Towards a minimal example of quantum nonlocality without inputs

Sadra Boreiri,[*] Antoine Girardin [⦿],[*] Bora Ulu, Patryk Lipka-Bartosik, Nicolas Brunner, and Pavel Sekatski

Department of Applied Physics, University of Geneva, 1211 Geneva, Switzerland

The network scenario offers interesting new perspectives on the phenomenon of quantum nonlocality. Notably, when considering networks with independent sources, it is possible to demonstrate quantum nonlocality without the need for measurement inputs, i.e., with all parties performing a fixed quantum measurement. Here we aim to find minimal examples of this effect. Focusing on the minimal case of the triangle network, we present examples involving output cardinalities of 3-3-3 and 3-3-2. A key element is a rigidity result for the parity token counting distribution, which represents a minimal example of rigidity for a classical distribution. Finally, we discuss the prospects of finding an example of quantum nonlocality in the triangle network with binary outputs and point out a connection to the Lovasz local lemma.

I. INTRODUCTION

The exploration of quantum nonlocality in networks has attracted growing attention in recent years (see, e.g., [1] for a recent review). This avenue of research opens interesting new perspectives and possibilities for quantum nonlocal correlations.

While quantum nonlocality has been investigated in a broad range of scenarios, including the multipartite case, the network scenario brings a main conceptual novelty. Specifically, the central idea is to consider networks where several sources distribute quantum resources to various subsets of the parties (nodes). In this sense, this model differs from the standard approach to multipartite nonlocality, where all parties are connected via a common source (see, e.g., [2]). The main assumption in the network scenario is then to consider that each source in the network is independent of all others [3–5]. From a more formal point of view, this implies that the relevant sets of possible correlations are nonconvex, as mixing two arbitrary strategies requires a source of randomness common to all parties. Hence, the standard methods aiming at the detection of quantum nonlocality with linear Bell tests are typically useless when discussing networks, and radically novel methods and concepts must be developed (see, e.g., [6–13]) .

Despite these challenges, recent works have brought significant insight into quantum nonlocality in networks. Notably, the few known examples of quantum nonlocal distributions show that the network scenario allows for novel forms of quantum nonlocality, that are possible only due to the network structure (in particular due to the independence of the sources). Remarkably, quantum nonlocality can be demonstrated in a network where none of the parties receive an input [4,5], that is, all the observers perform a single (fixed) measurement. This is in sharp contrast to the standard Bell scenario, where the presence of measurement inputs is fundamental. More specifically, in a scenario without inputs, the set of correlations obtainable with a classical source common to all parties coincides with the whole set of valid probability distributions. Consequently, there is no nonlocality without inputs in the standard Bell scenario.

A notable example of this phenomenon termed "quantum nonlocality without inputs" has been provided by Fritz in the triangle network [5], as shown in Fig. 1. While the example of Fritz can be viewed as an embedding of the well-known Clauser-Horne-Shimony-Holt (CHSH) Bell test in the triangle network (see also [14]), more recent work by Renou *et al.* [15], followed by other examples [16–19], suggested that a different form of quantum nonlocality without inputs can be also observed. Here, no obvious connection to standard Bell nonlocality can be made, suggesting a form of quantum nonlocality genuine to this network structure, as further examined in Ref. [20].

In this work, we investigate the question of finding a minimal example of quantum nonlocality without inputs. In the study of Bell nonlocality, the notion of minimal usually relates to the number of measurement inputs and outputs of the parties. In particular, the minimal scenario is that of the CHSH Bell inequality [21] where two parties perform two binary measurements each. It has been largely investigated and enables most of the applications of quantum nonlocality, such as device-independent quantum information processing, and most of the experimental demonstrations of quantum nonlocality.

In the context of networks, the notion of minimality can be defined in different ways. First, we aim to minimize the number of parties. Hence, we focus on networks with three parties. Among these, the only network that allows for nonlocality without inputs is the triangle network[1] (see Fig. 1). Next, as there are no measurement inputs, we want to minimize the number of measurement outputs. Previous examples of quantum nonlocality without inputs in the triangle network feature four-valued outputs. We use the notation 4-4-4 to denote the cardinality of such a scenario. To the best of our knowledge,

---

[1]Note that the other nontrivial network with three parties, the so-called bilocality network, does not allow for nonlocality without inputs; any distribution compatible with the bilocality network admits a local model [4].

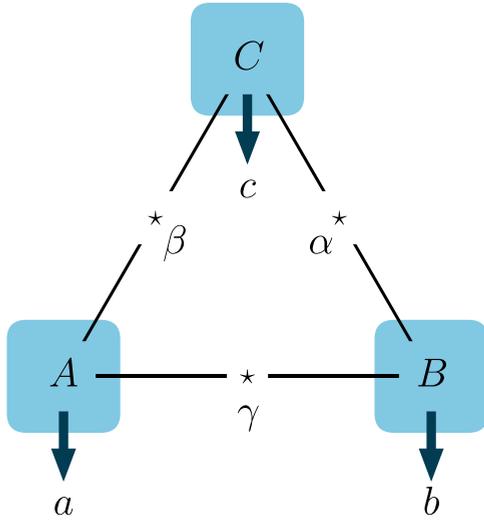[*]These authors contributed equally to this work.

FIG. 1. The triangle network features three distant parties. Each pair of parties is connected via a bipartite source. Importantly here, all parties perform a fixed measurement, i.e., they receive no input. In this work, we look for a minimal example (in terms of output cardinality) of quantum nonlocality.

the only improvement reported so far is a variant of the Fritz example given in Ref. [10], which features output cardinalities of 4-4-2.

Here we report on several examples of quantum nonlocality in the triangle network with output cardinalities 3-3-3 and 3-3-2. These examples are constructed from coarse graining of previous known examples. We provide analytic proofs of the nonlocality of these behaviors, and study their robustness to noise via numerical methods based on neural networks [22]. In particular, to prove the nonlocality of the 3-3-3 example we derive a rigidity result for network local models that we call parity token counting. We believe this is of independent interest, as it represents the minimal example (with cardinalities 2-2-2) of rigidity of a classical distribution. Finally, we discuss the prospects of finding an example of quantum nonlocality in the simplest triangle network with binary outputs.

II. PROBLEM AND METHODS

The triangle network consists of three parties (Fig. 1), namely, Alice, Bob, and Charlie. Each pair of parties is connected by a bipartite source. Importantly, the three sources are assumed to be independent of each other. Formally, independent sources prepare mutually independent random variables in the classical case and product quantum states in the quantum case. Importantly, the three parties do not have access to any common source. Upon receiving the physical resources the observers produce outputs (a, b, and c). In contrast to standard tests of Bell nonlocality, the observers receive no input in this setting. The statistics of the outputs are therefore given by the joint probability distribution $P(a, b, c)$. Despite the simplicity of the triangle network, finding quantum distributions that are provably nonlocal is a nontrivial task.

One of the central challenges in the study of network nonlocality is to find efficient methods for characterizing the set of distributions $P(a, b, c)$ which are obtainable from different

physical resources. In particular, the local set $\mathcal{L}$ consists of all distributions of the following form:

$$P_L(a, b, c) = \int d\alpha \, d\beta \, d\gamma \, P_A(a|\beta, \gamma) \, P_B(b|\gamma, \alpha) \, P_C(c|\alpha, \beta),$$
(1)

where α, β, γ are the local variables distributed by the sources [here $d\alpha \equiv d\mu(\alpha)$ means that the local variables are sampled from some underlying distributions $\mu(\alpha)$], and $P_A(a|\beta, \gamma)$, $P_B(b|\gamma, \alpha)$, and $P_C(c|\alpha, \beta)$ are the response functions of Alice, Bob, and Charlie, respectively. To prove that a distribution $P_Q(a, b, c)$ is not "triangle local," we thus need to demonstrate that it cannot be generated in the triangle network with classical sources. Formally, we need to show that $P_Q(a, b, c) \notin \mathcal{L}$.

The fundamental difficulty in verifying the existence of a local model according to Eq. (1) arises from the independence of the sources. Due to this independence, the local set $\mathcal{L}$ is nonconvex,[2] and the standard approach of characterizing the local set (a convex polytope) with linear Bell inequalities (facets of the polytope) cannot be applied. As a result, efficient bounds on the set of classical correlations (for example, in the form of nonlinear Bell inequalities) are still missing. Due to the lack of general tools, the known results on quantum network nonlocality without inputs are based on different arguments.

The example of Fritz is constructed by embedding a bipartite CHSH Bell test in the triangle network. It only involves a single quantum source, while the other two sources distribute the "inputs" for the measurement of the entangled quantum state. The nonlocality of the resulting distribution essentially follows from the violation of the CHSH inequality (or any other bipartite Bell inequality) [5,14].

The example of Renou *et al.* [15] which we refer to as the RGB4 distribution involves three quantum sources and entangled measurements. The proof of nonlocality of the RGB4 distribution relies on the concept of token counting rigidity [17,23]. It imposes severe constraints on the underlying classical model for certain triangle-local distributions, as discussed in the next section. It is also worth mentioning the construction of Gisin in Ref. [16] which is conjectured to be nonlocal based on numerical evidence [22], but for which a proof is still missing.

In this work, we look for minimal examples of quantum nonlocality in the triangle network. More specifically, we present examples of quantum distributions with low output cardinality and show that they do not admit a triangle-local model of Eq. (1) using a variety of methods. For our first example in Sec. III, with output cardinalities 3-3-3, we provide an analytic proof based on a relaxed version of the rigidity property, which we term "almost rigidity." Our second example in Sec. IV, with output cardinalities 3-3-2, is proven via the technique of inflation [6]. Loosely speaking, this method

---

[2]For a concrete example, consider the distribution where all parties always output "0," $P(000) = 1$, which is clearly triangle local. Similarly, the distribution where all parties output "1," $P(111) = 1$, is triangle local. However, the mixture of the two distributions violates the no-signaling condition [6], and is therefore not achievable in the triangle network.

considers larger (inflated) networks, some marginals of which coincide with the original triangle network. If one can show that the distribution on the inflated network cannot be consistently defined, given the marginals and some independence constraints, one can conclude that the original distribution is not triangle local. This technique also allows us to derive nonlinear Bell inequalities for detecting the corresponding distribution. More details can be found in Sec. IV and in Appendix E.

