}_q \leqslant \sqrt{2}$ in terms of the quantum model, i.e.,

$$\{P|\mathcal{R}_c\} \subseteq \{P|\mathcal{R}_q\} \subseteq \{P|\mathcal{R}_{\rm ns}\}. \quad (8)$$

For cyclic networks without independent agents, we can prove the inequality

$$|I_{n,k}| + |J_{n,k}| \leqslant 2 \quad (9)$$

for all the nonsignaling correlations P. This inequality defines also a star-convex set of P. Although one cannot distinguish two sets generated by the classical causal model and quantum model using the inequality $|I_{n,k}| + |J_{n,k}| \leqslant 1$ for the cyclic networks, fortunately, in most cases we only need to consider those networks with independent agents regarding the following two facts: One is that lots of applications require an acyclic network with independent agents such as generalized entanglement swapping for building a large-scale entanglement. The other is that one can obtain a reduced network with independent agents from each cyclic network without independent agents by omitting redundant entangled states.

In what follows, consider a general problem of determining whether a given nonsignaling correlation P is compatible with an IONBDAG with multiple sources (acyclic networks). By contracting the multiple indexes Λ into a vector, Eq. (4) is rewritten as

$$P = R_\Lambda \mu, \quad (10)$$

where R_Λ is a contracted matrix while $\mu = (\mu_1, \mu_2, \ldots, \mu_m)^T$ is a contracted vector over the contracted vector Λ. The compatibility problem of a given correlation vector P and fixed Bayesian network with respect to the inputs $\{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n\}$ is equivalent to determining $\mu_1, \mu_2, \ldots, \mu_m$ which satisfy inequality (9) and Eq. (10). From inequality (7), it is equivalent to solving the optimization problem

$$\min_{\forall \mu_i \geqslant 0, \|\mu\|=1} X\mu, \quad \text{s.t.,} \quad R_\lambda \mu = P, \quad |I_{n,k}|^{\frac{1}{k}} + |J_{n,k}|^{\frac{1}{k}} \leqslant c,$$
$$P(a_i|x) = P(a_i|x_i), \quad |I_{n,k}|, |J_{n,k}| \leqslant 1, \quad (11)$$

where X is an objective matrix function and $P(a_i|x) = P(a_i|x_i)$ are nonsignaling conditions. c is an adjustable parameter satisfying $1 \leqslant c \leqslant 2$, as shown in Fig. 2. One can choose different values of c in the optimization for specific goals. It is especially useful for exploring different classes of correlations such as quantum correlations for $c \leqslant \sqrt{2}$ or nonsignaling correlations going beyond quantum mechanics for $\sqrt{2} < c \leqslant 2$. If the optimization problem given in Eq. (11) is feasible, P is compatible with $\{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n\}$. Otherwise, P is not included in the causal set derived from the IONBDAG with the inputs $\{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n\}$.

Note that $g(P) = |I_{n,k}|^{1/k} + |J_{n,k}|^{1/k}$ defines a multivariable star-convex function (without Lipschitz guarantees), which has a unique global minimum (and star center) at the origin. The standard gradient method and variants fail to make further progress because the search point oscillates around a different axis. Fortunately, there is a polynomial-time complexity algorithm that makes use of the ellipsoid method. Generally, it repeatedly refines an ellipsoidal region containing the star center to search for a global optimum [53]. They introduce a randomized cutting plane algorithm refining a feasible region of exponentially decreasing volume by iteratively removing cuts. With this algorithm, one can efficiently solve the compatibility problem shown in Eq. (11). Thus, the so-called verification algorithm is theoretically useful to search for new causal correlations and Bell inequality.

III. BELL CLASSES OF MULTIPARTITE CAUSAL NETWORKS

An NBDAG $\mathcal{G}_1$ is nonsignaling, implying another NBDAG $\mathcal{G}_2$ if all nonsignaling correlations are compatible with $\mathcal{G}_1$ and also compatible with $\mathcal{G}_2$. If all causal relaxations in $\mathcal{G}_1$ are presented in $\mathcal{G}_2$, $\mathcal{G}_1$ is nonsignaling, implying $\mathcal{G}_2$. In addition, if $\mathcal{G}_1$ and $\mathcal{G}_2$ are nonsignaling mutually, they are nonsignaling equivalent. Similarly to a single-source network, if two NBDAGs are nonsignaling equivalent, the multipartite correlations produced are useful for the same information-theoretic protocols [17].

In what follows, let an NBDAG whose causal relaxations consist of input-to-output locality relaxations be an input-output NBDAG (IONBDAG). Each IONBDAG is described by the subset of inputs that consists of the parents of each output, as shown in Fig. 3. IONBDAGs propose generic representatives of all the possible causal relaxations in the nonsignaling framework. As an application, the following theorem presents a full classification of the tripartite NBDAG derived from the entanglement swapping.

Theorem 1. Consider an NBDAG with three nodes and two hidden variables. There are 15 nonsignaling causal Bell classes that are shown in Fig. 3.

The proof of Theorem 1 is completed by examining all the possible NBDAGs. Figure 2 provides a simplified causal hierarchy of nonsignaling Bell correlations. There, the three red-shaded IONBDAGs of {(1), (2), (1, 2, 3)}, {(1), (2, 3), (1, 2, 3)}, and {(1, 3), (2, 3), (1, 2, 3)} are equivalent to each other and collapse to the star class {(1), (2), (1, 2, 3)}. The two red-shaded classes of {(1), (1, 2, 3), (3)} and {(1), (1, 2, 3), (2, 3)} are nonsignaling equivalent. The two purple-shaded classes are new causal relations in comparison to these DAGs with one variable [17]. The eight gray-shaded classes are known not to reproduce all quantum correlations [17]. Similar classifications are available for small-scale networks or special networks such as chain-shaped networks and star-shaped networks.

Lemma 1. Let $\mathcal{G}_{\rm gen}$ and $\mathcal{G}_{\rm io}$ be two NBDAGs whose difference is that for $1 \leqslant i \neq j \leqslant n$ such that $\{a_j, x_j\} \subseteq S_{x_i}$ and $\{a_j, x_j\} \subseteq S_{a_i}$ for $\mathcal{G}_{\rm gen}$, whereas $\{x_j\} \subseteq S_{a_i}$ for $\mathcal{G}_{\rm io}$. Then $\mathcal{G}_{\rm gen}$ and $\mathcal{G}_{\rm io}$ are nonsignaling equivalent.

One example is schematically shown in Fig. 4(a) and Fig. 4(b), consisting of three agents and two variables, λ_1 and λ_2. There, the nonsignaling correlations produced by the general locality relaxation from two agents to another on the left side coincide with those produced by another input-to-output locality relaxation on the right side.

Proof of Lemma 1. We prove Lemma 1 for the particular case of NBDAGs with three agents and two sources, λ_1 and λ_2. A similar proof holds for general cases. It is sufficient to prove the implication relations between the NBDAGs shown in Fig. 4(a). The most general relaxation of the tripartite locality is schematically represented by an NBDAG $\mathcal{G}_{\rm gen}$ on the left side of Fig. 4(a). A simple NBDAG $\mathcal{G}_{\rm io}$ is shown on the right side of Fig. 4(a). There, all the causal relaxations given in $\mathcal{G}_{\rm io}$ belong to the set consisting of all the causal relaxations shown in $\mathcal{G}_{\rm gen}$. It follows that the NBDAG $\mathcal{G}_{\rm io}$ implies $\mathcal{G}_{\rm gen}$ in terms of the nonsignaling conditions [4]. In what follows, we need to prove the converse implication.

