

Received 26 October 2022; revised 9 January 2023; accepted 2 February 2023; date of publication 9 February 2023;
date of current version 14 March 2023.

Digital Object Identifier 10.1109/TQE.2023.3243849

Variational Quantum Optimization of Nonlocality in Noisy Quantum Networks

BRIAN DOOLITTLE¹ , THOMAS R. BROMLEY², NATHAN KILLORAN²,
AND ERIC CHITAMBAR³ 

¹Department of Physics, Loomis Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

²Xanadu, Toronto, ON M5G 2C8, Canada

³Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

Corresponding authors: Brian Doolittle; Eric Chitambar (e-mail: briand4@illinois.edu; echitamb@illinois.edu).

This work was supported by the National Science Foundation under Award DMR-1747426 and Award 2016136.

ABSTRACT The noise and complexity inherent to quantum communication networks leads to technical challenges in designing quantum network protocols using classical methods. We address this issue with a hybrid variational quantum optimization (VQO) framework that simulates quantum networks on quantum hardware and optimizes the simulation using differential programming. We maximize nonlocality in noisy quantum networks to showcase our VQO framework. Using a classical simulator, we investigate the noise robustness of quantum nonlocality. Our VQO methods reproduce known results and uncover novel phenomena. We find that maximally entangled states maximize nonlocality in the presence of unital qubit channels, while nonmaximally entangled states can maximize nonlocality in the presence of nonunital qubit channels. Thus, we show VQO to be a practical design tool for quantum networks even when run on a classical simulator. Finally, using IBM quantum computers, we demonstrate that our VQO framework can maximize nonlocality on noisy quantum hardware. In the long term, our VQO techniques show promise of scaling beyond classical approaches and can be deployed on quantum network hardware to optimize network protocols against their inherent noise.

INDEX TERMS Design and simulation tools, noisy intermediate-scale quantum (NISQ) algorithms and devices, quantum networking.

I. INTRODUCTION

Quantum information science and technology are progressing rapidly toward the quantum internet [1], [2], [3], [4], a global communication network of quantum processing devices linked by quantum communication. Quantum communication networks are predicted to revolutionize science and technology by providing advantages in distributed sensing [5], [6], [7], communications [8], [9], [10], network security [11], [12], [13], [14], [15], and distributed information processing [16], [17]. Unfortunately, the presence of noise in existing quantum hardware prevents the implementation of these technologies at the network scale. Nevertheless, we are at the forefront of the quantum internet because rudimentary quantum networks can be built using existing technology, and as new technologies emerge, these networks can be upgraded and scaled [3].

As quantum networks scale, technical challenges will emerge in their characterization, simulation, and optimization. A network protocol must be designed to be robust

against the noise present in quantum network devices. Unfortunately, the difficulty of characterizing quantum noise grows exponentially with the number of qubits [18], [19], [20]. Furthermore, simulating and optimizing large quantum systems is challenging. While tensor networks [21] are a promising classical tool for these tasks, they may fail to efficiently simulate or optimize networks that have complex entanglement structures or devices that have many qubits. Designing quantum networks will require characterization, simulation, and optimization of large quantum systems, but how can we develop the quantum internet if classical design tools do not scale?

Quantum problems often have quantum solutions. A solution with promising advantages in the quantum network design is variational quantum optimization (VQO) [22], [23], [24]. This hybrid algorithm combines the power of quantum simulation with classical differential programming. Hybrid, quantum-classical algorithms have demonstrated success across a wide range of simulation and optimization

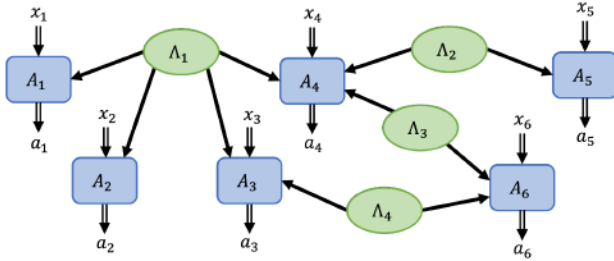


FIGURE 1. Example quantum network in the n -local setting. A quantum network can be represented by a directed acyclic graph. Sources (green ovals) correlate the nodes (blue rectangles). The double-line arrows depict the classical inputs x_j and outputs a_j . The solid arrows depict quantum communication between sources and nodes.

problems [24], [25], [26], and are predicted to yield practical advantages on noisy intermediate-scale quantum (NISQ) devices [27]. Furthermore, hybrid optimization techniques have been shown to be adaptive to hardware noise when used to maximize nonlocality in simple photonic quantum networks [28], [29].