Finally, in order to investigate further examples of quantum nonlocality in the triangle scenario, we use numerical methods based on machine learning. Concretely, we use a method developed in Ref. [22] based on neural networks (NN). The code used in this work is adapted from the code available in this last reference. Here, a generative neural network (which we refer to as "LHV-Net") is used to explore the space of triangle-local models. This is ensured by encoding the structure of the triangle network into the neural network. Given a target distribution $P_{\text{target}}(a, b, c)$, the LHV-Net aims at constructing a local distribution $P_{\text{NN}}(a, b, c)$ which is as close as possible to the target one. This is done by minimizing the loss function, which in most cases is the Euclidean distance

$$d(P_{\text{target}}, P_{\text{NN}}) = \sqrt{\sum_{a,b,c} [P_{\text{target}}(a, b, c) - P_{\text{NN}}(a, b, c)]^2}. \quad (2)$$

For a given target distribution, it is often useful to investigate a family of distributions obtained by adding noise to the initial one. This can be done by mixing the target distribution with another distribution that is triangle local (for example, by adding noise to the quantum model). The visibility parameter V controls the amount of noise, with $V = 1$ corresponding to the initial distribution. If the initial distribution P_{target} is indeed nonlocal, at some value V^* we expect to see a sharp transition when monitoring the distance $d(\cdot, \cdot)$ as a function of V. The critical visibility V^* then corresponds to the point where the noisy distribution becomes triangle local and gives an estimate of the noise robustness of the initial distribution.

III. QUANTUM NONLOCALITY WITH OUTPUT CARDINALITY 3-3-3

We start by presenting an example of quantum nonlocality in the triangle network with output cardinalities 3-3-3. This example is based on a coarse graining of the RGB4 distribution of Ref. [15], which has cardinalities 4-4-4. In order to demonstrate the nonlocality of the coarse-grained distribution, we first present a result on "almost rigidity" for the task of parity token counting. Then we give the target quantum distribution and use the almost-rigidity property to prove that it is not triangle local.

A. Almost rigidity of parity token counting distributions

In the present context, rigidity [17,23] is a property of some classes of classical distributions in a network. The property ensures that any model [i.e., local variables and response functions displayed on the right-hand side of Eq. (1)] underlying a rigid probability distribution $P_L(a, b, c)$ can be brought to a unique canonical form by means of local relabeling of the variables α, β, γ (to a discrete set of values) which do not conflict with the response function. To be more precise, we

now elaborate on the general form of triangle-local models in Eq. (1) and Fig. 1.

To fix the notation we say that each source samples local variables α, β, or γ according to distributions $\mu(\alpha)$, $\eta(\beta)$, and $\nu(\gamma)$. Note that without loss of generality, we assumed that each source distributes two copies of the same variable.[3] Moreover, without loss of generality, the response functions can be considered deterministic, i.e., $P_A(a|\beta, \gamma) = 0$ or 1, because any randomness required for the choice of the outputs (local to the party) can be delegated to one of the neighboring sources. Hence, without loss of generality, we use the three distributions of local variables $\mu(\alpha), \eta(\beta), \nu(\gamma)$ and the three deterministic response functions $a(\beta, \gamma), b(\alpha, \gamma), c(\alpha, \beta)$ to describe a generic model underlying triangle-local distributions $P_L(a, b, c)$.

Let us now construct a family of triangle-local models that we call parity token counting (PTC) strategies. Each source has a single token that can be sent to either one or the other party it connects to. For example, the source connecting Bob and Charlie prepares $(\alpha_B, \alpha_C) = (1, 0)$ with probability p_α and $(\alpha_B, \alpha_C) = (0, 1)$ with probability $1 - p_\alpha$. Each party then outputs the parity of the total number of tokens it receives, i.e., for Charlie $c(\alpha_C, \beta_C) = \alpha_C \oplus \beta_C$, where "$\oplus$" is understood as sum modulo two, so that all the outputs are binary. All possible PTC strategies we just described give rise to distributions $P_{\text{PTC}}(a, b, c)$ that live in a three-dimensional subset $\mathcal{L}_{\text{PTC}}$ of the local set for the 2-2-2 triangle. The set $\mathcal{L}_{\text{PTC}}$ can be parametrized by $(p_\alpha, p_\beta, p_\gamma)$, however, the strategies $(p_\alpha, p_\beta, p_\gamma)$ and $(1 - p_\alpha, 1 - p_\beta, 1 - p_\gamma)$ are related by flipping the values of all local variables and thus lead to the same distribution. Except for a subset of $\mathcal{L}_{\text{PTC}}$ of measure zero these two strategies are the only ones that simulate the distribution. The full characterization of the set $\mathcal{L}_{\text{PTC}}$, including the multiplicities of the PTC strategies underlying each distribution $P_{\text{PTC}}(a, b, c)$, can be found in Appendix A.

We now ask what possible models lead to a distribution $P_{\text{PTC}}(a, b, c)$, except for the strategies we just defined. In fact, we will now show that PTC strategies are essentially the only classical strategies that lead to a distribution $P_{\text{PTC}}(a, b, c)$. This is formalized by the following theorem.

Theorem 1. Let $P_{\text{PTC}}(a, b, c)$ be the distribution arising from parity token parity counting strategy $(p_\alpha, p_\beta, p_\gamma)$ on the triangle network, where each source distributes its only token to the connected parties with probabilities p_i and $1 - p_i$, and the parties output the parity of the number of received tokens. For any other strategy $(\mu(\alpha), \eta(\beta), \nu(\gamma), a(\beta, \gamma), b(\alpha, \gamma), c(\alpha, \beta))$ that achieves the distribution $P_{\text{PTC}}(a, b, c)$ there exist functions $T_i^j : \mathcal{S}_i \to \{0, 1\}$ ($\mathcal{S}_i$ is the set of all possible values of the local variable produced by the source S_i) for any $\mathcal{S}_i \to A_j$ such that

(i) $T_\alpha^b(\alpha) + T_\alpha^c(\alpha) = T_\beta^c(\beta) + T_\beta^a(\beta) = T_\gamma^a(\gamma) + T_\gamma^b(\gamma) = 1$.

---

[3]One can also consider sources sampling two variables that may take different values and sending them to different parties, e.g., (α_B, α_C) with α_B sent to Bob and α_C to Charlie. This case is, however, included in the previous one, as the source could alternatively distribute two copies of pair $\alpha = (\alpha_B, \alpha_C)$ to both parties while response functions ignore the respective half of the pair.

(ii) $a(\beta, \gamma) = T^a_\beta(\beta) \oplus T^a_\gamma(\gamma)$, $b(\gamma, \alpha) = T^b_\gamma(\gamma) \oplus T^b_\alpha(\alpha)$, $c(\alpha, \beta) = T^c_\alpha(\alpha) \oplus T^c_\beta(\beta)$.

In addition, if the distribution is such that $P(a = 1)$, $P(b = 1)$, $P(c = 1) \neq \frac{1}{2}$, the functions T^j_i can be chosen to fulfill

(iii) $(\mathbb{E}[T^b_\alpha], \mathbb{E}[T^c_\beta], \mathbb{E}[T^a_\gamma]) = (p_\alpha, p_\beta, p_\gamma)$,

where $\mathbb{E}[T^b_\alpha] = \int d\alpha \, T^b_\alpha(\alpha)$.

The proof is given in Appendix B. We now discuss how to interpret it. First, note that the six local compression functions T^j_i map the local variable at each output of each source into a bit, e.g.,

$$
\begin{aligned}
(T^b_\alpha, T^c_\alpha) : \mathcal{S}_\alpha &\to \{0, 1\}^{\times 2}, \\
\alpha &\mapsto (T^b_\alpha(\alpha), T^c_\alpha(\alpha)),
\end{aligned}
\tag{3}
$$

for the two outputs of the source S_α. Furthermore, the condition (i) guarantees that for any α the only possible values are $(\alpha_B, \alpha_C) = (T^b_\alpha(\alpha), T^c_\alpha(\alpha)) = (1, 0)$ or $(0, 1)$, so that after the application of the compression functions to the outputs of a source it becomes a PTC source. On top of this, condition (ii) guarantees that the response functions of the original strategy are consistent with such compression, e.g., $c(\alpha, \beta) = \alpha_C \oplus \beta_C = T^c_\alpha(\alpha) \oplus T^c_\beta(\beta)$. That is, the response functions can only depend on the local variables via the values of the bits after compression. In other words, any information encoded in α other than the single bit given by $(T^b_\alpha(\alpha), T^c_\alpha(\alpha))$ is ignored by the response functions.

Finally (iii) guarantees that in most cases there is a unique PTC strategy (up to the choice of the functions T^j_i) underlying each PTC distribution. Nevertheless, for the following discussion, we will be interested in distributions with unbiased outputs for which (iii) does not apply. The cases where at least one output is unbiased (a subset of $\mathcal{L}_{\text{PTC}}$ of measure zero) are degenerate, i.e., there are infinitely many PTC strategies that lead to say same distribution (see Appendix A). Since (iii) does not hold for all distributions we refer to Theorem 1 as almost rigidity.

Two things are worth mentioning before moving on. First, the almost-rigidity property of PTC distributions holds for all networks satisfying the no double common source (NDCS) condition [24]. Second, PTC rigidity works for binary outputs, which is in contrast to token counting and color matching distributions requiring at least ternary outputs [23]. So, in some sense, it provides the minimal example of network rigidity. This is also why it can be used to detect nonlocality in the 3-3-3 triangle as we now show.