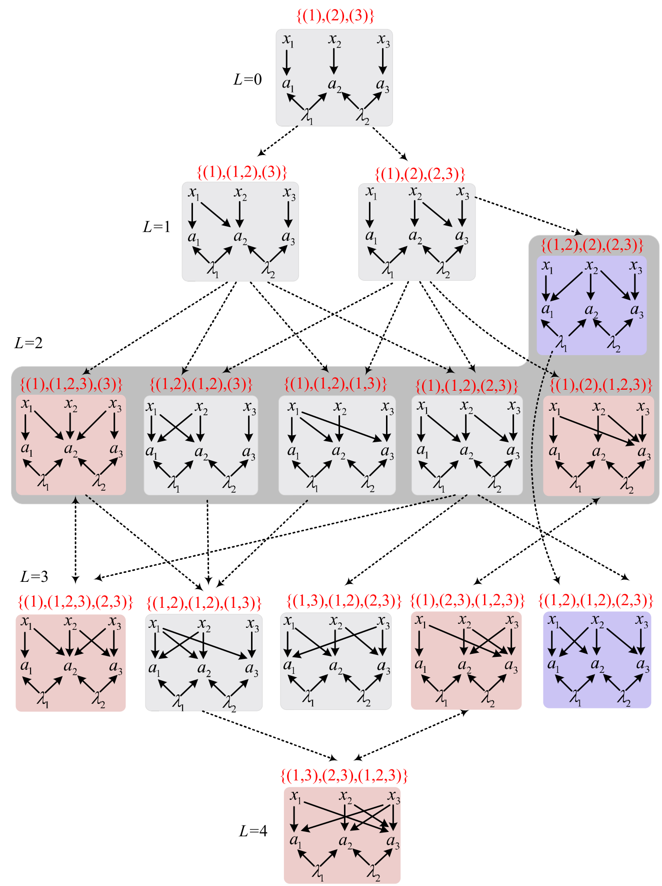

FIG. 3. Hierarchy (causal implications) of nonsignaling causal classes of tripartite Bell correlations. a_i and x_i are the respective input and output of one agent, $i = 1, 2, 3$. λ_1 and λ_2 are two independent hidden variables. Each class is represented by an IONBDAG that is labeled by a set $\{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3\}$, where each vector $\mathbf{I}_i$ consists of the parents of the output a_i. Each level of the hierarchy is defined by the total number L of the input-to-output locality relaxations. The dashed black arrow from one IONBDAG in one level to the followed level denotes that the latter nonsignaling implies the former. The bidirectional arrow represents the equivalent IONBDAGs.

Note that any joint probability distribution which is compatible with the NBDAG $\mathcal{G}_{\rm gen}$ can be rewritten as

$$
\begin{aligned}
P(a|x) &= \sum_{\lambda_1,\lambda_2} P(a, \lambda_1, \lambda_2|x) \\
&= \sum_{\lambda_1,\lambda_2} p(\lambda_1|x)p(\lambda_2|x)P(a|x, \lambda_1, \lambda_2) \qquad (12)
\end{aligned}
$$

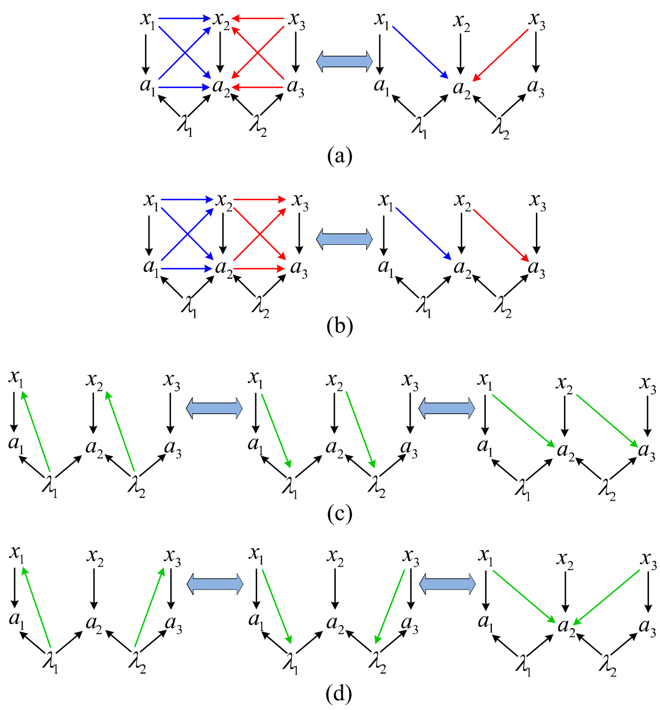

FIG. 4. Schematic of locality relaxation of a Bayesian network from the entanglement swapping. (a) Nonsignaling correlations produced by the general locality relaxation from two independent agents to another on the left side. a_i and x_i are the respective input and output of one agent, $i = 1, 2, 3$. λ_1 and λ_2 are two hidden variables. (b) Nonsignaling correlations produced by the general locality relaxation from two agents to another on the left side. (c) The relaxations of measurement independence on the left side and in the center produce the same set of nonsignaling correlations as the input-broadcasting model for two agents to two different agents on the right side. (d) The relaxations of measurement independence on the left side and in the center produce the same nonsignaling correlations.

$$= \sum_{\lambda_1,\lambda_2} p(\lambda_1|x)p(\lambda_2|x)p(a_1|x_1, x_2, \lambda_1) \times p(a_2|x, \lambda_1, \lambda_2)p(a_3|x_2, x_3, \lambda_1, \lambda_2) \quad (13)$$

$$= \sum_{\lambda_1,\lambda_2} p(\lambda_1|x)p(\lambda_2|x)p(a_1|x_1, \lambda_1) \times p(a_2|x, \lambda_1, \lambda_2)p(a_3|x_3, \lambda_1, \lambda_2) \quad (14)$$

$$= \sum_{\lambda_1,\lambda_2} p(\lambda_1)p(\lambda_2)p(a_1|x_1, \lambda_1) \times p(a_2|x, \lambda_1, \lambda_2)p(a_3|x_3, \lambda_1, \lambda_2). \quad (15)$$

This is an explicit expression of generic correlations produced by Bayesian networks with respect to the NBDAG $\mathcal{G}_{\rm io}$ shown in Fig. 4(a), where $a = a_1a_2a_3$, $x = x_1x_2x_3$. Equation (12) follows from the independence of two sources, λ_1 and λ_2. Equation (13) follows from the nonsignaling conditions: $p(a_1|x_1, x_2, x_3, \lambda_1, \lambda_2) = p(a_1|x_1, x_2, \lambda_1)$ and $p(a_3|x_1, x_2, x_3, \lambda_1, \lambda_2) = p(a_3|x_2, x_3, \lambda_2)$. To obtain Eq. (14), a new variable, $\hat{\lambda}_1$ (with more outputs), is defined for representing two variables, λ_1 and x_2, where x_2 is deterministic. Similarly, one can define a variable $\hat{\lambda}_2$ for representing two variables, λ_2 and x_2. Equation (15) follows from the independence of variables.

Note that in Eq. (13) if a new variable, $\hat{\lambda}_2$, is used to represent two conditional variables, λ_2 and x_3, for the variable a_2, i.e., $p(a_2|x, \lambda_1, \lambda_2) = p(a_2|x_1, x_2, \lambda_2)$, one can prove another structure shown on the right side of Fig. 4(a). A similar result holds for the NBDAG shown in Fig. 4(b). Consequently, the proof is complete. □

Lemma 2. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ and $\mathcal{G}_b$ be three NBDAGs whose differences are $\lambda_j \in S_{x_i}$ for $\mathcal{G}_1$, $x_i \in S_{\lambda_j}$ for $\mathcal{G}_2$, and $x_i \in S_{a_k}$, with all $k \in \mathcal{I}_j$ for $\mathcal{G}_b$ and $1 \leqslant i \leqslant n$, where $\mathcal{I}_j$ satisfies that $\{x_s, s \in \mathcal{I}_j\} \subseteq S_{a_k}$. Then $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_b$ are nonsignaling equivalent.

The proof of Lemma 2 is forward and easily completed. Two examples are shown in Fig. 4(c) and Fig. 4(d). Here, the relaxations of measurement independence both on the left side and in the center produce the same set of nonsignaling correlations as the input-broadcasting model for two agents to two different agents on the right side. The proof can be completed by considering the subnetworks consisting of one hidden variable [for example, subnetwork $\{x_1, x_2, a_1, a_2, \lambda_1\}$ or $\{x_2, x_3, a_2, a_3, \lambda_2\}$ of the NBDAG on the left side of Fig. 4(c)] using the recent measurement-independence relaxation [17].

Remarkably, Lemmas 1 and 2 imply that every causal relaxation on a GLHV model is accounted for by an input-to-output locality relaxation when any nonsignaling correlations are concerned. Thus, Lemmas 1 and 2 are useful for reducing the total number of examined NBDAGs. For example, all the NBDAGs of 15 different ways of connecting directed edges from one agent to another are collectively grouped into a single IONBDAG due to Lemma 1, where there are 15 instances of general locality relaxations similar to those shown in Fig. 4(a). All NBDAGs with directed edges from hidden variables to any of the inputs are further grouped together into an IONBDAG due to Lemma 2.