To demonstrate the utility of VQO in a quantum network design, we apply it to maximizing nonlocal correlations in the n -local setting where quantum network devices are linked only by entanglement (see Fig. 1). These quantum networks are an essential step toward the quantum internet [1], [2], [3], [4] because they enable entanglement distribution [30], [31], [32] and long-distance quantum communication over repeater chains [33], [34]. The entanglement in the n -local setting can be used to create stronger-than-classical correlations between network devices [35], [36], [37], [38], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50]. These non- n -local correlations are important for certifying quantum devices [51], [52], [53], [54], [55], securely running protocols on uncharacterized devices [13], [15], [56], [57], [58], [59], [60], [61], and distributed information processing [16].

Unfortunately, non- n -locality deteriorates in the presence of noise [62], [63], [64], making it difficult to preserve in physical systems. To build quantum networks that are robust to noise, it is crucial to understand how noise affects non- n -local correlations. Classical machine learning has shown success in solving complex optimization problems that arise in the study of nonlocality [65], [66], [67]. While VQO techniques share many similarities with classical machine learning, our approach is distinct from the numerical approaches considered in previous works. To demonstrate the utility of VQO, we apply it to the practical network design problem of maximizing nonlocality in noisy quantum networks.

The rest of this article is organized as follows. In Section II, we give an overview of quantum networks in the n -local setting. In Section III, we describe our VQO framework for noisy quantum networks. In Section IV, we discuss quantum non- n -locality, its noise robustness, and how it can be maximized using VQO. In Section V, we apply VQO on a classical simulator to verify that it can successfully

maximize non- n -locality in the presence of noise and report upon interesting theoretical findings. In Section VI, we use VQO to maximize non- n -locality on noisy IBM quantum hardware. In summary, we use our VQO framework to reproduce known nonlocality results, obtain new nonlocality results, and show its potential promise of practical advantage on quantum hardware.

We make our techniques accessible, transparent, and reproducible through open-source software. Our framework for VQO of quantum networks is built upon PennyLane [68] and released as a Python package called qNetVO [69]. Our data and numerics are available on GitHub in a supplementary codebase [70].

II. NETWORKS IN THE N -LOCAL SETTING

In this section, we outline the theoretical model of quantum and classical networks in the n -local setting. The experienced reader may proceed to Section III, where our VQO framework is discussed in detail.

A. CLASSICAL NETWORKS

A classical network in the n -local setting consists of n independent sources $\Lambda_1, \dots, \Lambda_n$ that distribute randomness to m nonsignaling devices $A_1, \dots, A_m$ (see Fig. 1). The i th source outputs classical value μ_i drawn randomly from the distribution Ω^{Λ_i} with probability $P(\mu_i)$ and sends μ_i to all linked devices. All sources in the network are assumed to be independent such that $P(\vec{\mu}) = \prod_{i=1}^n P(\mu_i)$, where $\vec{\mu} = (\mu_i)_{i=1}^n$ contains the random value output from each independent source. The sources $\{\Lambda_i\}_{i=1}^n$ each distribute their random value to the nodes $\{A_j\}_{j=1}^m$ through a collection of links $\{L_k\}_{k=1}^l$. The node A_j may receive multiple random values; thus, we denote with $\vec{\mu}^{A_j} \subseteq \vec{\mu}$ the set of random values received by node A_j .

The j th node has classical input and output alphabets $\mathcal{X}_j := \{1, \dots, |\mathcal{X}_j|\}$ and $\mathcal{A}_j := \{1, \dots, |\mathcal{A}_j|\}$, respectively, where the input and output alphabets for the entire network are denoted $\mathcal{X} := \mathcal{X}_1 \times \dots \times \mathcal{X}_m$ and $\mathcal{A} := \mathcal{A}_1 \times \dots \times \mathcal{A}_m$, respectively. Hence, the network processes the classical input $\vec{x} \in \mathcal{X}$ to produce the classical output $\vec{a} \in \mathcal{A}$, where $\vec{x} = (x_j \in \mathcal{X}_j)_{j=1}^m$ and $\vec{a} = (a_j \in \mathcal{A}_j)_{j=1}^m$. Since, the nodes are nonsignaling, their joint probability distribution must satisfy

$$P(\vec{a}|\vec{x}, \vec{\mu}) = \prod_{j=1}^m P(a_j|x_j, \vec{\mu}^{A_j}) \quad (1)$$

for all $\vec{x} \in \mathcal{X}$, $\vec{a} \in \mathcal{A}$, and $\vec{\mu} \in \{\Omega^{\Lambda_i}\}_{i=1}^n$.