B. Quantum strategy

Consider each source distributing a maximally entangled two-qubit state

$$
|\psi^+\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle).
\tag{4}
$$

Each party performs the following two-qubit joint measurement with ternary outputs $a, b, c \in \{\bar{0}, \bar{1}_0, \bar{1}_1\}$:

$$
\begin{aligned}
\bar{0} &: |00\rangle\langle00| + |11\rangle\langle11|, \\
\bar{1}_0 &: |\bar{1}_0\rangle\langle\bar{1}_0|, \\
\bar{1}_1 &: |\bar{1}_1\rangle\langle\bar{1}_1|,
\end{aligned}
\tag{5}
$$

where we define $|\bar{1}_0\rangle = u|01\rangle + v|10\rangle$, $|\bar{1}_1\rangle = v|01\rangle - u|10\rangle$, with $0 < v < u < 1$ real parameters that satisfy $u^2 + v^2 = 1$. This strategy is similar to the RGB4 construction, with the difference that the positive operator-valued measurement (POVM) elements $|00\rangle\langle00|$ and $|11\rangle\langle11|$ are here coarse grained into a unique output $\bar{0}$.

Combining the above states and measurements, we obtain a quantum distribution denoted $P^{333}_Q(a, b, c)$ and given in Appendix C, for which we can prove the following.

Theorem 2. The quantum distribution $P^{333}_Q(a, b, c)$ is not triangle local [incompatible with any model of the form (1)] for $u^2_{\max} < u^2 < 1$, where $u^2_{\max} = \frac{-3+(9+6\sqrt{2})^{2/3}}{2(9+6\sqrt{3})^{1/3}} \approx 0.785$.

Below we sketch the proof of this theorem, while all details are given in Appendix C. The structure of the proof is similar to Ref. [15], but it is based on the novel PTC rigidity result which is essential to make the proof work for ternary outputs. First, we observe that if the outputs $\bar{1}_0$ and $\bar{1}_1$ are coarse grained for all parties the distribution resulting from $P^{333}_Q(a, b, c)$ is triangle local and PTC. The rigidity of PTC distributions (Theorem 1) imposes severe constraints on the underlying classical models. Now, it can be shown that if the distribution $P^{333}_Q(a, b, c)$ was simulated by a classical model (before coarse graining) these constraints would be impossible to satisfy. Therefore, $P^{333}_Q(a, b, c)$ is necessarily not triangle local.

C. Noise robustness

The above result applies to the (noiseless) distribution $P^{333}_Q(a, b, c)$, but does not extend to the case where noise is added, as the coarse-grained version of the noisy distribution is no longer PTC in general. Thus, to investigate noise robustness we must resort to other methods.

A first possibility would be to use the inflation method. Unfortunately, using it we could not prove the nonlocality of the above 3-3-3 quantum distribution. Note that the inflation method is able to detect nonlocality of the original RGB4 distribution, albeit with an extremely small noise tolerance [25].

Instead, we move to numerical methods and use LHV-Net. For the noise model, we add white noise to the entangled states of Eq. (4) produced by each source, which now prepares Werner states

$$
\rho_V = V|\psi^+\rangle\langle\psi^+| + (1 - V)\mathbb{1}/4.
\tag{6}
$$

Leaving the measurement in Eq. (5) unchanged, we obtain a family of distributions $P^{333}_{Q|V}$.

In Fig. 2, we plot the minimal distance [see Eq. (2)] found by LHV-Net, when taking $P^{333}_{Q|V}$ for the target distribution. We observe that, as V decreases, the minimal distance undergoes a transition around the critical visibility $V^* \approx 0.99$, demonstrating a small robustness to noise. Note that, for the original RGB4 distribution (with four-valued outputs), the estimated noise tolerance (also via LHV-Net) is much larger [22]. Hence, it seems that the coarse graining severely decreases the robustness to noise in this case.

Moreover, LHV-Net also allows us to investigate different coarse grainings of the RGB4 distribution. In particular, we find that when combining the outputs in the following way, $0 : |00\rangle\langle00| + |\bar{1}_0\rangle\langle\bar{1}_0|$, $1 : |11\rangle\langle11|$ and $2 : |\bar{1}_1\rangle\langle\bar{1}_1|$, the
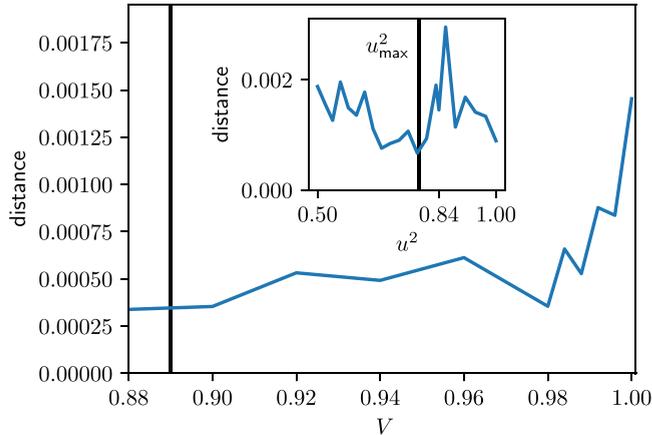
FIG. 2. Noise robustness of the 3-3-3 quantum distribution. The plot shows the minimal distance $d(P_{target}, P_{NN})$ found by the neural network (LHV-Net) between the target distribution and the closest local distribution, for different levels of noise, as quantified by the visibility parameter V. Here, the measurements are given by setting $u^2 = 0.84$, and we consider two possible coarse grainings, the blue curve is for the measurements given in Eq. (5), the orange curve coarse grain as $0 : |00\rangle\langle00| + |\bar{1}_0\rangle\langle\bar{1}_0|$, $1 : |11\rangle\langle11|$, and $2 : |\bar{1}_1\rangle\langle\bar{1}_1|$. The vertical line represents the critical visibilities for the original four-output RGB4 distribution as found in [22]. Note that the coarse-grained 3-3-3 distributions appear to be much less robust to noise. The inset shows the distance for different values of the measurement parameter u^2. The local distributions found by LHV-Net are very close to the three-output RGB4 distributions except around $u^2 = 0.84$.

resulting 3-3-3 distribution also appears to be nonlocal. The noise robustness is similar to the previous case. It is also worth mentioning that the coarse graining $0 : |00\rangle\langle00|$, $1 : |\bar{1}_0\rangle\langle\bar{1}_0| + |\bar{1}_1\rangle\langle\bar{1}_1|$, and $2 : |11\rangle\langle11|$ results in product measurements and gives a classical distribution.

Finally, we also investigated binary coarse grainings of the RGB4 distribution, but according to LHV-Net, all resulting distributions appear to be triangle local. Of course, this is only a numerical result and it should be considered carefully. It could be in principle that the resulting distribution is nonlocal but extremely close to the set of triangle-local distributions.

D. Elegant distribution

Using LHV-Net, we also investigated coarse graining of the so-called "elegant distribution" proposed by Gisin [16]. The distribution is based on all sources producing a Bell state, here chosen to be the singlet state $|\psi^-\rangle = 1/\sqrt{2}(|01\rangle - |10\rangle)$. Each party performs an entangling measurement with the four outcomes corresponding to projectors $\{|\Phi_1\rangle\langle\Phi_1|, \ldots, |\Phi_4\rangle\langle\Phi_4|\}$ on two-qubit states

$$|\Phi_j\rangle = \sqrt{\frac{3}{2}} |m_j, -m_j\rangle + i\frac{\sqrt{3}-1}{2} |\psi^-\rangle, \qquad (7)$$

with the four vectors $\{|m_j\rangle\}_{j=1,2,3,4}$ forming a tetrahedron on the Bloch sphere. A notable difference with the RGB4 measurement in Eq. (5) is that here the states $|\Phi_j\rangle$ corresponding to all measurement outcomes are entangled, with the same level of entanglement (they are all related by local unitary
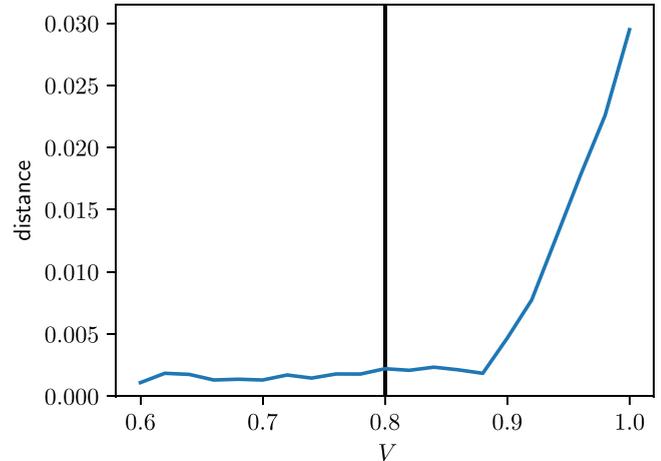

FIG. 3. Noise robustness for the coarse-grained elegant distribution. The plot shows the distance $d(P_{target}, P_{NN})$ between the target distribution and the closest local distribution found by LHV-Net, for different values of the visibility V. The vertical line represents the critical visibility for the original four-output distributions as found in [22]. Again, it seems that coarse graining reduces robustness to noise.

transformations). The resulting distribution is symmetric under permutation of outputs and parties and is given by

$$P(a = b = c) = \frac{25}{256},$$
$$P(a = b \neq c) = \frac{1}{256},$$
$$P(a \neq b \neq c \neq a) = \frac{5}{256}. \qquad (8)$$

There is currently no known proof that the elegant distribution is not triangle local.

Given the level of symmetry of this distribution, it is enough to consider a single coarse graining to a 3-3-3 distribution, with the constraint that each party performs the same coarse graining. Here LHV-Net also suggests that the resulting distribution remains nonlocal, as shown in Fig. 3. Again, it seems that the noise tolerance is reduced compared to the original 4-4-4 distribution.