Proof of Theorem 1. Inspired by the method in [17], to prove Theorem 1 we first present the causal hierarchy of a network with $n = 3$ and $k = 2$ shown in Fig. 5 from Lemmas 1 and 2, where the symmetry of two agents, A_1 and A_3, has been used to reduce causal classes.

A. Case 1. Nonsignaling borning causal classes

In Fig. 4, we first prove that the orange boxes are nonsignaling borning. It can be completed by proving that $\{(1), (1, 2), (1, 2, 3)\}$ and $\{(1), (1, 2, 3), (1, 3)\}$ in the third level are nonsignaling borning. Consider the arbitrary tripartite correlation $P(a|x)$ with $a = a_1a_2a_3$, $x = x_1x_2x_3$. It can be decomposed as

$$P(a|x) = p(a_2|x, a_1, a_3)p(a_3|x, a_1)p(a_1|x) \quad (16)$$

$$= p(a_2|x, a_1, a_3)p(a_3|x_1, x_2, a_1)p(a_1|x_1), \quad (17)$$

where Eq. (16) follows from Bayes' rule, and Eq. (17) follows from the nonsignaling constraints given by Eq. (3). Now, from Eq. (2), $p(a_2|x_1, x_2, x_3, a_1, a_3)p(a_3|x_1, x_2, a_1)p(a_1|x_1)$ is a special correlation produced by a generalized Bayesian network with respect to an NBDAG with locality relaxations from agent A_1 to agent A_2 or A_3 and from A_3 to A_2. From Lemma 1, these correlations are always within the causal Bell

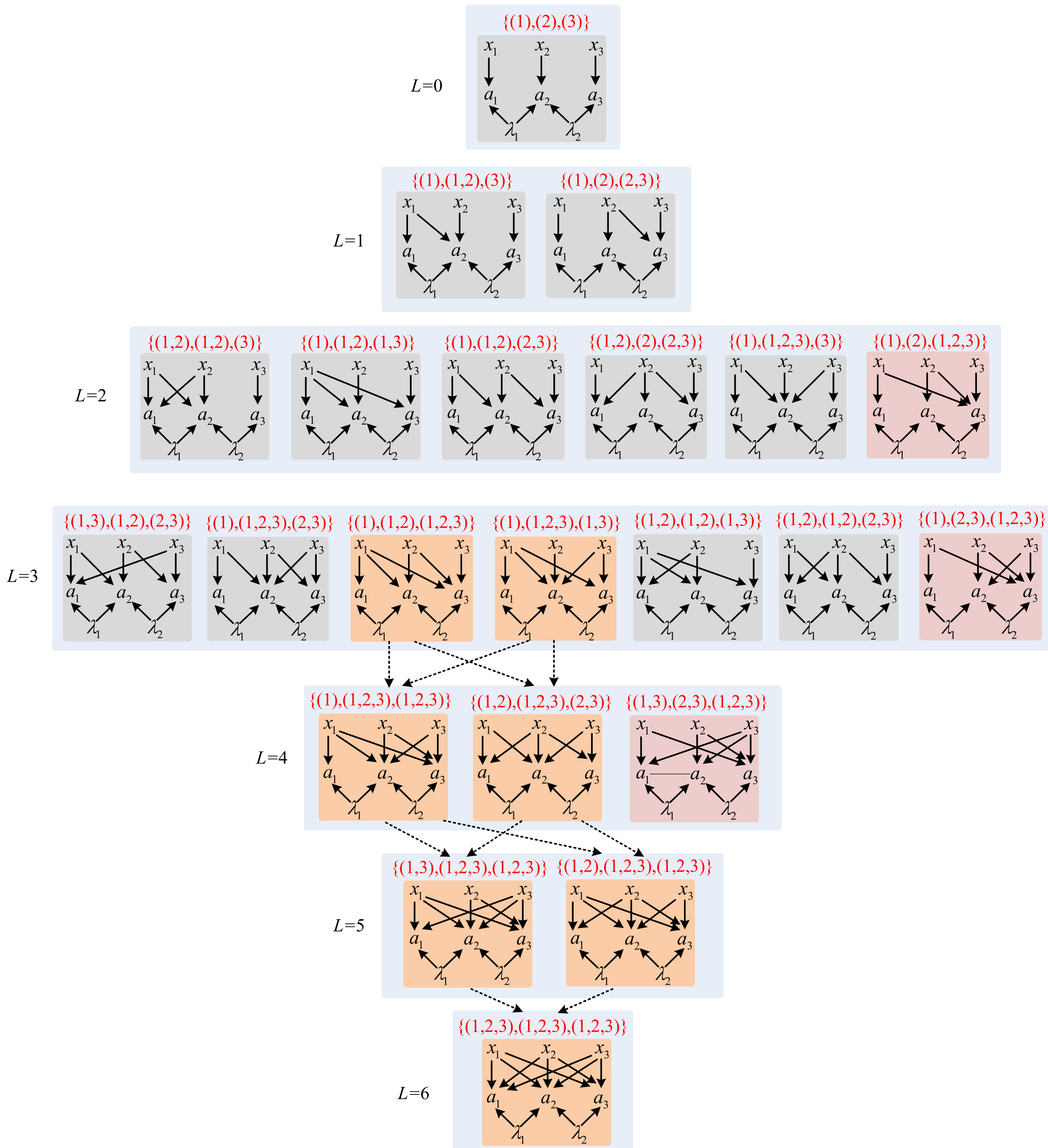

FIG. 5. Causal hierarchy of nonsignaling classes of Bell correlations from a network with $n = 3$ and $k = 2$. Each class is represented by an IONBDAG and labeled by a set $\{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3\}$, where each vector $\mathbf{I}_i$ is composed of the parent inputs of a_i. Each level of the hierarchy has the total number L of input-to-output locality relaxations. The dashed black arrow from one IONBDAG in one level to the followed level represents that the latter nonsignaling implies the former.

class $\{(1), (1, 2, 3), (1, 3)\}$. A similar result can be proved for $\{(1), (1, 2), (1, 2, 3)\}$.

B. Case 2. Two equivalences of causal classes

Now, we prove the equivalence of three red boxes, i.e., $\{(1), (2), (1, 2, 3)\} \longleftrightarrow \{(1), (2, 3), (1, 2, 3)\} \longleftrightarrow \{(1, 3), (2, 3), (1, 2, 3)\}$. Similar results hold for $\{(1), (1, 2, 3), (3)\} \longleftrightarrow \{(1), (1, 2, 3), (2, 3)\}$. It is sufficient to prove the implication relationships $\{(1), (2), (1, 2, 3)\} \leftarrow \{(1), (2, 3), (1, 2, 3)\} \leftarrow \{(1, 3), (2, 3), (1, 2, 3)\}$.

We first prove that these Bayesian networks, $\{(1), (2), (1, 2, 3)\}$, $\{(1), (2, 3), (1, 2, 3)\}$, and $\{(1, 3), (2, 3), (1, 2, 3)\}$, generate the same marginal correlations $P(a_1, a_2|x_1, x_2)$. Consider the arbitrary correlation $P(a|x)$ produced by a generalized tripartite Bayesian network with respect to the inputs $\{(1, 3), (2, 3), (1, 2, 3)\}$. Then the marginal correlations of agents A_1 and A_2 have the following components:

$$
\begin{aligned}
&P(a_1, a_2|x_1, x_2)\\
&\quad := \sum_{\substack{a_3, x_3,\\ \lambda_1, \lambda_2}} P(a, x_3|x_1, x_2, \lambda_1, \lambda_2)p(\lambda_1)p(\lambda_2)\\
&\qquad \times p(a_3|x, \lambda_1, \lambda_2)p(x_3|x_1, x_2, \lambda_1, \lambda_2)p(\lambda_1)p(\lambda_2) \quad (18)
\end{aligned}
$$

$$= \sum_{\substack{a_3,x_3,\\ \lambda_1,\lambda_2}} p(a_1|x_1, x_3, \lambda_1)p(a_2|x_2, x_3, \lambda_1, \lambda_2) \times p(a_3|x, \lambda_2)p(x_3|\lambda_1, \lambda_2)p(\lambda_1)p(\lambda_2)$$

$$= \sum_{x_3,\lambda_1,\lambda_2} p(a_1|x_1, x_3, \lambda_1)p(a_2|x_2, x_3, \lambda_1, \lambda_2) \times P(x_3, \lambda_1, \lambda_2) \quad (19)$$

$$= \sum_{\lambda_1',\lambda_2} p(a_1|x_1, \lambda_1')p(a_2|x_2, \lambda_1', \lambda_2)p(\lambda_1')p(\lambda_2) \quad (20)$$

$$= \sum_{\lambda_1'} p(a_1|x_1, \lambda_1')p(a_2|x_2, \lambda_1')p(\lambda_1'). \quad (21)$$