We characterize networks using only their input–output statistics. Hence, we consider a scenario where many identical and independent experiments are performed. In each experiment, a classical input $\vec{x} \in \mathcal{X}$ is drawn from a uniform random distribution. The network processes the input $\vec{x}$ to produce the output $\vec{a} \in \mathcal{A}$. After many repetitions, an approximate conditional probability distribution $\{P(\vec{a}|\vec{x})\}_{\vec{a} \in \mathcal{A}, \vec{x} \in \mathcal{X}}$ is constructed. These conditional probabilities fully characterize the network and are represented as a column stochastic

matrix referred to as the *network behavior*

$$\mathbf{P} = \sum_{\vec{a} \in \mathcal{A}} \sum_{\vec{x} \in \mathcal{X}} P(\vec{a}|\vec{x}) |\vec{a}\rangle \langle \vec{x}| \quad (2)$$

where $\{|\vec{x}\rangle\}_{\vec{x} \in \mathcal{X}}$ and $\{|\vec{a}\rangle\}_{\vec{a} \in \mathcal{A}}$ form classical orthonormal bases over the input and output sets, respectively. The transition probabilities $P(\vec{a}|\vec{x})$ decompose as [43]

$$P(\vec{a}|\vec{x}) = \sum_{\mu_1 \in \Omega^{\Lambda_1}} \cdots \sum_{\mu_n \in \Omega^{\Lambda_n}} \prod_{i=1}^n P(\mu_i) \prod_{j=1}^m P(a_j|x_j, \vec{\mu}^{A_j}). \quad (3)$$

B. QUANTUM NETWORKS

In the quantum setting, each source prepares a quantum state and sends it to the linked nodes. Each node receives and measures the quantum state sent from each of its linked sources. The node outputs a classical value indicating the result of the measurement.

This work models a quantum network in the n -local setting as an N -qubit system where each qubit is indexed by an integer $q_i \in [N]$. An M -qubit subsystem with $M < N$ is referenced using the sequence $(q_i)_{i=1}^M \subset [N]$. Sources Λ_i , nodes A_j , and links L_k are uniquely described by their local qubits, $\Lambda_i, A_j, L_k \subset [N]$. Qubits cannot be shared between multiple sources, links, or nodes; hence, $\Lambda_i \cap \Lambda_{i'} = \emptyset \ \forall i \neq i'$ and similarly for A_j and L_k .

In the quantum setting, the n sources collectively prepare the state $|\psi^{\text{Net}}\rangle = \bigotimes_{i=1}^n |\psi^{\Lambda_i}\rangle$, where $|\psi^{\text{Net}}\rangle \in \mathcal{H}_{\text{Prep}}^{\text{Net}}$. We define $\mathcal{H}_{\text{Prep}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{H}_{\text{Prep}}^{\Lambda_i}$ as the joint Hilbert space where source independence is ensured by the separability across states. For mixed states, we denote the density operator as $\rho^{\Lambda_i} \in D(\mathcal{H}_{\text{Prep}}^{\Lambda_i})$.

Quantum channels model the noisy link between source and measurement devices in the network. A quantum channel is represented by the completely-positive trace-preserving (CPTP) map [71] $\mathcal{N}^{L_k} : D(\mathcal{H}_{\text{Prep}}^{L_k}) \rightarrow D(\mathcal{H}_{\text{Meas}}^{L_k})$, where the L_k denotes the qubits on which the channel acts. For convenience, the input and output Hilbert spaces have equal dimensions. Independent quantum channels combine to describe the network noise as $\mathcal{N}^{\text{Net}} = \bigotimes_{k=1}^l \mathcal{N}^{L_k}$, where, in the noiseless case, $\mathcal{N}^{\text{Net}}(\rho^{\text{Net}}) = \text{id}(\rho^{\text{Net}}) = \rho^{\text{Net}}$. Note that noise can also be applied to the source or measurement devices.