Finally, note that when using LHV-Net to analyze different ternary coarse grainings for the parties or binary coarse grainings, it appears that all resulting distributions are triangle local.

IV. QUANTUM NONLOCALITY WITH OUTPUT CARDINALITY 3-3-2

We now present an example of quantum nonlocality in the triangle that uses even smaller cardinalities for the outputs, namely, 3-3-2. This example is constructed from coarse graining the Fritz distribution [5].

Let us first recall the model of Fritz. The idea is to have Alice and Bob perform a standard CHSH Bell test [21]. Hence, they share a singlet Bell state $|\psi^-\rangle$. Now, the binary measurement inputs that are required for both Alice and Bob for testing the CHSH inequality are provided by the two extra sources. Specifically, the β source provides a uniformly random bit $x = 0, 1$ while the α source produces a uniformly random bit $y = 0, 1$. Upon receiving their effective inputs x
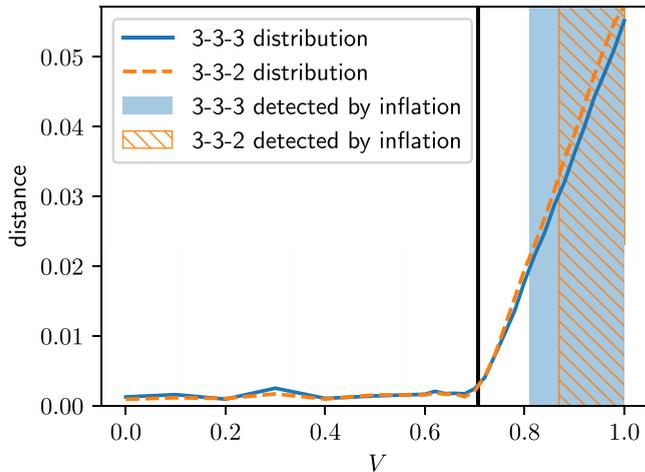
FIG. 4. Noise robustness of the coarse-grained Fritz distributions. The plot shows the distance $d(P_{\text{target}}, P_{\text{NN}})$ between the target distribution and the closest local distribution found by LHV-Net, for different values of the visibility V. Both 3-3-2 and 3-3-3 cases are shown. The vertical line represents the critical visibility ($V^* = 1/\sqrt{2}$) for the original four-output distributions (see [22]). The colored regions indicate the visibilities detected via the web inflation (see Appendix E for details).

and y, Alice and Bob perform the corresponding local Pauli measurements [σ_z or σ_x for Alice and $(-\sigma_z - \sigma_x)/\sqrt{2}$ or $(-\sigma_z + \sigma_x)/\sqrt{2}$ for Bob] and obtain binary outputs a' and b'. Finally, Alice outputs $a = (a', x)$, Bob $b = (b', y)$, and Charlie $c = (x, y)$. Note that it is crucial that Charlie also broadcasts both effective inputs x and y in order to ensure the condition of measurement independence.

Fritz showed that since the values of the outputs x (y) are perfectly correlated between Charlie and Alice (Charlie and Bob), these outputs have to be independent of the source connecting Alice and Bob [5]. Therefore, the resulting distribution $P_F(a, b, c)$ is nonlocal whenever the conditional distribution $P(a', b'|x, y)$ violates the CHSH Bell inequality. When adding noise to the singlet state shared by Alice and Bob [similarly to Eq. (6)], one finds a critical visibility of $V^* = 1/\sqrt{2}$.

We first consider the following coarse graining. On Alice's side, the output is ternary and given by $a = 0$ if $a' = 0$, $a = 1$ if $x = 0$ and $a' = 1$, and $a = 2$ if $x = 1$ and $a' = 1$. Similarly, on Bob's side, we define a ternary output according to $b = 0$ if $x = 0$ and $b' = 0$, $b = 1$ if $y = 1$ and $b' = 0$, and $b = 2$ if $b' = 1$. Finally, on Charlie, we get a binary output: $c = xy$, i.e., $c = 1$ only if $x = y = 1$. Note that we consider here uniform distributions for both effective inputs to be biased, specifically $p(x = 1) = p(y = 1) = \frac{1}{2}$, which renders the binary output c biased. We have checked that changing these input distributions only marginally influences the result.

The resulting distribution, with cardinality 3-3-2 is nonlocal, which we can prove using the inflation technique. Specifically, we consider the so-called web inflation, which allows us to construct a nonlinear inequality for detecting the above coarse-grained distribution for visibility greater than ≈ 0.87. This value is higher compared to the critical visibility of the original Fritz distribution. This appears to be a limi-

tation of our inflation since using LHV-Net, we find that the coarse-grained 3-3-2 distribution still has a critical visibility close to the original value $1/\sqrt{2} \simeq 0.71$. These results are shown in Fig. 4.

More generally, there exist many different possible coarse grainings of the Fritz distribution that lead to interesting 3-3-2 or 3-3-3 distributions. In Appendix E, we provide a detailed analysis of all combinations resulting in triangle-nonlocal distributions and discuss their noise tolerance.

Finally, we also tried here to obtain quantum nonlocality from coarse graining to a binary distribution, but in all cases, LHV-Net can reproduce the resulting distribution with excellent accuracy.

V. TRIANGLE WITH BINARY OUTPUTS

Finding an example of quantum nonlocality in the triangle with binary outputs would of course provide a minimal example of network nonlocality. At this point, it is still an open question whether this is possible or not. We obtained relatively convincing numerical evidence that all examples discussed above become triangle local when coarse grained to binary outputs. It also seems that the rigidity-based proofs of nonlocality, along the lines of Theorem 2, cannot work for binary quantum distributions since they require to elaborate on a coarse-groaned version of the original distribution.

We also investigated another approach based on a connection to the Lovasz local lemma. The latter states that when considering events that are almost independent and individually not very likely, there is always a nonvanishing probability that none of them occurs. A refined version called variational Lovasz local lemma (VLLL) [26] applies to events described by a fixed event-variable graph, or to network-local models with binary outputs in the language of this paper. When applied to the case of the triangle network, Theorem 4 in [26] implies that for any triangle-local strategy

$$P(j = 0) < p^* \quad \forall\, j = a, b, c \implies P(a = b = c = 1) > 0,$$
(9)

where $p^* = (\sqrt{5} - 1)/2 \approx 0.38$. Interestingly, the above relation is tight and can be saturated with a classical strategy using only two sources[5] (i.e., corresponding to the bilocality network). The implication (9) is not necessarily true for quantum strategies. Moreover, finding a quantum example $P_Q(a, b, c)$ that violates the VLLL [the condition (9)] would reveal nonlocality in the 2-2-2 triangle. We have investigated this question numerically, but could not find an instance of a quantum violation.

---

[5]Consider a source distributing the shared variable $\beta = 0, 1$ to A and C with probability $P(\beta = 1) = p^*$. Similarly, B and C are connected via a source sending $\alpha = 0, 1$ with probability $P(\alpha = 1) = p^*$. Then consider that A outputs β, B outputs α, while C outputs $c = \alpha\beta \oplus 1$. It is straightforward to check that the resulting distribution saturates the condition given in Eq. (9). That is, it leads to $P(a = 0) = P(b = 0) = P(c = 0) = p^*$ and $P(a = b = c = 1) = 0$.

VI. DISCUSSION

We investigated the question of finding a minimal example of quantum nonlocality without inputs in a network. We focused on the triangle network, as this involves the minimal number of parties, i.e., three. We found two classes of examples involving, respectively, output cardinalities 3-3-3 and 3-3-2. The examples were constructed from coarse graining of previously known examples. To prove some of our results, we considered the scenario of parity token counting, for which we proved an almost-rigidity property, which can be of independent interest as it represents a minimal example of rigidity of a classical distribution. We also discussed the noise robustness of these examples.

The main question left open here is whether there exists an example of quantum nonlocality in the minimal triangle network with binary outputs. For all examples we found, it appears that further coarse graining to obtain binary outputs leads to triangle-local distributions. We also established a connection to the Lovasz local lemma, but could not find a quantum nonlocal distribution with this approach. We note, however, that a slightly more complicated network, namely, a ring network with four parties (i.e., a square) allows for quantum nonlocality without inputs and binary outputs [27].

In addition to the binary outputs question, it would be interesting to find an example of nonlocality without inputs in the 3-2-2 triangle network. Here the set of correlation $P(a, b, c)$ is described by $11 = 3 \times 2 \times 2 - 1$ real parameters. This is lower than the $12 = 2 \times 2 \times 2 \times 2 - 4$ parameters needed to describe the set of correlations $P(a, b|x, y)$ in the CHSH scenario. Hence, from the perspective of the dimension of the correlation set, an example of quantum nonlocality in the triangle with output cardinality 3-2-2 would be "more minimal" than the minimal example of standard Bell nonlocality.

Finally, another interesting question is whether the 3-3-3 examples we constructed here feature some stronger form of network correlations, for example, genuine network nonlocality [20] or full network nonlocality [28]. Since our 3-3-2 example is constructed from the Fritz distribution, these can be neither genuine network nonlocal nor full network nonlocal.