Here, Bayes' rule has been used in Eq. (18). Equation (2) has been used to get Eq. (19). Equation (20) follows from the normalization equality $\sum_{a_3} p(a_3|x, \lambda_2) = 1$ and Bayes' rule $p(x_3|\lambda_1, \lambda_2)p(\lambda_1)p(\lambda_2) = p(x_3, \lambda_1, \lambda_2)$. Equation (21) follows from a redefined variable, $\lambda_1' := (x_3, \lambda_1)$. Equation (21) is obtained from the normalization equality $\sum_{\lambda_2} p(\lambda_2)p(a_2|x_2, \lambda_1', \lambda_2) = p(a_2|x_2, \lambda_1')$. Equation (21) defines bipartite correlations of a GLHV model for two agents, A_1 and A_2. These correlations are the same as those generated from Bayesian networks, $\{(1), (2), (1, 2, 3)\}$ and $\{(1), (2, 3), (1, 2, 3)\}$.

In what follows, we prove that three Bayesian networks, $\{(1), (2), (1, 2, 3)\}$, $\{(1), (2, 3), (1, 2, 3)\}$, and $\{(1, 3), (2, 3), (1, 2, 3)\}$, can generate the same nonsignaling correlation. Consider the arbitrary nonsignaling correlation $P(a|x)$. Then it holds that

$$P(a|x) = \sum_{\lambda_1,\lambda_2} P(a|x, \lambda_1, \lambda_2)p(\lambda_1)p(\lambda_2)$$

$$= \sum_{\lambda_1,\lambda_2} p(a_3|a_1, a_2, x, \lambda_2)P(a_1, a_2|x, \lambda_1, \lambda_2) \cdot p(\lambda_1)p(\lambda_2) \quad (22)$$

$$= \sum_{\lambda_2} p(a_3|a_1, a_2, x, \lambda_2)P(a_1, a_2|x_1, x_2, \lambda_2)p(\lambda_2) \quad (23)$$

for any a_i, x_j. Here, Eq. (22) follows from Bayes' rule. Equation (23) is from the nonsignaling condition and the normalization equality $\sum_{\lambda_1} p(\lambda_1)P(a_1, a_2|x, \lambda_1, \lambda_2) = P(a_1, a_2|x, \lambda_2)$.

Assume that P is produced by a Bayesian network with respect to one of three NBDAGs: $\{(1), (2), (1, 2, 3)\}$, $\{(1), (2, 3), (1, 2, 3)\}$, and $\{(1, 3), (2, 3), (1, 2, 3)\}$. From Eq. (21) the marginal distribution $P(a_1, a_2|x_1, x_2, \lambda_2) = P(a_1, a_2|x_1, x_2)/p(\lambda_2)$ in Eq. (23) for two agents, A_1 and A_2, defines the same correlation for three NBDAGs. Moreover, the marginal distribution $p(a_3|a_1, a_2, x, \lambda_2)$ given in Eq. (23) spans the same set of the conditional probability distributions given by a_1, a_2, x, and λ_2. The reason is that for each NBDAG A_3 knows the other's inputs and the variable λ_2. Thus, A_3 can reproduce all the conditional distributions of $p(a_3|a_1, a_2, x, \lambda_2)$. Hence, from Eq. (23), the arrows from A_3 to A_1 or A_1 do not generate new nonsignaling correlations. This means that the three Bayesian networks, $\{(1), (2), (1, 2, 3)\}$, $\{(1), (2, 3), (1, 2, 3)\}$, and $\{(1, 3), (2, 3), (1, 2, 3)\}$, generate the same nonsignaling correlations.

C. Case 3. Other causal classes

We consider eight gray-shaded classes. All the causal classes denoted by $\{(1), (2), (3)\}$, $\{(1), (1, 2), (3)\}$, $\{(1), (2), (2, 3)\}$, $\{(1, 2), (1, 2), (3)\}$, $\{(1), (1, 2), (1, 3)\}$, $\{(1), (1, 2), (2, 3)\}$, and $\{(1, 2), (1, 2), (1, 3)\}$ are partially paired correlations [51], which satisfy the following Svetlichny inequality [11]:

$$-\langle A_0B_0C_0\rangle + \langle A_0B_0C_1\rangle + \langle A_0B_1C_0\rangle + \langle A_0B_1C_1\rangle + \langle A_1B_0C_0\rangle + \langle A_1B_0C_1\rangle + \langle A_1B_1C_0\rangle - \langle A_1B_1C_1\rangle \leqslant 4. \quad (24)$$

This inequality is violated by quantum correlations obtained from local measurements on an entangled quantum state [11]. Unfortunately, it may be useless to verify distributed entangled states.

For the class represented by $\{(1, 3), (1, 2), (2, 3)\}$, it has been proved that all the compatible correlations satisfy the following inequality [17]:

$$\langle A_0B_0C_0\rangle + \langle A_0B_0C_1\rangle + \langle A_0B_1C_0\rangle + \langle A_0B_1C_1\rangle + \langle A_1B_0C_0\rangle + \langle A_1B_0C_1\rangle + \langle A_1B_1C_0\rangle - \langle A_1B_1C_1\rangle \leqslant 6. \quad (25)$$

This inequality is violated up to the algebraic maximal value 8 by the nonsignaling correlations as

$$P(a_1, a_2, a_3|x_1, x_2, x_3) = \tfrac{1}{4}\delta_{a_1\oplus a_2\oplus a_3, x_1\times(x_2\oplus x_3)}, \quad (26)$$

which is originally identified in Ref. [52]. This proves that $\{(1, 3), (1, 2), (2, 3)\}$ is nonsignaling.

Generally, for DAGs with multiple sources, it is difficult to classify all the external nonsignaling correlations using the standard convex-optimization tools [53]. Actually, these nonsignaling correlations consist of a star-convex set such as those proved in Sec. II. Thus, the star-convex optimization [54] is useful for exploring new nonsignaling causal classes for a specific network. □

IV. DEVICE-INDEPENDENT INFORMATION PROCESSING IN GENERAL QUANTUM NETWORKS

Consider a general network $\mathcal{N}$ consisting of n agents, $\mathsf{A}_1, \ldots, \mathsf{A}_n$, who share m independent hidden sources. $\mathcal{N}$ is k independent if there are k agents without prior-shared sources. Each agent A_i performs local measurements with dichotomic input, denoted $x_i \in \{0, 1\}$, and obtains dichotomic output, denoted $a_i \in \{-1, 1\}$. Similar results hold for multiple inputs and outputs using linear superposition of different inputs and outputs. The schematic causal relations are shown in an NBDAG in Fig. 6(a). If all sources λ_i are equivalent random variables, the classically achievable n-partite correlations satisfy the nonlinear inequality [29]

$$\mathcal{R}_k := |I_{n,k}|^{\frac{1}{k}} + |J_{n,k}|^{\frac{1}{k}} \leqslant 1, \quad (27)$$

where $I_{n,k}$ and $J_{n,k}$ are linear superpositions of correlations [29]. Similarly to the standard device-independent informa-