The measurement occurs at node A_j and is modeled as a projection-valued measure (PVM) $\{\Pi_{a_j|x_j}^{A_j}\}_{a_j \in A_j}$ that forms a set of orthogonal projectors satisfying $\sum_{a_j \in A_j} \Pi_{a_j|x_j}^{A_j} = \mathbb{I}^{A_j}$. The measurement device A_j receives and measures the qubits $\rho^{A_j} \in D(\mathcal{H}_{\text{Meas}}^{A_j})$. In aggregate, the network applies the projector $\Pi_{\vec{a}|\vec{x}}^{\text{Net}} = \bigotimes_{j=1}^m \Pi_{a_j|x_j}^{A_j}$, where the PVM applied at each node is conditioned upon the classical input $x_j \in \mathcal{X}_k$. Upon measurement, the classical output $\vec{a}$ is obtained with probability

$$P(\vec{a}|\vec{x}) = \text{Tr} \left[\Pi_{\vec{a}|\vec{x}}^{\text{Net}} \mathcal{N}^{\text{Net}}(\rho^{\text{Net}}) \right] \quad (4)$$

where any permutations needed to map $\mathcal{H}_{\text{Prep}}^{\text{Net}}$ to $\mathcal{H}_{\text{Meas}}^{\text{Net}}$ are included implicitly.

III. VQO OF NOISY QUANTUM NETWORKS

This section introduces our VQO framework for quantum networks. At a high level, our framework simulates the quantum network using a tunable quantum circuit. The quantum circuit parameters are optimized using differential programming techniques. Our VQO framework is implemented as a Python package called qNetVO: the Quantum Network Variational Optimizer [69]. The qNetVO software is built upon PennyLane, which is a free and open-source quantum differential programming framework [68]. PennyLane enables our VQO framework to be easily run on a wide range of quantum devices and classical simulators.

A. SIMULATING NOISY QUANTUM NETWORKS

A quantum network can be simulated on a quantum computer. In the simulation, quantum circuits model the state preparations, the communication, and the measurements of network devices. A quantum network simulation is constructed modularly by combining the quantum circuit models for each source, receiver, and noisy link. The quantum network circuit is then run on a quantum computer to produce statistics that are, ideally, indistinguishable from the simulated network.

Formally, the quantum network circuit is represented by the unitary operator $U_{\vec{x}}^{\text{Net}}$, where the subscript $\vec{x} \in \mathcal{X}$ is the classical input on which the network's dynamics are conditioned. The unitary $U_{\vec{x}}^{\text{Net}}$ acts upon the N -qubit zero state $|0\rangle^N$ and is measured in the computational basis $\{|\vec{z}\rangle\}_{\vec{z} \in \mathcal{Z}}$, where $\mathcal{Z} := \{0, 1\}^N$ is the set of all N -bit strings. When a fault-tolerant quantum computer executes the ansatz circuit, the bit string $\vec{z}$ is output with probability

$$P(\vec{z}|\vec{x}) = \left| \langle \vec{z} | U_{\vec{x}}^{\text{Net}} | 0 \rangle^N \right|^2. \quad (5)$$

Using (2) and (5), we express the parameterized quantum circuit behavior $\mathbf{P}_{\text{QC}}$ as the column stochastic matrix

$$\mathbf{P}_{\text{QC}} = \sum_{\vec{z} \in \mathcal{Z}} \sum_{\vec{x} \in \mathcal{X}} \left| \langle \vec{z} | U_{\vec{x}}^{\text{Net}} | 0 \rangle^N \right|^2 |\vec{z}\rangle \langle \vec{x}|. \quad (6)$$

If the simulated quantum network does not output an N -bit string, then a classical postprocessing map $\mathbf{L} : \mathcal{Z} \rightarrow \mathcal{A}$ is needed to map the $|\mathcal{Z}|$ outputs of the quantum circuit to the $|\mathcal{A}|$ outputs of the quantum network. That is, the quantum network behavior $\mathbf{P}_{\text{Net}}$ is constructed as

$$\mathbf{P}_{\text{Net}} = \mathbf{L} \mathbf{P}_{\text{QC}} \quad (7)$$

where $\mathbf{P}_{\text{QC}}$ is defined in (6), and the postprocessing map is represented as a column stochastic matrix

$$\mathbf{L} = \sum_{\vec{a} \in \mathcal{A}} \sum_{\vec{z} \in \mathcal{Z}} P(\vec{a}|\vec{z}) |\vec{a}\rangle \langle \vec{z}|. \quad (8)$$

On a quantum computer, the network behavior $\mathbf{P}_{\text{Net}}$ is obtained by repeatedly executing the ansatz circuit $U_{\vec{x}}^{\text{Net}}$ across

all inputs $\vec{x} \in \mathcal{X}$ to estimate the probabilities $P(\vec{z}|\vec{x})$ for all inputs. Then, the postprocessing map $\mathbf{L}$ is applied to the quantum circuit probabilities.