APPENDIX A: THE SET $\mathcal{L}_{\text{PTC}}$ OF PARITY TOKEN COUNTING DISTRIBUTIONS IN THE 2-2-2 TRIANGLE

PTC local models defined in the main text are constructed with bipartite sources sampling:

$$(\alpha_B, \alpha_C) = \begin{cases} (1, 0) & \text{with probability} & p_\alpha, \\ (0, 1) & \text{with probability} & 1 - p_\alpha, \end{cases}$$

$$(\beta_C, \beta_A) = \begin{cases} (1, 0) & \text{with probability} & p_\beta, \\ (0, 1) & \text{with probability} & 1 - p_\beta, \end{cases} \quad \text{(A1)}$$

$$(\gamma_A, \gamma_B) = \begin{cases} (1, 0) & \text{with probability} & p_\gamma. \\ (0, 1) & \text{with probability} & 1 - p_\gamma \end{cases}$$

and fixed binary response functions

$$a(\beta_A, \gamma_A) = \beta_A \oplus \gamma_A, \quad b(\gamma_B, \alpha_B) = \gamma_B \oplus \alpha_B,$$
$$c(\alpha_C, \beta_C) = \alpha_C \oplus \beta_C. \quad \text{(A2)}$$

Obviously, the set $\mathcal{L}_{\text{PTC}}$ of PTC distributions is parametrized by the three values $(p_\alpha, p_\beta, p_\gamma)$ and is at most of dimension 3. $\mathcal{L}_{\text{PTC}}$ is a subset of the seven-dimensional local set $\mathcal{L}_{2\text{-}2\text{-}2}$ for the 2-2-2 triangle. The exact dependence of the probability distribution $P_{\text{PTC}}(a, b, c)$ on the parameters of the sources are given in Table I. From there one can easily verify that $\mathcal{L}_{\text{PTC}}$ is indeed three dimensional, e.g., by noting that $P_{\text{PTC}}(1, 0, 0)$, $P_{\text{PTC}}(0, 1, 0)$, and $P_{\text{PTC}}(0, 0, 1)$ are linearly independent polynomials of $(p_\alpha, p_\beta, p_\gamma)$. One can also check that the transformation $(p_\alpha, p_\beta, p_\gamma) \to (1 - p_\alpha, 1 - p_\beta, 1 - p_\gamma)$ does not affect the distribution.

To describe the set $\mathcal{L}_{\text{PTC}}$ it is convenient to introduce the one-party correlators

$$E_A = 2P(a = 1) - 1, \quad E_B = 2P(b = 1) - 1,$$
$$E_C = 2P(c = 1) - 1, \quad \text{(A3)}$$

which satisfy the equations

$$P(a = 1) = \frac{1 + E_A}{2} = p_\beta p_\gamma + (1 - p_\beta)(1 - p_\gamma) \quad \text{(A4)}$$

upon party permutations. Given these equations it is straightforward to rewrite the full distribution $P_{\text{PTC}}(a, b, c)$ in terms of the correlators (E_A, E_B, E_C). It is as given in Table II. Since any PTC distribution can be reconstructed completely given the values of the correlators, it is natural to ask which values (E_A, E_B, E_C) are actually achievable by some PTC strategy. To answer this question we will now distinguish three disjoint subsets of $\mathcal{L}_{\text{PTC}}$ depending on the values of correlators.

1. The generic case $E_A, E_B, E_C \neq 0$. Without loss of generality we permute the parties such that $E_A \geqslant E_B \geqslant E_C$. Formally solving the equations (A4) for all parties gives two

TABLE I. The table of probabilities $P_{\text{PTC}}(a, b, c)$ as a function of $(p_\alpha, p_\beta, p_\gamma)$.

$(b, c)\backslash a$	0	1
(0,0)	0	$p_\alpha(1 - p_\beta)(1 - p_\gamma) + (1 - p_\alpha)p_\beta p_\gamma$
(0,1)	$p_\alpha p_\beta(1 - p_\gamma) + (1 - p_\alpha)(1 - p_\beta)p_\gamma$	0
(1,0)	$p_\alpha(1 - p_\beta)p_\gamma + (1 - p_\alpha)p_\beta(1 - p_\gamma)$	0
(1,1)	0	$p_\alpha p_\beta p_\gamma + (1 - p_\alpha)(1 - p_\beta)(1 - p_\gamma)$

TABLE II. The table of probabilities $P_{\text{PTC}}(a, b, c)$ as a function of (E_A, E_B, E_C).

$(b, c)\backslash a$	0	1
(0,0)	0	$\frac{1}{4}(1 + E_A - E_B - E_C)$
(0,1)	$\frac{1}{4}(1 - E_A - E_B + E_C)$	0
(1,0)	$\frac{1}{4}(1 - E_A + E_B - E_C)$	0
(1,1)	0	$\frac{1}{4}(1 + E_A + E_B + E_C)$

solutions

$$
\begin{pmatrix} p_\alpha \\ p_\beta \\ p_\gamma \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \pm \sqrt{\frac{E_B E_C}{E_A}} \\ 1 \pm \sqrt{\frac{E_A E_C}{E_B}} \\ 1 \pm \sqrt{\frac{E_A E_B}{E_C}} \end{pmatrix}. \tag{A5}
$$

Note that the two solutions are related by the transformation $(p_\alpha, p_\beta, p_\gamma) \rightarrow (1 - p_\alpha, 1 - p_\beta, 1 - p_\gamma)$ mentioned above. This degeneracy is thus easy to understand, and we can focus on the solution with the "+" sign. The solution in Eq. (A5) makes physical sense if and only if the numbers on the right-hand side are probabilities, that is, $(*) \sqrt{\frac{E_B E_C}{E_A}} \in \mathbb{R}$ and $(**) \sqrt{\frac{E_B E_C}{E_A}} \leqslant 1$ for all party permutations. The reality conditions imply that only an even number of correlators E_A, E_B, E_C can be negative, leading to two different cases ($E_A \geqslant E_B \geqslant E_C \geqslant 0$) or ($E_A \geqslant 0 > E_B \geqslant E_C$) that fulfill $(*)$.

1a. For ($E_A \geqslant E_B \geqslant E_C > 0$), the fractions are ordered as $\sqrt{\frac{E_A E_B}{E_C}} \geqslant \sqrt{\frac{E_A E_C}{E_B}} \geqslant \sqrt{\frac{E_B E_C}{E_A}}$, hence, the condition $(**)$ is verified iff $\sqrt{\frac{E_A E_B}{E_C}} \leqslant 1$ or simply $E_A E_B \leqslant E_C$.

1b. For ($E_A > 0 > E_B \geqslant E_C$), $E_C \leqslant E_B$ and thus $|E_C| \geqslant |E_B|$. The fractions are partially ordered $\sqrt{\frac{E_A E_B}{E_C}} \leqslant \sqrt{\frac{E_A E_C}{E_B}}$ and the condition $(**)$ is fulfilled iff both $\sqrt{\frac{E_A E_C}{E_B}}, \sqrt{\frac{E_B E_C}{E_A}} \leqslant 1$. The two inequalities can be rewritten as $E_A|E_C| \leqslant |E_B| \leqslant \frac{E_A}{|E_C|}$.

2. The measure zero case $E_A E_B E_C = 0$. Without loss of generality let us assume that $E_A = 0$. By Eq. (A4) we have

$$
\frac{1}{2} = p_\beta p_\gamma + (1 - p_\beta)(1 - p_\gamma) \Rightarrow p_\beta = \frac{1}{2} \quad \text{or} \quad p_\gamma = \frac{1}{2}. \tag{A6}
$$

Let us consider the case where $p_\beta = \frac{1}{2}$. This automatically implies that $E_C = 0$ as well. Furthermore, with the help of Table I one computes the full probability distribution

$$
P_{\text{PTC}}(1, 0, 0) = P_{\text{PTC}}(0, 0, 1)
$$
$$
= \frac{1}{2}[p_\alpha(1 - p_\gamma) + (1 - p_\alpha)p_\gamma] = \frac{1 - E_B}{4},
$$
$$
P_{\text{PTC}}(0, 1, 0) = P_{\text{PTC}}(1, 1, 1)
$$
$$
= \frac{1}{2}[p_\alpha p_\gamma + (1 - p_\alpha)(1 - p_\gamma)] = \frac{1 + E_B}{4}. \tag{A7}
$$

There is thus a single degree of freedom left, and it can be parametrized by the value of the remaining correlator E_B. Its value

$$
E_B = 2[p_\alpha p_\gamma + (1 - p_\alpha)(1 - p_\gamma)] - 1 \tag{A8}
$$

can be chosen freely in the interval $[-1, 1]$, and any choice of p_α and p_γ fulfilling Eq. (A8) realizes it. Hence, we conclude that there exists a PTC distribution with $(E_A, E_B, E_C) = (0, E_B, 0)$ (and any permutation of parties), and there are infinitely many PTC strategies that achieve it. Again we can distinguish two cases.

2a. For a non-negative value $E_B \geqslant 0$ one can exchange the parties A and B and realize that $(E_A \geqslant 0, 0)$ is a continuation of the case (1a) where ($E_A \geqslant E_B \geqslant E_C > 0$).

2b. Similarly, for a negative value $E_B < 0$ one exchanges B with C and notice that $(0, 0, E_C < 0)$ is the continuation of the case (1b) where ($E_A > 0 > E_B \geqslant E_C$).

We now combine all the cases 1 and 2 in the following observation.

Theorem 3. There exists a PTC distribution $P_{\text{PTC}}(a, b, c)$ with the correlator values $E_A \geqslant E_B \geqslant E_C$ (where $E_A = 2P(a = 1) - 1$) if and only if one of the following is true:

1a. $E_A, E_B, E_C > 0$ and $E_C \geqslant E_A E_B$,
1b. $E_A > 0$; $E_B, E_C < 0$ and $E_A|E_C| \leqslant |E_B| \leqslant \frac{E_A}{|E_C|}$,
2a. $E_A \geqslant 0$ and $E_B = E_C = 0$,
2b. $E_A = E_B = 0$ and $E_C < 0$.

The PTC distribution is unique and given in Table II. In addition, in the cases 1a and 1b the distribution can be simulated with exactly two strategies $(p_\alpha, p_\beta, p_\gamma) = (\frac{1}{2}(1 + \sqrt{\frac{E_B E_C}{E_A}}), \frac{1}{2}(1 + \sqrt{\frac{E_A E_C}{E_B}}), \frac{1}{2}(1 + \sqrt{\frac{E_A E_B}{E_C}}))$ and $(1 - p_\alpha, 1 - p_\beta, 1 - p_\gamma)$. In the cases 2a and 2b there are infinitely many PTC strategies (one-parameter family) that simulate the distribution.