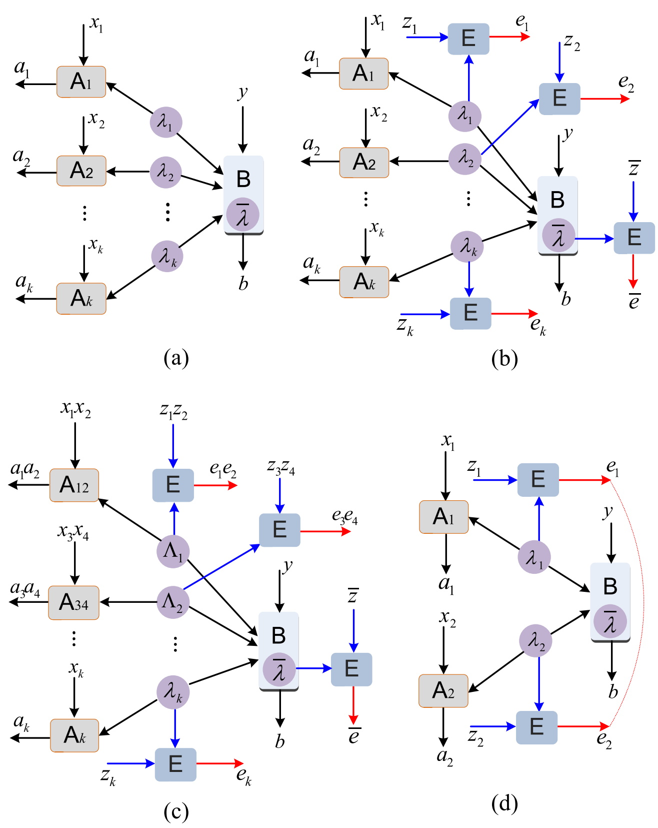

FIG. 6. (a) A k-independent network in terms of the GLHV model. A_i and B share some independent sources $\Lambda_i = \lambda_{j_1} \ldots \lambda_{j_{\ell_i}}$, where Λ_i's satisfy $\cup_{i=1}^n \Lambda_i = \lambda_1 \ldots \lambda_m$. Assume $\Lambda_i = \lambda_i$ for simplicity. $\overline{\lambda} = \lambda_{k+1} \ldots \lambda_m$. B includes all the other agents, $\mathsf{A}_{k+1}, \ldots, \mathsf{A}_n$. (b) An IONBDAG in terms of the device-independent model. An eavesdropper holds some systems, each of which is correlated with one source, λ_i. z_i and e_i denote the respective input and outcome of the measurement on each eavesdropper's system λ_i. $\overline{e} = e_{k+1} \ldots e_n$. (c) An eavesdropper holds some systems, some of which are correlated with multiple sources in Λ_i, which may be correlated into a new variable $\hat{\lambda}_i$. Take $\Lambda_1 = \{\lambda_1, \lambda_2\}$ as an example. The eavesdropper can correlate λ_1 and λ_2 into a new variable, $\hat{\lambda}_1$. (d) An eavesdropper holds a system that correlates with λ_1 and λ_2 (represented by red lines).

tion processing in Bell networks, the adversary is limited to recovering privacy information of legal agents. Assume herein that the eavesdropper holds m independent systems, each of which is correlated with one of m sources, as shown in Fig. 6(b). The eavesdropper's systems can be postquantum (nonsignaling). The output e_i of each eavesdropper's system depends on its input z_i and correlated sources. To complete a general network task, it is reasonable to permit independent agents A_i with $i \in \mathcal{I}$ to classically communicate with each other. Informally, a violation of inequality (27) provides an upper bound of an eavesdropper's information about the outcomes of independent agents. Denote the variational distance of two probability distributions, $\{p(x)\}$ and $\{q(x)\}$, as $D(p, q) = \frac{1}{2}\sum_x |p(x) - q(x)|$.

Theorem 2. The total information about independent agents' outcome recovered by an eavesdropper satisfies the inequality

$$D(P(e|a_i, i \in \mathcal{I}; x, z), \prod_{i=1}^{m} p(e_i|z_i)) \leqslant k(2 - \mathcal{R}_k) \tag{28}$$

if all the variables λ_i with $i \in \mathcal{I}$ are independent, where $e = e_1 \ldots e_m$, $x = x_1 \ldots x_n$, and $z = z_1 \ldots z_m$.

Proof. Consider a general network $\mathcal{N}$ consisting of n nodes (or agents), $\mathsf{A}_1, \mathsf{A}_2, \ldots, \mathsf{A}_n$, who share m independent hidden sources. $\mathcal{N}$ is k independent if there are k spacelike separated agents without prior-shared hidden sources. Each agent A_i performs local measurements with dichotomic input, denoted $x_i \in \{0, 1\}$, and then obtains dichotomic output, denoted $a_i \in \{-1, 1\}$. Similar results hold for multiple inputs and outputs using linear superposition of different inputs and outputs. The schematic causal relations of all the agents' inputs and outputs are shown in an NBDAG in Fig. 3(a). If all the sources λ_i are equivalent random variables, the classically achievable n-partite correlations satisfy the nonlinear inequality [29]

$$\mathcal{R}_k := |I_{n,k}|^{\frac{1}{k}} + |J_{n,k}|^{\frac{1}{k}} \leqslant 1, \tag{29}$$

where

$$I_{n,k} = \frac{1}{2^k} \sum_{x_i, i \in \mathcal{I}} \langle a_{x_1} a_{x_2} \ldots a_{x_n} \rangle_{\overline{\mathcal{I}}}^{0}, \tag{30}$$

$$J_{n,k} = \frac{1}{2^k} \sum_{x_i, i \in \mathcal{I}} (-1)^{\sum_{j \in \mathcal{I}} x_j} \langle a_{x_1} a_{x_2} \ldots a_{x_n} \rangle_{\overline{\mathcal{I}}}^{1}, \tag{31}$$

which are defined in Eq. (7), $\mathcal{I} = \{i_1, i_2, \ldots, i_k\}$ denotes the indexes of independent agents A_{i_j}, $\overline{\mathcal{I}} = \{1, 2, \ldots, n\} \setminus \mathcal{I}$ denotes the complement set of $\mathcal{I}$, $\langle a_{x_1} a_{x_2} \ldots a_{x_n} \rangle_{\overline{\mathcal{I}}}^{0} = \sum_{a_1, \ldots, a_n} (-1)^{\sum_{i=1}^n a_i} P(a|x_{\mathcal{I}}; x_s = 0, s \in \overline{\mathcal{I}})$, and $\langle a_{x_1} a_{x_2} \ldots a_{x_n} \rangle_{\overline{\mathcal{I}}}^{1} = \sum_{a_1, \ldots, a_n} (-1)^{\sum_{i=1}^n a_i} P(a|x_{\mathcal{I}}; x_s = 1, s \in \overline{\mathcal{I}})$.

Now, consider a quantum realization of $\mathcal{N}$, where $\mathcal{N}$ consists of generalized EPR states [2] or GHZ states [42]. Each observer A_i performs local two-valued positive-operator-valued-measurements defined by positive semidefinite operators. We proved that the expectation of quantum correlations among spacelike separated observers satisfies the Cirel'son bound [29]

$$1 < |I_{n,k}^q|^{\frac{1}{k}} + |J_{n,k}^q|^{\frac{1}{k}} \leqslant \sqrt{2}, \tag{32}$$

which violates inequality (29), where $I_{n,k}^q$ and $J_{n,k}^q$ are the corresponding quantities of $I_{n,k}$ and $J_{n,k}$ constructed by quantum correlations. This nonlinear Bell-type inequality is useful for verifying multisource quantum networks [29].

Consider the conditional distribution $P(a, b, e|x, y, z)$, where a and b are binary random variables and x and y are s-valued ($s \geqslant 2$), satisfying the nonsignaling conditions

$$\begin{aligned} P(a, b|x, y, z) &= P(a, b|x, y), \\ P(a, e|x, y, z) &= P(a, e|x, z), \\ P(b, e|x, y, z) &= P(b, e|y, z). \end{aligned} \tag{33}$$

It easily implies new nonsignaling conditions:

$$p(a|b, x, y, z) = p(a|b, x, y), \quad p(b|c, x, y, z) = p(b|c, y, z),$$
$$p(c|a, x, y, z) = p(c|b, x, z). \tag{34}$$

We only prove $p(a|b,x,y,z)=p(a|b,x,y)$ as an example, which is obtained from the equalities

$$\begin{aligned} p(b|x,y,z)p(a|b,x,y,z) &= P(a,b|x,y,z) \\ &= P(a,b|x,y) \\ &= p(b|x,y)p(a|b,x,y) \end{aligned} \quad (35)$$

and $p(b|x,y,z)=p(b|x,y)$ (nonsignaling condition).