The network simulation circuit can be parameterized by a set of real-valued settings $\Theta_{\vec{x}}$ as $U_{\vec{x}}^{\text{Net}} = U^{\text{Net}}(\Theta_{\vec{x}})$. Each scalar value $\theta \in \Theta_{\vec{x}}$ is continuous, differentiable, and controls dynamics of the network simulation circuit $U_{\vec{x}}^{\text{Net}}$. For each input $\vec{x} \in \mathcal{X}$, the collection of settings $\Theta_{\vec{x}}$ may be unique; hence, we can parameterize the network behavior $\Theta = \{\Theta_{\vec{x}}\}_{\vec{x} \in \mathcal{X}}$, and it follows from (5) that

$$P_{\text{QC}}(\Theta) = \sum_{\vec{z} \in \mathcal{Z}} \sum_{\vec{x} \in \mathcal{X}} \left| \langle \vec{z} | U^{\text{Net}}(\Theta_{\vec{x}}) | 0 \rangle^N \right|^2 |\vec{z}\rangle \langle \vec{x}|. \quad (9)$$

While this work focuses on networks in the n -local setting, we remark that the simulation framework can apply more generally. Hence, our approach and the tools provided by the qNetVO software [69] can be applied beyond the scope of this work.

1) SIMULATING NOISELESS N -LOCAL QUANTUM NETWORKS

A noiseless quantum network ansatz decomposes into preparation and measurement layers

$$U^{\text{Net}}(\Theta_{\vec{x}}) = U^{\text{Meas}}(\vec{\theta}_{\vec{x}}) U^{\text{Prep}}(\vec{\phi}) \quad (10)$$

where each layer is modularized as

$$U^{\text{Prep}}(\vec{\phi}) = \bigotimes_{i=1}^n U^{\Lambda_i}(\vec{\phi}_i) \quad (11)$$

$$U^{\text{Meas}}(\vec{\theta}_{\vec{x}}) = \bigotimes_{j=1}^m U^{A_j}(\vec{\theta}_{x_j}) \quad (12)$$

and the circuit parameters are organized as $\Theta_{\vec{x}} = \{\vec{\phi}, \vec{\theta}_{\vec{x}}\}$, where

$$\vec{\phi} = \left(\vec{\phi}_i \right)_{i=1}^n \quad \vec{\theta}_{\vec{x}} = \left(\vec{\theta}_{x_j} \right)_{j=1}^m \quad (13)$$

parameterize the state preparations and measurements, respectively. The network state preparation and measurements are expressed as

$$|\psi^{\text{Net}}\rangle = U^{\text{Prep}}(\vec{\phi}) |0\rangle^N \quad (14)$$

$$\Pi_{\vec{z}|\vec{x}}^{\text{Net}} = \left(U^{\text{Meas}}(\vec{\theta}_{\vec{x}}) \right)^\dagger |\vec{z}\rangle \langle \vec{z}| U^{\text{Meas}}(\vec{\theta}_{\vec{x}}). \quad (15)$$

Combining (14) and (15) with (5) yields

$$P(\vec{z}|\vec{y}) = \langle \psi^{\text{Net}} | \Pi_{\vec{z}|\vec{x}}^{\text{Net}} | \psi^{\text{Net}} \rangle. \quad (16)$$

In Table 2 of Appendix B, we provide a list of preparation and measurement ansatzes used throughout this work.

2) SIMULATING NOISY QUANTUM NETWORKS

In a quantum circuit, the nonunitary dynamics of a noisy channel $\mathcal{N}$ are simulated using unitary evolution. By Stinespring's theorem [72], any quantum channel can be expressed in the system-environment representation [71], [72]

$$\mathcal{N}(\rho) = \text{Tr}_E \left[U_{\mathcal{N}}(\rho \otimes |0 \dots 0\rangle \langle 0 \dots 0|^E) U_{\mathcal{N}}^\dagger \right] \quad (17)$$