Obviously, the observation remains true for any permutation of the parties. This gives a full characterization of the set $\mathcal{L}_{\text{PTC}}$ and the multiplicity of the underlying PTC strategies.

APPENDIX B: PROOF OF ALMOST RIGIDITY OF PARITY TOKEN COUNTING DISTRIBUTIONS

Proof. Take any value of the outputs $a_0 b_0 c_0$ that has nonzero probability and consider the values $\alpha_0, \beta_0, \gamma_0$ of the local variables that lead to $a_0 b_0 c_0$. Now, we define the derivative of an output with respect to a local variable as follows:

$$
\nabla_{\alpha_0} b(\gamma, \alpha) = b(\gamma, \alpha) \ominus b(\gamma, \alpha_0). \tag{B1}
$$

Here by $\ominus$ we denote the difference modulo 2, $x \ominus y = (x - y) \mod 2$.

Since $a \oplus b \oplus c = 1$ for any PTC distribution, we have

$$
\nabla_{\alpha_0} a(\beta, \gamma) \oplus \nabla_{\alpha_0} b(\gamma, \alpha) \oplus \nabla_{\alpha_0} c(\alpha, \beta) = 0.
$$

The output $a(\beta, \gamma)$ does not depend on α, therefore, we have $\nabla_{\alpha_0} a(\beta, \gamma) = a(\beta, \gamma) - a(\beta, \gamma) = 0$ and $\nabla_{\alpha_0} b(\gamma, \alpha) = \nabla_{\alpha_0} c(\alpha, \beta)$. Here $\nabla_{\alpha_0} b(\gamma, \alpha)$ is independent of β and $\nabla_{\alpha_0} c(\alpha, \beta)$ is independent of γ, being always equal they can only depend on α. We can thus define

$$
f_{\alpha_0}(\alpha) = \nabla_{\alpha_0} b(\gamma, \alpha) = \nabla_{\alpha_0} c(\alpha, \beta). \tag{B2}
$$

In a similar manner, one defines $f_{\beta_0}(\beta) = \nabla_{\beta_0} a(\beta, \gamma) = \nabla_{\beta_0} c(\alpha, \beta)$ and $f_{\gamma_0}(\gamma) = \nabla_{\gamma_0} a(\beta, \gamma) = \nabla_{\gamma_0} b(\alpha, \gamma)$.

Now, following Eq. (B2) and the definition of the derivative (B1), we have

$$b(\gamma, \alpha) = b(\gamma, \alpha_0) \oplus f_{\alpha_0}(\alpha),$$
$$b(\gamma, \alpha) = b(\gamma_0, \alpha) \oplus f_{\gamma_0}(\gamma).$$

Combining the two gives

$$b(\gamma, \alpha) = b(\gamma_0, \alpha_0) \oplus f_{\alpha_0}(\alpha) \oplus f_{\gamma_0}(\gamma)$$
$$= \underbrace{b_0 \oplus f_{\alpha_0}(\alpha)}_{\equiv T_\alpha^b(\alpha)} \oplus \underbrace{f_{\gamma_0}(\gamma)}_{\equiv T_\gamma^b(\gamma)},$$

where we defined the functions $T_\alpha^b(\alpha)$ and $T_\gamma^b(\gamma)$. Similarly, we can define

$$a(\beta, \gamma) = a_0 \oplus f_{\gamma_0}(\gamma) \oplus f_{\beta_0}(\beta)$$
$$= \underbrace{1 \oplus f_{\gamma_0}(\gamma)}_{\equiv T_\gamma^a(\gamma)} \oplus \underbrace{1 \oplus a_0 \oplus f_{\beta_0}(\beta)}_{\equiv T_\beta^a(\beta)},$$

$$c(\alpha, \beta) = c_0 \oplus f_{\beta_0}(\beta) \oplus f_{\alpha_0}(\alpha)$$
$$= \underbrace{1 \oplus b_0 \oplus c_0 \oplus f_{\beta_0}(\beta)}_{\equiv T_\beta^c(\beta)} \oplus \underbrace{1 \oplus b_0 \oplus f_{\alpha_0}(\alpha)}_{\equiv T_\alpha^c(\alpha)}.$$

Manifestly, defined in the above way the function T_j^i fulfills (ii). To prove (i) simply note that

$$T_\gamma^b(\gamma) \oplus T_\gamma^a(\gamma) = f_{\gamma_0}(\gamma) \oplus 1 \oplus f_{\gamma_0}(\gamma) = 1,$$

$$T_\alpha^b(\alpha) \oplus T_\alpha^c(\alpha) = b_0 \oplus f_{\alpha_0}(\alpha) \oplus 1 \oplus b_0 \oplus f_{\alpha_0}(\alpha) = 1,$$

$$T_\beta^a(\beta) \oplus T_\beta^c(\beta) = 1 \oplus a_0 \oplus f_{\beta_0}(\beta) \oplus 1 \oplus b_0 \oplus c_0 \oplus f_{\beta_0}(\beta) = 1,$$

where we used $a_0 \oplus b_0 \oplus c_0 = 1$ for the last equality.

Finally, let us prove (iii) for the distributions with $P(a = 1), P(b = 1), P(c = 1) \neq \frac{1}{2}$. In Appendix A (Theorem 3) we proved that in this case ($E_A, E_B, E_C \neq 0$) there are only two PTC strategies that achieve $P_{\text{PTC}}(a, b, c)$, namely, $(p_\alpha, p_\beta, p_\gamma)$ and $(1 - p_\alpha, 1 - p_\beta, 1 - p_\gamma)$. We thus know that there are functions T_i^j that satisfy (i) and (ii) and can only lead to two possibilities

$$(\mathbb{E}[T_\alpha^b], \mathbb{E}[T_\beta^c], \mathbb{E}[T_\gamma^a]) = (p_\alpha, p_\beta, p_\gamma) \quad \text{or}$$
$$(\mathbb{E}[T_\alpha^b], \mathbb{E}[T_\beta^c], \mathbb{E}[T_\gamma^a]) = (1 - p_\alpha, 1 - p_\beta, 1 - p_\gamma). \quad \text{(B3)}$$

In the first case (iii) holds automatically. In the second one can define new compression functions $\bar{T}_i^j = T_i^j \oplus 1$ with "flipped" token directions. These functions also satisfy (i) and (ii) and give $\mathbb{E}[\bar{T}_i^j] = 1 - \mathbb{E}[T_i^j]$. Hence, using them as compression functions gives back the original strategy $(p_\alpha, p_\beta, p_\gamma)$, and proves (iii). ∎

APPENDIX C: PROOF OF MAIN RESULT: THEOREM 2

The resulting probability distribution P_Q is given by

$$P_Q(a, b, c) = \text{Tr}[(a \otimes b \otimes c)|\Psi\rangle\langle\Psi|], \quad a, b, c \in \{\bar{0}, \bar{1}_0, \bar{1}_1\},$$

where in $|\Psi\rangle$, the respective Hilbert spaces are suitably ordered according to the triangle configuration:

$$|\Psi\rangle_{A_1 A_2 B_1 B_2 C_1 C_2} \equiv |\psi^+\rangle_{A_2 B_1} \otimes |\psi^+\rangle_{B_2 C_1} \otimes |\psi^+\rangle_{C_2 A_1}.$$
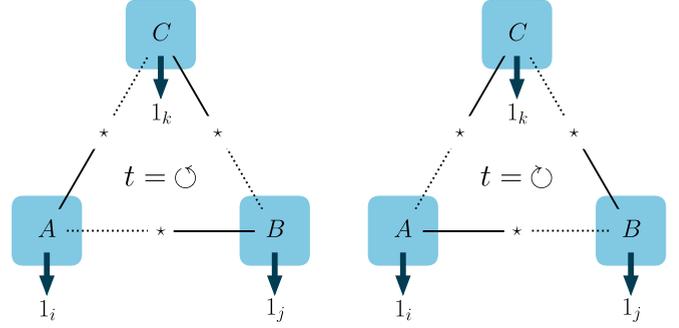

FIG. 5. In uniform token counting strategy over the triangle, when all parties obtain $\bar{1}$, the token paths are undetermined between $t = \circlearrowleft$ and $\circlearrowright$. In a quantum scenario the system could be in a superposition state of $t = \circlearrowleft$ and $\circlearrowright$. However, in a classical scenario the underlying trajectory of tokens is either $t = \circlearrowleft$ or $t = \circlearrowright$.

Introducing the notation $u_0 = -v_1 = u$ and $v_0 = u_1 = v$ (such that $|\bar{1}_t\rangle = u_t |01\rangle + v_t |10\rangle$) we have that

$$P_Q(\bar{1}_i, \bar{0}, \bar{0}) = \frac{1}{8}, \quad \text{(C1)}$$

$$P_Q(\bar{1}_i, \bar{1}_j, \bar{1}_k) = \frac{1}{8}(u_i u_j u_k + v_i v_j v_k)^2 \quad \text{(C2)}$$

and similar relations by permuting the parties. We prove by contradiction, assuming the existence of a trilocal model reproducing $P_Q(a, b, c)$, we identify conditions that this model should satisfy, leading to a contradiction for certain choices of the measurement parameter u.

Notice that the measurement can be performed in two steps. In the first step, the parties measure the operators in the basis $\bar{0} = |00\rangle\langle00| + |11\rangle\langle11|$, $\bar{1} = |01\rangle\langle01| + |10\rangle\langle10|$, where $\bar{n}$ is the parity of received tokens, obtaining the distribution P_{parity}. Second, whenever $\bar{1}$ is obtained, the party performs an additional measurement given by $\bar{1}_0$, $\bar{1}_1$, which are the only entangled measurement operators in this setting.

As the operators $\{\bar{0}, \bar{1}\}$, are diagonal in the basis of token numbers, substituting the quantum sources with the sources $\frac{1}{2}(|10\rangle\langle10| + |01\rangle\langle01|)$ has no effect on P_{parity}. This is equivalent to the sources delivering its token to each of the connected parties with probability $\frac{1}{2}$ which is a classical token counting strategy. As a result, any potential classical strategy that simulates P_Q imposes a classical strategy for P_{parity}, which must be the classical parity token counting strategy, due to the almost-rigidity property.