From Fig. 6(b), we get the conditional independence relations

$$\begin{aligned} P(e|a_i, i\in\mathcal{I};x,z) = &\prod_{i=1}^{k} p(e_i|x_{\lambda_i}, a_{\lambda_i}, z_i) \\ &\times \prod_{j=k+1}^{m} p(e_j|x_{\lambda_j}, a_{\lambda_j}, z_j), \end{aligned} \quad (36)$$

where we have assumed for convenience that the sources $\lambda_1,\lambda_2,\dots,\lambda_k$ are shared by all independent agents of $\mathsf{A}_1,\mathsf{A}_2,\dots,\mathsf{A}_k$. All the other sources $\lambda_{k+1},\lambda_{k+2},\dots,\lambda_m$ are shared by the agents included in B as shown in Fig. 6(b). x_{λ_i} and a_{λ_i} denote the respective inputs and outputs of all the agents who have shared the source λ_i, $i=1,2,\dots,m$. A similar proof holds for other cases by combining the shared sources into a new one for each agent.

Note that all the agents included in B are not permitted for classical communications. From the nonsignaling conditions shown in Eqs. (33) and (34), we can rewrite Eq. (36) as

$$P(e|a_i, i\in\mathcal{I};x,z) = \prod_{i=1}^{k} p(e_i|x_{\lambda_i}, a_{\lambda_i}, z_i) \prod_{j=k+1}^{m} p(e_j|z_j). \quad (37)$$

Consider the left side of inequality (28). From Eq. (37) it can be decomposed as follows:

$$\begin{aligned} &D(P(e|a_i, i\in\mathcal{I};x,z), \prod_{i=1}^{m} p(e_i|z_i)) \\ &= D(\prod_{i=1}^{k} p(e_i|x_{\lambda_i}, a_{\lambda_i}, z_i) \prod_{j=k}^{m} p(e_j|z_j), \prod_{i=1}^{m} p(e_i|z_i)) \\ &= \sum_{e_{k+1},\dots,e_m} D(\prod_{i=1}^{k} p(e_i|x_{\lambda_i}, a_{\lambda_i}, z_i), \prod_{i=1}^{k} p(e_i|z_i)) \\ &\quad \times \prod_{j=k+1}^{m} p(e_j|z_j) \\ &= D(\prod_{i=1}^{k} p(e_i|x_{\lambda_i}, a_{\lambda_i}, z_i), \prod_{i=1}^{k} p(e_i|z_i)) \end{aligned} \quad (38)$$

$$\begin{aligned} &\leqslant \frac{1}{2}\sum_{e_1,\dots,e_k} p(e_1|x_{\lambda_1}, a_{\lambda_1}, z_1)|\prod_{i=2}^{k} p(e_i|x_{\lambda_i}, a_{\lambda_i}, z_i) \\ &\quad -\prod_{i=2}^{k} p(e_i|z_i)| + \frac{1}{2}\sum_{e_1,\dots,e_k} |p(e_1|x_{\lambda_1}, a_{\lambda_1}, z_1) \\ &\quad - p(e_1|z_1)|\prod_{i=2}^{k} p(e_i|z_i) \end{aligned} \quad (39)$$

$$\begin{aligned} &\leqslant \frac{1}{2}\sum_{e_1,\dots,e_k} |\prod_{j=2}^{k} p(e_j|x_{\lambda_j}, a_{\lambda_j}, z_j) - \prod_{j=2}^{k} p(e_j|z_j)| \\ &\quad + D(p(e_1|x_{\lambda_1}, a_{\lambda_1}, z_1), p(e_1|z_1)) \end{aligned} \quad (40)$$

$$\leqslant \sum_{i=1}^{k} D(p(e_i|x_{\lambda_i}, a_{\lambda_i}, z_i), p(e_i|z_i)) \quad (41)$$

$$\leqslant \sum_{i=1}^{k} I_2(P(a_i, b_i|x_i, y_i)). \quad (42)$$

In Eq. (38), we have used the normalization conditions $\sum_{e_j} p(e_j|z_j)=1$ for $j=k+1, k+2,\dots,m$. Here, inequality (39) follows from the triangle inequality $|x-y|\leqslant|x-z|+|z-y|$. Inequality (40) follows from the normalization conditions $\sum_{e_1} p(e_1|x_{\lambda_1},a_{\lambda_1},z_1)=1$ and $\sum_{e_j} p(e_j|z_j)=1$ for $j=2,3,\dots,k$. In inequality (41), we have iteratively used inequality (40) for $\frac{1}{2}\sum_{e_2,\dots,e_k}|\prod_{j=2}^{k} p(e_j|x_{\lambda_j},a_{\lambda_j},z_j)-\prod_{j=2}^{k} p(e_j|z_j)|$. In inequality (42), I_2 is from the chained Bell inequality [8] with two measurement settings, defined as

$$\begin{aligned} I_2(P(a_i,b_i|x_i,y_i)) := &P(a_i=b_i|x_i=1, y_i=2) \\ &+ \sum_{|x_i-y_i|=1} P(a_i\neq b_i|x_i,y_i), \end{aligned} \quad (43)$$

where x_i and y_i denote the respective input of agents A_i and the related agent B_i included in B, and a_i and b_i denote the respective output of agent A_i and the related agent B_i included in B. Inequality (42) follows from the inequality $D(p(e|a,x,z),p(e|z))\leqslant I_k(P(a,b|x,y))$ [55–57] and the general form

$$D(p(e|x,a,z),p(e|z))\leqslant I_k(P(a_i,a_j|x_i,x_j)), \quad (44)$$

with $I_k(P(a,b|x,y))=P(a=b|x=1;y=k)+\sum_{|x-y|=1}P(a\neq b|x;y)$, where all the agents A_i and the potential eavesdropper are correlated by one source. Inequality (44) can be proved by following the same procedure [56] and the fact that $P(a|x,z)$ is a conditional probability distribution for given inputs x and z.

Now, consider a quantum network $\mathcal{N}$ in which all the agents have binary inputs and outputs (a similar result holds for the multiple inputs and outputs [45]). Especially, as proved in Ref. [29], all the independent observers A_i perform separable measurements $A_i^{x_i}=A_{i,0}^{x_i}\otimes A_{i,1}^{x_i}$, while all the other agents B_j included in B perform separable measurements $B_j^{y_j}=B_{j,0}^{y_j}\otimes B_{j,1}^{y_j}$ on local systems. We can get

$$\langle A_1^{x_1}A_2^{x_2}\dots A_k^{x_k}B^{y}\rangle=\prod_{i=1}^{k}\langle A_i^{x_i}B_i^{y_i}\rangle. \quad (45)$$

From the definitions of $I_{n,k}$ and $J_{n,k}$ shown in Eqs. (30) and (31), respectively, it follows that

$$\begin{aligned} I_{n,k} &= \frac{1}{2^k}\prod_{i=1}^{k}\left(\langle A_j^0B_j^0\rangle+\langle A_j^1B_j^0\rangle\right), \\ J_{n,k} &= \frac{1}{2^k}\prod_{i=1}^{k}\left(\langle A_j^0B_j^1\rangle-\langle A_j^1B_j^1\rangle\right). \end{aligned} \quad (46)$$

From Eq. (46) and the arithmetic-geometric inequality $(\prod_{i=1}^n x_i)^{1/n} \leqslant \frac{1}{n}\sum_{i=1}^n x_i$, we get

$$\begin{aligned}\mathcal{R}_k &= |I_{n,k}|^{\frac{1}{k}} + |J_{n,k}|^{\frac{1}{k}} \\ &\leqslant \frac{1}{2}\sum_{i=1}^{k}\left(\left|\langle A_j^0B_j^0\rangle + \langle A_j^1B_j^0\rangle\right| + \left|\langle A_j^0B_j^1\rangle - \langle A_j^1B_j^1\rangle\right|\right) \\ &:= \frac{1}{2k}\sum_{i=1}^{k} C_2^{A_iB_i},\end{aligned} \tag{47}$$

where $C_2^{A_iB_i} := |\langle A_j^0B_j^0\rangle + \langle A_j^1B_j^0\rangle| + |\langle A_j^0B_j^1\rangle - \langle A_j^1B_j^1\rangle|$ is a special quantity used in the CHSH inequality [58].