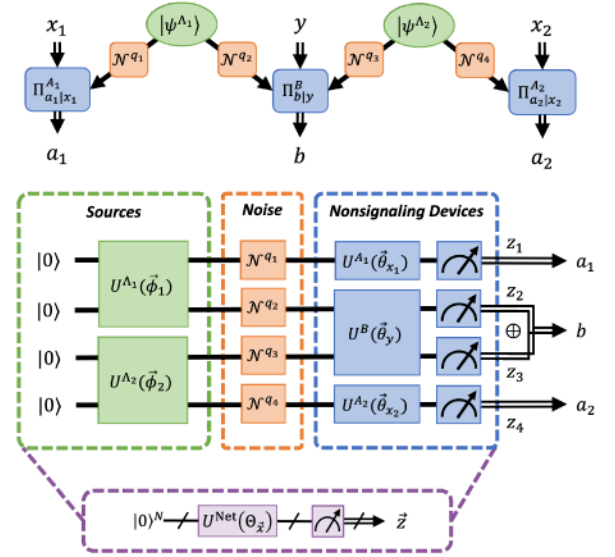


FIGURE 2. Noisy network ansatz circuit. (Top) Quantum network with two sources (green), communication noise (orange), and PVM measurements (blue). (Bottom) Three-layer quantum circuit ansatz that simulates the network. In the source layer, the quantum state $|\psi^{A_i}\rangle$ is prepared. In the noise layer, a static noise model $\mathcal{N}^{q_k}$ is applied. In the measurement layer, the projector $\Pi_{z_j|x_j}^{A_j}$ is a measurement in the computational basis. The ansatz circuit is combined into a single unitary $U^{\text{Net}}(\Theta_{\vec{x}})$ and the outcome probabilities are calculated using (5). When a measurement includes multiple qubits, an XOR operation is used to map the raw output to a single bit, e.g., $b = z_2 \oplus z_3$.

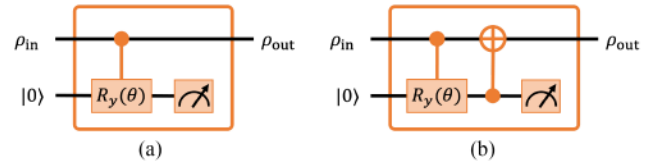


FIGURE 3. Qubit dephasing and amplitude damping circuits [71]. (a) Qubit dephasing channel $\mathcal{P}_\gamma$ is implemented using one ancillary qubit and a controlled rotation about the y -axis. (b) Qubit amplitude damping channel $\mathcal{A}_\gamma$ is implemented using one ancillary qubit and a controlled rotation about the y -axis followed by a CNOT gate. In both circuits, the rotation parameter θ relates to the noise parameter γ as $\theta = 2 \sin^{-1}(\sqrt{\gamma})$. (a) Dephasing. (b) Amplitude damping.

where $E \subseteq [N']$ is an ancillary set of qubits that represent the environment, $\text{Tr}_E[\cdot]$ denotes the partial trace over the environment, and $U_{\mathcal{N}}$ is a unitary applied to both system and environment. Hence, noise can be implemented in a quantum circuit as a unitary gate $U_{\mathcal{N}}$ applied across the system S and an ancillary environment E (see Fig. 3). The Hilbert space dimension of the environment is bounded as $d_E \leq d_S^2$, where d_S is the dimension of the system. Thus, a simulation of a noisy N -qubit network can require up to $2N$ ancillary qubits.

In some cases, it may be difficult to implement the unitary $U_{\mathcal{N}}$ as a quantum circuit. An alternative approach to model the noise using the operator-sum representation of a quantum channel is [71], [73]

$$\mathcal{N}(\rho) = \sum_i K_i \rho K_i^\dagger, \text{ where } \sum_i K_i^\dagger K_i = \mathbb{I} \quad (18)$$

where the CPTP map $\mathcal{N}$ is expressed in terms of its Kraus operators $\{K_i\}_i$. The noise is simulated without using ancillary qubits by directly applying the Kraus operators to the network state as $\mathcal{N}^{\text{Net}}(\rho^{\text{Net}}) = \sum_k K_k \rho^{\text{Net}} K_k^\dagger$. The drawback is that quantum hardware typically does not support nonunitary gates such as Kraus operators. In practice, simulating noise using the operator-sum representation requires the use of a classical simulator such as PennyLane's "default.mixed" mixed state simulator [68].

Certain detector errors can be modeled using a classical postprocessing map $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{A}'$ that acts upon the network behavior as $\mathbf{P}'_{\text{Net}} = \mathbf{E}^{\text{Net}} \mathbf{P}_{\text{Net}}$. Here, $\mathbf{E} = \sum_{\vec{a}', \vec{a}} P(\vec{a}' | \vec{a}) |\vec{a}'\rangle\langle\vec{a}|$ is a column stochastic matrix whose elements $P(\vec{a}' | \vec{a})$ describe the probability that the ideal network output $\vec{a}$ transitions to the output $\vec{a}'$. This noise model describes unflagged detector errors that occur without the experimenter's knowledge.