Therefore, any potential classical simulating strategy for P_Q involves a new classical hidden variable which shows the movement of each individual token. For example, considering the case when all parties output $\bar{1}$, the source's tokens must be transmitted in either the clockwise ($t = \circlearrowright$) or in the anticlockwise ($t = \circlearrowleft$) direction, as illustrated in Fig. 5. Therefore, if a trilocal model existed, one should be able to define the following joint probability distribution for $t \in \{\circlearrowright, \circlearrowleft\}$:

$$q(i, j, k, t) = \text{Pr}(a = \bar{1}_i, \ b = \bar{1}_j, \ c = \bar{1}_k, \ t \mid$$

$$\text{all parties output one token})$$

$$= 4 \text{Pr}(a = \bar{1}_i, \ b = \bar{1}_j, \ c = \bar{1}_k, \ t),$$

where the second equality is due to the fact that $\text{Pr}(\text{all parties output one token}) = \frac{1}{4}$ and it includes the event
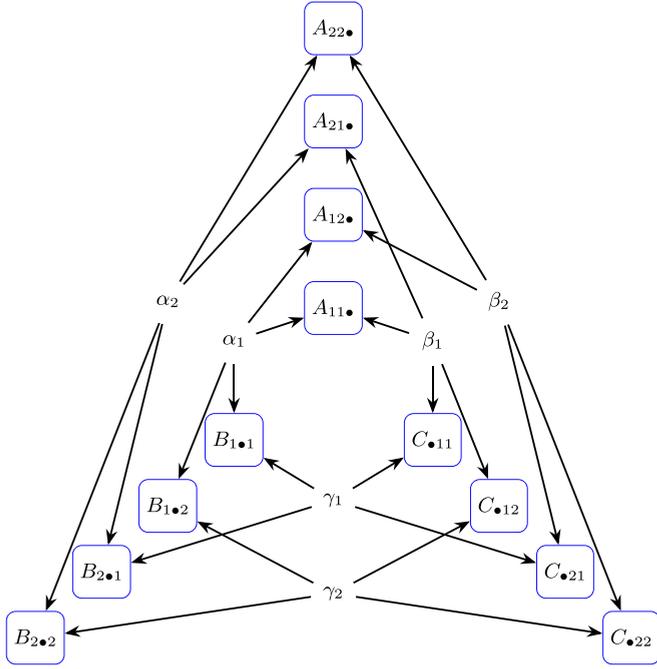
FIG. 6. The inflated triangle scenario where α, β, and γ are the sources and the rest are observable variables. The inflation order is set to be 2. The copies are created such that $A_{\alpha=x,\beta=y,\gamma=z} = A_{xyz}$ where $x, y, z \in \{1, 2\}$ and the symbol $\bullet$ means that the variable in question does not have a dependency on that particular source.

of $\{a = \bar{1}_i, b = \bar{1}_j, c = \bar{1}_k, t\}$. Therefore, we can explicitly compute some of the marginals of $q(i, j, k, t)$, for example, it should satisfy

$$q(i, j, k) = \sum_t q(i, j, k, t) = \frac{1}{2}(u_i u_j u_k + v_i v_j v_k)^2, \quad \text{(C3)}$$

$$q(i, t = \circlearrowright) = \sum_{j,k} q(i, j, k, t = \circlearrowright) = \frac{1}{2} u_i^2, \quad \text{(C4)}$$

and similar constraints on $q(i, t = \circlearrowleft)$, $q(j, t)$, and $q(k, t)$.

Equation (C3) is straightforward from Eq. (C2), and Eq. (C4) can be derived from the fact that Alice's output must be independent of α:

$$q(i, t = \circlearrowright) = 4 \Pr(a = \bar{1}_i, \alpha \mapsto B, \gamma \mapsto A, \beta \mapsto C)$$
$$= 4 \Pr(a = \bar{1}_i, \alpha \mapsto C, \gamma \mapsto A, \beta \mapsto C)$$
$$= 4 \Pr_{\text{RGB4}}(a = \bar{1}_i, b = \bar{0}, c = \bar{2}) = \frac{1}{2} u_i^2,$$

where, for example, $\alpha \mapsto B$ means that the source α, which is shared between B and C, sends its only token to the party B.

Assuming the existence of a trilocal model, one should be able to define a distribution $q(i, j, k, t)$ that is consistent with all of the marginals (C3), (C4). Note that this satisfying marginal constraints problem is indeed a linear program (LP). The authors of Ref. [15] demonstrated that this LP has no solution for some specific choices of the operators $\{\bar{1}_0, \bar{1}_1\}$, notably when $0.785 \approx u_{\max}^2 < u^2 < 1$ which concludes the proof. For completeness, we mention this proof in detail in Appendix D.

A key distinction between classical and quantum token counting strategies is that in the quantum approach, when all parties have received one token, the global state is in a coherent superposition of tokens cycling clockwise and anti-clockwise. However, in the classical method the whole system is either in state of tokens cycling clockwise or anticlockwise with some probability. For example,

$$\text{Classical} : \frac{1}{2}(|\circlearrowright\rangle\langle\circlearrowright| + |\circlearrowleft\rangle\langle\circlearrowleft|),$$

$$\text{Quantum} : \frac{1}{\sqrt{2}}(|\circlearrowright\rangle + |\circlearrowleft\rangle).$$

The creation of a global coherent superposition in the network through local measurements in quantum token counting strategies is ultimately what allows these approaches to create network nonlocality.

APPENDIX D: THERE IS NO $q(i, j, k, t)$ COMPATIBLE WITH THE MARGINALS

Lemma 1. There exists no valid distribution $q(i, j, k, t)$ satisfying the mentioned marginals (C3), (C4) and similar ones on $q(i, t = \circlearrowright)$, $q(j, t)$, and $q(k, t)$, when $0.785 \approx u_{\max}^2 < u^2 < 1$. Meaning that, this linear program has no feasible solution.

Proof. Let us introduce the probability distribution $\tilde{q}$ which is the symmetrization of q over i, j, k:

$$\tilde{q}(i, j, k, t) = \frac{1}{6}[q(i, j, k, t) + q(j, k, i, t) + q(k, i, j, t)$$
$$+ q(i, k, j, t) + q(k, j, i, t) + q(j, i, k, t)].$$

Clearly, $\tilde{q}$ still satisfies the same given constraints as q. Let $\xi_{ijk} := \tilde{q}(i, j, k, t = \circlearrowright) - \tilde{q}(i, j, k, t = \circlearrowleft)$. We have

$$\tilde{q}(i, j, k, t = \circlearrowright) = \frac{1}{2}[\tilde{q}(i, j, k) + \xi_{ijk}],$$
$$\tilde{q}(i, j, k, t = \circlearrowleft) = \frac{1}{2}[\tilde{q}(i, j, k) - \xi_{ijk}].$$

For simplicity, we write $\xi_0 := \xi_{000}$, $\xi_1 := \xi_{100} = \xi_{010} = \xi_{001}$, $\xi_2 := \xi_{110} = \xi_{101} = \xi_{011}$, and $\xi_3 := \xi_{111}$.

Hence, with (C3), we can write the marginals

$$\tilde{q}(k, t = \circlearrowright) = \frac{1}{2} \sum_{i,j} \left(\frac{(u_i u_j u_k + v_i v_j v_k)^2}{2} + \xi_{ijk} \right)$$

$$= \frac{1}{2} \left(\frac{u_k^2 + v_k^2}{2} + \sum_{ij} \xi_{ijk} \right),$$

$$\tilde{q}(k, t = \circlearrowleft) = \frac{1}{2} \sum_{i,j} \left(\frac{(u_i u_j u_k + v_i v_j v_k)^2}{2} - \xi_{ijk} \right)$$

$$= \frac{1}{2} \left(\frac{u_k^2 + v_k^2}{2} - \sum_{ij} \xi_{ijk} \right).$$

As $\sum_{ij} \xi_{ij0} = \xi_0 + 2\xi_1 + \xi_2$, $\sum_{ij} \xi_{ij1} = \xi_1 + 2\xi_2 + \xi_3$, we deduce

$$\xi_0 = \frac{u^2 - v^2}{2} - 2\xi_1 - \xi_2 = u^2 - \frac{1}{2} - 2\xi_1 - \xi_2,$$

$$\xi_3 = \frac{v^2 - u^2}{2} - \xi_1 - 2\xi_2 = \frac{1}{2} - u^2 - \xi_1 - 2\xi_2.$$

We have the positivity conditions

$$0 \leqslant \tilde{q}(0,0,0,\circlearrowleft) = \frac{(u^3+v^3)^2}{4} - \frac{\xi_0}{2},$$

$$0 \leqslant \tilde{q}(1,1,1,\circlearrowright) = \frac{(v^3-u^3)^2}{4} + \frac{\xi_3}{2},$$

hence,

$$\xi_2 \geqslant u^2 - \frac{1+(u^3+v^3)^2}{2} - 2\xi_1,$$

$$\xi_2 \leqslant \frac{1+(v^3-u^3)^2}{4} - \frac{u^2}{2} - \frac{\xi_1}{2},$$

leading to

$$\xi_1 \geqslant u^2 - \frac{2[1+(u^3+v^3)^2]+1+(v^3-u^3)^2}{6}.$$

Finally, we use the positivity condition

$$0 \leqslant \tilde{q}(0,0,1,\circlearrowright) = \frac{(u^2v-uv^2)^2}{4} - \frac{\xi_1}{2}$$

to get

$$3(u^2v-uv^2)^2-6u^2+2[1+(u^3+v^3)^2]+1+(v^3-u^3)^2 \geqslant 0.$$

For $v = \sqrt{(1-u^2)}$ and $v < u$, this inequality leads us to $u_{max}^2 \geqslant u^2$ with $u_{max}^2 = \frac{-3+(9+6\sqrt{2})^{2/3}}{2(9+6\sqrt{3})^{1/3}} \approx 0.785$, which leads us to contradiction for $u_{max}^2 < u^2 < 1$. ∎

APPENDIX E: PROOF OF NONLOCALITY BY INFLATION

To prove the nonlocality of several variations of the coarse-grained Fritz distribution and to corroborate the results of the neural network a Web inflation of the triangle scenario was used [6]. Each source was inflated up to two copies (see Fig. 6).