By using $\langle AB\rangle = 2p(A = B) - 1$, one can prove [17] that

$$I_2(e_i) = 2 - \tfrac{1}{2}C_2^{A_iB_i}. \tag{48}$$

Combining Eqs. (47) and (48), the right side of inequality (42) is evaluated as

$$\begin{aligned}\sum_{i=1}^{k} I_2(e_i) &= \sum_{i=1}^{k}\left(2 - \frac{1}{2}C_2^{A_iB_i}\right) \\ &\leqslant k(2 - \mathcal{R}_k),\end{aligned} \tag{49}$$

which completes the proof. □

If the eavesdropper can correlate sources, λ_i's, inequality (28) will then be extended from a similar proof. An example is shown in Fig. 6(c). Here, one first combines the correlated sources λ_1 and λ_2 (λ_3 and λ_4) into a new one, $\hat{\lambda}_1$ ($\hat{\lambda}_2$). Define $\hat{e} = (e_1e_2, e_3e_4, e_5, \ldots, e_m)$. A similar result holds by replacing the left side of inequality (28) with $D(P(\hat{e}|a_i, i \in \mathcal{I}; x, z), p(e_1e_2|z_1z_2)p(e_3e_4|z_3z_4)\prod_{i=3}^m p(e_i|z_i))$. For example, assume that the eavesdropper holds two uncorrelated sources, λ_1 and λ_2, after readjusting the network. It is then sufficient to use a new nonlinear inequality, $\mathcal{R}_2 = \sqrt{|I_{n,2}|} + \sqrt{|J_{n,2}|} \leqslant 1$, by considering two independent agents who own the respective sources λ_1 and λ_2 [29], where $I_{n,2}$ and $J_{n,2}$ are new quantities with respect to two independent agents [29]. Hence, a new inequality, $D(P(e_1, e_2|a_i, i \in \mathcal{I}; x, z_1, z_2), p(e_1|z_1)p(e_2|z_2)) \leqslant 2(2 - \mathcal{R}_2)$, follows for the eavesdropping information from a proof similar to that above.

Note that for a network consisting of white noisy sources of EPR states or GHZ states [29], the visibility from inequality (27) is still unchanged in comparison to those networks with a single entangled source in terms of the CHSH inequality [57]. So, similarly to the standard Bell network, noisy sources cannot strengthen the security in a general network in terms of the nonlinear inequality (27). Hence, all agents can make use of some strategies such as nonseparable measurements or different forms of inequality (27) to against leaking information. Nevertheless, the result fails to feature the strongest eavesdropper who can correlate all sources, as shown in Fig. 6(d), which are reduced to a single-source network [34–36].

V. SOME EXAMPLES OF DEVICE-INDEPENDENT INFORMATION PROCESSING

A. Chain-shaped networks

A long-distance chain-shaped network is schematically shown in Fig. 7(a). We have shown that multipartite quantum correlations of long-distance entanglement distributions violate inequality (27) for all the bipartite entangled pure states as resources [29], where $k = \lceil n/2 \rceil$ denotes the number of independent observers and $\lceil x \rceil$ denotes the smallest integer no less than x. The maximal violation is achieved for EPR states. From Theorem 2, if an eavesdropper holds n independent systems each of them is correlated with one of n sources $\lambda_1, \lambda_2, \ldots, \lambda_n$. Each system can be measured by a device with input z_i and output e_i. In the experiment, each agent A_i first outputs a_i depending on the input x_i and shared sources, $i = 2, 3, \ldots, n$. Then, each agent A_j outputs a_j depending on its input x_j and shared sources, $j = 1, n+1$. z_i and e_i denote the respective input and outcome of the measurement on each eavesdropper's system λ_i. If we permit agents A_1 and A_{n+1} to communicate with each other, Theorem 2 reduces to a recent result [17] for classical simulation. Generally, if all the independent agents $\mathsf{A}_1, \mathsf{A}_3, \ldots, \mathsf{A}_{n+1}$ for an even n ($\mathsf{A}_1, \mathsf{A}_3, \ldots, \mathsf{A}_{n-2}, \mathsf{A}_{n+1}$ for an odd n) can communicate with each other, from Theorem 2 we obtain the upper bound $k(2 - \mathcal{R}_k)$ for the classical correlations for these independent agents and the eavesdropper.

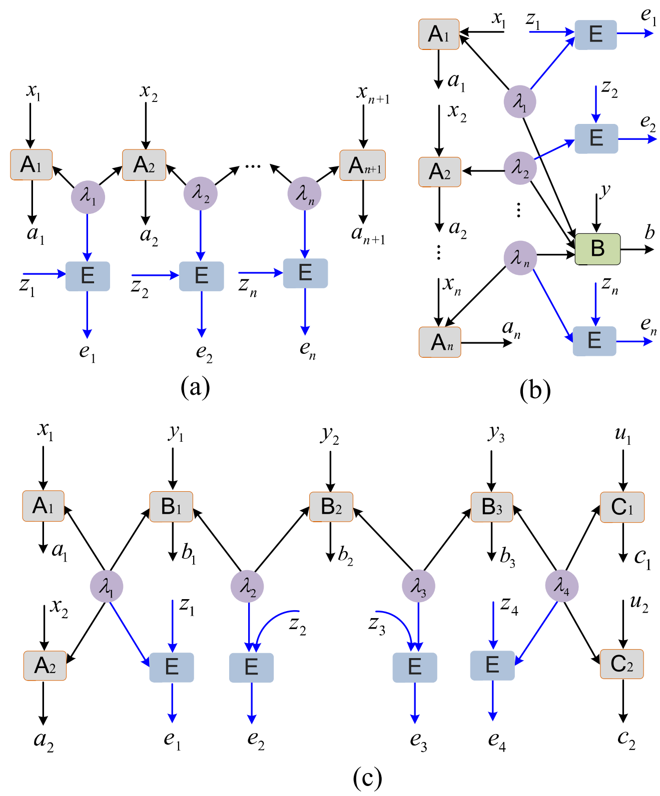

FIG. 7. Schematic IONBDAGs in terms of the device-independent information processing model. (a) Chain-shaped network. There are n independent hidden sources $\lambda_1, \lambda_2, \ldots, \lambda_n$. Each spacelike separated agent A_i shares some sources, λ_j's. (b) Star-shaped network. There are n independent hidden sources $\lambda_1, \lambda_2, \ldots, \lambda_n$. Each pair of spacelike separated agents A_i and B shares one source λ_i. (c) Hybrid chain-shaped network. There are four independent sources $\lambda_1, \lambda_2, \ldots, \lambda_4$. Each spacelike separated agent A_i, B_j, or C_s shares some sources. One eavesdropper holds some systems, each of which is correlated with one source λ_i.

B. Star-shaped networks

A general star-shaped network [26] is schematically shown in Fig. 7(b). It is proved that multipartite quantum correlations violate inequality (27) with $k = n$ [29] when the network consists of generalized EPR states. For device-independent information processing [44], assume that an eavesdropper holds n independent systems, where each system is correlated with one source λ_i and can be measured by a device with input z_i and output e_i. In the experiment, agent B first outputs b depending on its input y and shared sources. Then, each agent A_i outputs a_i depending on the input x_i and shared sources, $i = 1, 2, \dots, n$. z_i and e_i denote the respective input and outcome of the measurement on each eavesdropper's system λ_i. Assume that all the sources $\lambda_1, \lambda_2, \dots, \lambda_n$ are not correlated by the eavesdropper. Theorem 2 gives an upper bound of the leaking information of all the agents' outputs [44]. Otherwise, assume that partial sources $\lambda_{i_1}, \lambda_{i_2}, \dots, \lambda_{i_k}$ are not correlated. We can obtain from Theorem 2 the upper bound $k(2 - \mathcal{R}_k)$ of the eavesdropper's information, where $\mathcal{R}_k$ depends on all the independent agents $\mathsf{A}_{i_1}, \mathsf{A}_{i_2}, \dots, \mathsf{A}_{i_k}$ chosen for constructing the nonlinear inequality (27) given in Ref. [29].

C. Hybrid chain-shaped networks

Differently from the standard chain-shaped network shown in Fig. 7(a), a new network consisting of multipartite resources is shown in Fig. 7(c). A previous result [29] shows that multipartite quantum correlations violate inequality (27) with $k = 3$ when all resources consist of generalized EPR states and GHZ states, where A_1, B_2, C_1 are independent observers who have no preshared entanglement [29]. In the experiment, each agent B_i first outputs b_i depending on the input y_i and shared sources, $i = 1, 2, 3$. Then, agents A_i and C_j output one respective bit a_i and c_j. z_i and e_i denote the possible input and outcome of the measurement on each eavesdropper's system λ_i. Assume that an eavesdropper has four independent systems, each of which is correlated with one source. When A_1, B_2, C_1 are allowed to communicate with each other, Theorem 2 provides the upper bound $3(2 - \mathcal{R}_3)$ of the information relevant to these agents' outputs. Similar results hold for partially correlated hidden sources. For example, if λ_1 and λ_3 or λ_2 and λ_4 are correlated, from Theorem 2 we can also obtain the upper bound $2(2 - \mathcal{R}_2)$ of leakage information for an eavesdropper, where $\mathcal{R}_2$ depends on two independent agents, A_1 and C_1, for constructing the nonlinear inequality (27) given in Ref. [29].