It is important to distinguish the noise models in our quantum network simulation from the noise present in quantum hardware. In the simulation, our objective is to precisely reproduce the noise present in a quantum network. However, evaluating a simulation circuit on noisy quantum hardware adds more noise into the simulation. Thus, to achieve precise noise modeling, we must rely upon classical simulators or fault-tolerant quantum computers. If run on noisy quantum hardware, careful consideration must be made regarding how the hardware's noise affects the simulation.

B. OPTIMIZING QUANTUM NETWORKS

The goal of our VQO framework is to find the optimal settings Θ^* that yield a network behavior $\mathbf{P}_{\text{Net}}(\Theta^*)$ that is optimal for a particular task, e.g., violating a Bell inequality. We define a problem-specific cost function $\text{Cost}(\mathbf{P}_{\text{Net}}(\Theta))$ that quantifies the network's performance at this task. The optimization objective is then expressed as a minimization of the cost function

$$\Theta^* = \arg \min_{\Theta} \text{Cost}(\mathbf{P}_{\text{Net}}(\Theta)). \quad (19)$$

The cost function can quantify a wide range of network properties such as entropic quantities, the distance to a desired network behavior, or the winning probability of a multipartite game.

We solve the optimization problem in (19) using gradient descent [74], [75] to find local optima of the cost function by traversing the path of steepest descent. Formally, gradient descent is an iterative procedure where, in each step, the settings Θ are updated as

$$\Theta' = \Theta - \eta \nabla_{\Theta} \text{Cost}(\mathbf{P}_{\text{Net}}(\Theta)) \quad (20)$$

where $\eta \in \mathbb{R}$ is the step size and $\nabla_{\Theta} \text{Cost}(\Theta)$ is the gradient of the cost function evaluated at Θ . The gradient $\nabla_{\Theta} \text{Cost}(\mathbf{P}_{\text{Net}}(\Theta))$ is evaluated numerically using automatic differentiation [76], [77].

In the qNetVO software, we use PennyLane to automatically differentiate quantum network circuit ansatzes. When

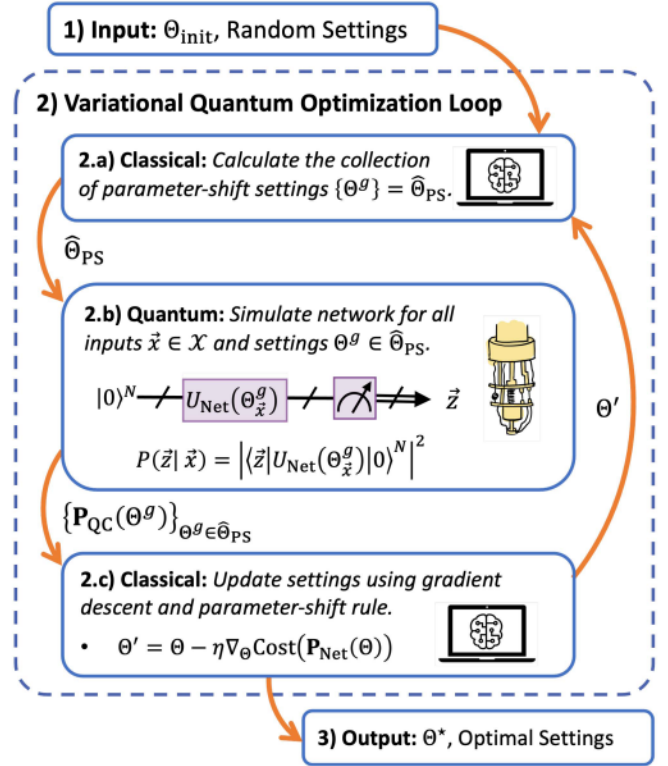


FIGURE 4. VQO. In this hybrid algorithm, the optimal settings Θ^* that minimize the $\text{Cost}(\mathbf{P}_{\text{Net}}(\Theta))$ are computed. Steps 2.a) and 2.c) are performed on a classical computer, while step 2.b) is performed on a quantum computer.

simulated on classical hardware, our software evaluates the gradient of a quantum circuit using the chain rule in an algorithm known as *back propagation* [77]. On quantum hardware, our software computes the gradient of a quantum circuit using the parameter-shift rule [78]. In practice, the parameter-shift rule first runs the network simulation over a collection of “shifted” settings $\{\Theta^g\}_g = \hat{\Theta}_{\text{PS}}$ and then, the classical optimizer uses the resulting circuit behaviors $\{\mathbf{P}_{\text{QC}}(\Theta^g)\}_{\Theta^g \in \hat{\Theta}_{\text{PS}}}$ to construct the gradient.