The following two maximal expressible sets were used to run the inflation linear program. Due to the factorization of these probabilities, the inequalities obtained from the dual solution would contain the corresponding quadratic and cubic elements:

$$P(A_{11\bullet}A_{22\bullet}B_{1\bullet1}B_{2\bullet2}C_{\bullet11}C_{\bullet22}) = P(A_{11\bullet}B_{1\bullet1}C_{\bullet11})P(A_{22\bullet}B_{2\bullet2}C_{\bullet22}), \tag{E1}$$

$$P(A_{12\bullet}B_{2\bullet1}C_{\bullet12}) = P(A_{12\bullet})P(B_{2\bullet1})P(C_{\bullet12}). \tag{E2}$$

The inflation linear program showed a varying sensitivity for different combinations of Alice's and Bob's outputs in detecting nonlocality. Therefore, in some cases using only (E1) was sufficient. Following is a table that presents the results of the inflation linear program for some coarse-grained Fritz distributions:

A	B	C	Noise tolerance		A	B	C	Noise tolerance	
0[00 01]	**0**[00 **1**[01	**0**[00 01]	≈0.87		**0**[00 01]	**0**[00 **1**[01	**0**[-00 -01 -10	≈0.87	(E3)
1[10 **2**[11	**2**[10 11	**1**[10 **2**[11			**1**[10 **2**[11	**2**[10 11	**1**[11		
A	B	C	Noise tolerance		A	B	C	Noise tolerance	
0[00 01]	**0**[00 01]	**0**[00 01]	≈0.81		**0**[00 01]	**0**[00 01]	**0**[-00 -01 -10	≈0.97	(E4)
1[10 **2**[11	**1**[10 **2**[11	**1**[10 **2**[11			**1**[10 **2**[11	**1**[10 **2**[11	**1**[11		
A	B	C	Noise tolerance		A	B	C	Noise tolerance	
0[00 01]	**0**[00 **1**[01 10	**0**[00 01]	≈0.85		**0**[00 01]	**0**[00 **1**[01 10	**0**[-00 -01 -10	≈0.96	(E5)
1[10 **2**[11	**2**[11	**1**[10 **2**[11			**1**[10 **2**[11	**2**[11	**1**[11		

The boldface numbers designate the new, merged outputs and A, B, C are Alice, Bob, and Charlie, respectively, and we use the notation $a'x$ for the outputs of A, $b'y$ for B, and xy for C. In order to obtain the noise tolerances in (E3) and (E4) both expressible sets (E1) and (E2) were implemented in the linear program, however, in (E5) only the expressible set in (E1) was used. Additionally, to obtain the noise tolerance for both of the distributions in (E4) noncertificate-type constraints were implemented in the linear program. These equality constraints are of the form

$$P(A_{11\bullet}A_{22\bullet}B_{2\bullet1}B_{2\bullet2}C_{\bullet21}C_{\bullet22}) = \underbrace{P(A_{11\bullet})}_{\text{injectable}} \underbrace{P(A_{22\bullet}B_{2\bullet1}B_{2\bullet2}C_{\bullet21}C_{\bullet22})}_{\text{not injectable}}.$$

In total there are three such semiknowable (semi-injectable) sets with the factorized injectable parts [6] being $P(B_{1\bullet 1})$ and $P(C_{\bullet 11})$ for the other two semiknowable sets, respectively. Merging Alice's first two outputs, doing the same for Bob's middle two outputs and Charlie's first three outputs as depicted in the right side of (E5) the following inequality was obtained from the linear program:

$$P(000)[P(020) + P(101) + P(111) + P(120) + P(201) + P(211) + P(220)]P(001)[P(020) + P(021) + P(120) + P(121)$$
$$+ P(220) + P(221)] + P(101)[P(110) + P(120) + P(121) + P(200) + P(210) + P(220)] + P(010)[P(101) + P(201)]$$
$$+ P(020)[P(101) + P(201)] + P(021)[P(101) + P(201)] + P(100)[P(101) + P(111)] + P(110)[P(210) + P(220)]$$
$$+ P(201)[P(220) + P(221)] + P(120)[P(200) + P(201) + P(210) + P(220)] + P(121)[P(200) + P(201)$$
$$+ P(210) + P(220)] + P(111)[P(200) + P(210) + P(220)] + P(200)P(220) - P(100)P(221) \geqslant 0,$$

where $P(ABC) = P(ijk)$ with $i, j \in \{0, 1, 2\}$ and $k \in \{0, 1\}$. However, since the distribution at the right side of (E5) was obtained by merging outcomes in the original 4-output Fritz distribution, many of the probabilities in this inequality are zero. Therefore, the inequality violated by this distribution was in fact

$$P(000)[P(020) + P(120) + P(211)] + P(120)[P(200) + P(210)] + P(110)P(210) - P(100)P(221) \geqslant 0.$$

The violation of this inequality by the noiseless distribution (E5) is $\approx -1.6 \times 10^{-3}$.

---

[1] A. Tavakoli, A. Pozas-Kerstjens, M.-X. Luo, and M.-O. Renou, Bell nonlocality in networks, Rep. Prog. Phys. **85**, 056001 (2022).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[3] C. Branciard, N. Gisin, and S. Pironio, Characterizing the Nonlocal Correlations Created via Entanglement Swapping, Phys. Rev. Lett. **104**, 170401 (2010).

[4] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, Bilocal versus nonbilocal correlations in entanglement-swapping experiments, Phys. Rev. A **85**, 032119 (2012) .

[5] T. Fritz, Beyond Bell's theorem: Correlation scenarios, New J. Phys. **14**, 103001 (2012).

[6] E. Wolfe, R. W. Spekkens, and T. Fritz, The inflation technique for causal inference with latent variables, J. Causal Inference **7**, 20170020 (2019).

[7] M. Navascués, E. Wolfe, D. Rosset, and A. Pozas-Kerstjens, Genuine Network Multipartite Entanglement, Phys. Rev. Lett. **125**, 240505 (2020).

[8] E. Wolfe, A. Pozas-Kerstjens, M. Grinberg, D. Rosset, A. Acín, and M. Navascues, Quantum Inflation: A General Approach to Quantum Causal Compatibility, Phys. Rev. X **11**, 021043 (2021).

[9] J. Åberg, R. Nery, C. Duarte, and R. Chaves, Semidefinite Tests for Quantum Network Topologies, Phys. Rev. Lett. **125**, 110505 (2020).

[10] M. Weilenmann and R. Colbeck, Non-shannon inequalities in the entropy vector approach to causal structures, Quantum **2**, 57 (2018).

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

[12] M.-O. Renou, Y. Wang, S. Boreiri, S. Beigi, N. Gisin, and N. Brunner, Limits on Correlations in Networks for Quantum and No-Signaling Resources, Phys. Rev. Lett. **123**, 070403 (2019).

[13] N. Gisin, J.-D. Bancal, Y. Cai, P. Remy, A. Tavakoli, E. Z. Cruzeiro, S. Popescu, and N. Brunner, Constraints on nonlocality in networks from no-signaling and independence, Nat. Commun. **11**, 2378 (2020).

[14] T. C. Fraser and E. Wolfe, Causal compatibility inequalities admitting quantum violations in the triangle structure, Phys. Rev. A **98**, 022113 (2018).

[15] M.-O. Renou, E. Bäumer, S. Boreiri, N. Brunner, N. Gisin, and S. Beigi, Genuine quantum nonlocality in the triangle network, Phys. Rev. Lett. **123**, 140401 (2019).

[16] N. Gisin, Entanglement 25 years after quantum teleportation: Testing joint measurements in quantum networks, Entropy **21**, 325 (2019).

[17] M.-O. Renou and S. Beigi, Nonlocality for Generic Networks, Phys. Rev. Lett. **128**, 060401 (2022).

[18] P. Abiuso, T. Kriváchy, E.-C. Boghiu, M.-O. Renou, A. Pozas-Kerstjens, and A. Acín, Single-photon nonlocality in quantum networks, Phys. Rev. Res. **4**, L012041 (2022).

[19] A. Pozas-Kerstjens, N. Gisin, and M.-O. Renou, Proofs of Network Quantum Nonlocality in Continuous Families of Distributions, Phys. Rev. Lett. **130**, 090201 (2023) .

[20] I. Šupić, J.-D. Bancal, Y. Cai, and N. Brunner, Genuine network quantum nonlocality and self-testing, Phys. Rev. A **105**, 022206 (2022).

[21] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[22] T. Kriváchy, Y. Cai, D. Cavalcanti, A. Tavakoli, N. Gisin, and N. Brunner, A neural network oracle for quantum nonlocality problems in networks, npj Quantum Inf. **6**, 70 (2020).

[23] M.-O. Renou and S. Beigi, Network nonlocality via rigidity of token counting and color matching, Phys. Rev. A **105**, 022408 (2022).

[24] S. Boreiri *et al.* (unpublished).

[25] E. Wolfe (private communication).

[26] K. He, L. Li, X. Liu, Y. Wang, and M. Xia, Variable-version lovász local lemma: Beyond shearer's bound, in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, NJ, 2017), pp. 451–462.

[27] A. Restivo, D. Rosset, and N. Brunner (unpublished).

[28] A. Pozas-Kerstjens, N. Gisin, and A. Tavakoli, Full Network Nonlocality, Phys. Rev. Lett. **128**, 010403 (2022).