VI. DISCUSSION AND CONCLUSIONS

Multipartite Bell causal correlations with multiple independent sources consist of star-convex sets which may inspire interesting applications in deep learning or artificial intelligence. From Theorem 1 the compatible nonsignaling correlations are featured in a simple input-output causal network using only locality relaxations. This framework is useful for identifying new multipartite causal structures that cannot reproduce quantum correlations. Another application is to derive new Bell-type inequalities [17] and quantum causal networks [10,47].

From Theorem 2 the eavesdropper's information relevant to independent observers' outcome is bounded by the violation of inequality (27). The result is reasonable because the statistics from separable measurements provides the maximal nonmultilocality by maximally violating the inequality [29]. This achievement suggests a device-independent key distribution in acyclic networks going beyond the standard Bell network. This is interesting for multipartite communication or quantum Internet [19]. An interesting problem is to establish a full security proof of these applications going beyond the bound provided.

In conclusion, we have presented a framework to characterize nonsignaling causal correlations by relaxing different assumptions on multisource networks. This model implies a star-convex set of correlations and is further exemplified by classifying all nonsignaling correlations of the entanglement swapping network. For large-scale applications in the presence of an eavesdropper, a unified device-independent information processing model is presented to bound the leaking information on all acyclic networks by making use of explicit nonlinear Bell inequalities. These results are both fundamentally interesting in Bell theory and significantly applicable in quantum information processing and communication complexity.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grants No. 61772437 and No. 61702427), Sichuan Youth Science and Technique Foundation (Grant No. 2017JQ0048), Fundamental Research Funds for the Central Universities (Grant No. 2018GF07), and Chuying Fellowship.

[1] J. S. Bell, Phys. Phys. Fiz. **1**, 195 (1964).
[2] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[3] J. v. Neumann, *In Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1931).
[4] C. Simon, V. Bužek, and N. Gisin, Phys. Rev. Lett. **87**, 170405 (2001).
[5] C. J. Wood and R. W. Spekkens, New J. Phys. **17**, 033002 (2015).
[6] M. J. W. Hall, Phys. Rev. Lett. **105**, 250404 (2010).
[7] G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, and N. Gisin, Phys. Rev. Lett. **113**, 190402 (2014).
[8] R. Chaves, R. Kueng, J. B. Brask, and D. Gross, Phys. Rev. Lett. **114**, 140403 (2015).
[9] C. M. Lee and R. W. Spekkens, J. Causal Infer. **5**, 20160013 (2017).
[10] M. J. W. Hall, Phys. Rev. A **84**, 022102 (2011).
[11] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).

[12] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Phys. Rev. Lett. **88**, 170405 (2002).
[13] M. Seevinck and G. Svetlichny, Phys. Rev. Lett. **89**, 060401 (2002).
[14] L. Aolita, R. Gallego, A. Cabello, and A. Acín, Phys. Rev. Lett. **108**, 100401 (2012).
[15] J.-D. Bancal, C. Branciard, N. Gisin, and S. Pironio, Phys. Rev. Lett. **103**, 090503 (2009).
[16] J.-D. Bancal, N. Brunner, N. Gisin, and Y.-C. Liang, Phys. Rev. Lett. **106**, 020405 (2011).
[17] R. Chaves, D. Cavalcanti, and L. Aolita, Quantum **1**, 23 (2017).
[18] L.-M. Duan, M. Lukin, J. I. Cirac, and P. Zoller, Nature **414**, 413 (2001).
[19] H. J. Kimble, Nature (London) **453**, 1023 (2008).
[20] S. Ritter, C. Nölleke, C. Hahn, A. Reiserer, A. Neuzner, M. Uphoff, M. Mücke, E. Figueroa, J. Bochmann, and G. Rempe, Nature **484**, 195 (2012).
[21] Q.-C. Sun, Y.-L. Mao, S.-J. Chen, W. Zhang, Y.-F. Jiang, Y.-B. Zhang, W.-J. Zhang, S. Miki, T. Yamashita, H. Terai, X. Jiang, T.-Y. Chen, L.-X. You, X.-F. Chen, Z. Wang, J.-Y. Fan, Q. Zhang and J.-W. Pan, Nat. Photon. **10**, 671 (2016).
[22] R. Valivarthi, M. G. Puigibert, Q. Zhou, G. H. Aguilar, V. B. Verma, F. Marsili, M. D. Shaw, S. W. Nam, D. Oblak, and W. Tittel, Nat. Photon. **10**, 676 (2016).
[23] M. Zukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).
[24] A. J. Short, S. Popescu, and N. Gisin, Phys. Rev. A **73**, 012101 (2006).
[25] C. Branciard, N. Gisin, and S. Pironio, Phys. Rev. Lett. **104**, 170401 (2010).
[26] A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acín, Phys. Rev. A **90**, 062109 (2014).
[27] R. Chaves, Phys. Rev. Lett. **116**, 010402 (2016).
[28] D. Rosset, C. Branciard, T. J. Barnea, G. Pütz, N. Brunner, and N. Gisin, Phys. Rev. Lett. **116**, 010403 (2016).
[29] M.-X. Luo, Phys. Rev. Lett. **120**, 140402 (2018).
[30] M.-X. Luo, npj Quantum Info. **5**, 91 (2019).
[31] J. Pearl, *Causality* (Cambridge University Press, New York, 2009).
[32] L. Masanes, Phys. Rev. Lett. **96**, 150501 (2006).
[33] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[34] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005).
[35] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).
[36] U. Vazirani and T. Vidick, Phys. Rev. Lett. **113**, 140501 (2014).
[37] S. Pironio, A. Acín, S. Massar, A. B. de La Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).
[38] R. Colbeck and R. Renner, Nat. Phys. **8**, 450 (2012).
[39] R. Ramanathan, F. G. S. L. Brandão, K. Horodecki, M. Horodecki, P. Horodecki, and H. Wojewódka, Phys. Rev. Lett. **117**, 230501 (2016).
[40] J. Anders and D. E. Browne, Phys. Rev. Lett. **102**, 050502 (2009).
[41] M. J. Hoban, E. T. Campbell, K. Loukopoulos, and D. E. Browne, New J. Phys. **13**, 023014 (2011).
[42] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going beyond Bell's Theorem, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer Academic, Dordrecht, Netherlands, 1989), p. 69.
[43] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).
[44] C. M. Lee and M. J. Hoban, Phys. Rev. Lett. **120**, 020504 (2018).
[45] C. Branciard, N. Brunner, H. Buhrman, R. Cleve, N. Gisin, S. Portmann, D. Rosset, and M. Szegedy, Phys. Rev. Lett. **109**, 100401 (2012).
[46] R. Cleve and H. Buhrman, Phys. Rev. A **56**, 1201 (1997).
[47] J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens, Phys. Rev. X **7**, 031021 (2017).
[48] J. Barrett and N. Gisin, Phys. Rev. Lett. **106**, 100406 (2011).
[49] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).
[50] O. Regev and B. Toner, SIAM J. Comput. **39**, 1562 (2010).
[51] N. S. Jones, N. Linden, and S. Massar, Phys. Rev. A **71**, 042329 (2005).
[52] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Phys. Rev. A **71**, 022101 (2005).
[53] J. Borwein and A. Lewis, *Convex Analysis and Nonlinear Optimization* (Springer, New York, 2000).
[54] J. C. H. Lee and P. Valiant, Optimizing star-convex functions, *IEEE 57th Annual Symposium on Foundations of Computer Science, New Brunswick, NJ, 2016* (IEEE, New York, 2016).
[55] J. Barrett, A. Kent, and S. Pironio, Phys. Rev. Lett. **97**, 170409 (2006).
[56] J. Barrett, R. Colbeck, and A. Kent, Phys. Rev. A **86**, 062326 (2012).
[57] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner, and V. Scarani, Phys. Rev. Lett. **107**, 050502 (2011).
[58] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).