C. VQO ALGORITHM

This section outlines our framework's hybrid VQO procedure. The described algorithm is implemented by the qNetVO software and can be run using remote quantum hardware or a classical simulator [69]. As input, the VQO algorithm requires a parameterized network ansatz $U_{\text{Net}}(\Theta_{\vec{x}})$ and cost function $\text{Cost}(\mathbf{P}_{\text{Net}}(\Theta))$. As hyperparameters, the algorithm accepts the step size η and the number of gradient descent iterations `num_steps`. The VQO algorithm then proceeds as follows (see Fig. 4).

- 1) The network settings Θ_{init} are randomly initialized.
- 2) The hybrid VQO loop repeats `num_steps` times. The following actions are performed in each step.
 - a) The classical optimizer constructs the collection of parameter-shift settings $\{\Theta^g\}_g = \hat{\Theta}_{\text{PS}}$.

- b) The quantum computer evaluates $P_{QC}(\Theta^g)$ for all settings in $\Theta^g \in \hat{\Theta}_{PS}$.
 - c) The classical optimizer evaluates the gradient $\nabla_{\Theta} \text{Cost}(P_{Net}(\Theta))$ using the parameter-shift rule and updates the network settings using gradient descent.
- 3) The VQO algorithm exits after `num_steps` iterations and outputs the optimal settings Θ' .

Our VQO algorithm is not guaranteed to find Θ^* , the global minimum of the cost function. The algorithm instead outputs settings Θ' that upper bounds the global minimum as $\text{Cost}(\Theta') \geq \text{Cost}(\Theta^*)$. The bound can be improved by repeating the optimization many times with randomly initialized settings. Additionally, hyperparameters such as the step size η can often be adjusted to improve the tightness of the bound $\text{Cost}(\Theta') \geq \text{Cost}(\Theta^*)$.

IV. QUANTUM NON- n -LOCALITY

A. QUANTIFYING NON- n -LOCALITY

In the n -local scenario, entanglement between quantum network devices can be used to create nonclassical correlations referred to as non- n -local correlations. To formalize this concept, we first introduce some notation. Let $\mathcal{L}^{Net}$ and $\mathcal{Q}^{Net}$ denote, respectively, the sets of classical and quantum network behaviors while P_{Net} denotes a network behavior as defined in (2). If $P_{Net} \in \mathcal{L}^{Net}$, then the behavior is classical and its probabilities $P(\vec{a}|\vec{x})$ all can decompose as (3). Likewise, if $P_{Net} \in \mathcal{Q}^{Net}$, then the behavior is quantum and its probabilities all can decompose as (4). In the noiseless case, $\mathcal{L}^{Net} \subseteq \mathcal{Q}^{Net}$. Formally, a quantum network behavior $P \in \mathcal{Q}^{Net}$ is *non- n -local* if and only if $P \notin \mathcal{L}^{Net}$, where both classical and quantum networks have the same topology [37], [38].

To decide whether a network behavior P_{Net} is non- n -local or not, we use the fact that that $\mathcal{L}^{Net}$ is closed, connected, and bound by a set of inequalities referred to as n -local Bell inequalities [37], [38]. We express an n -local Bell inequality as

$$S_{Bell}(\mathbf{P}) \leq \beta \quad (21)$$

where $S_{Bell}(\cdot)$ is a function referred to as the *Bell score* and β is the upper bound for all $\mathbf{P} \in \mathcal{L}^{Net}$. All classical behaviors $\mathbf{P} \in \mathcal{L}^{Net}$ satisfy the Bell inequalities that bound $\mathcal{L}^{Net}$. In general, the Bell score $S_{Bell}(\mathbf{P})$ is a nonlinear function of the probabilities $P(\vec{a}|\vec{x})$ [38], [39], [40], [41], [42], [43], [43], [44], [45], [79], and [80].

The Bell inequalities bounding $\mathcal{L}^{Net}$ can be used to witness quantum non- n -locality. That is, if $S_{Bell}(\mathbf{P}) \not\leq \beta$, then $\mathbf{P} \notin \mathcal{L}^{Net}$. In this case, the Bell inequality is violated and the behavior $\mathbf{P}$ is witnessed to be *non- n -local*. Therefore, the study of quantum non- n -locality reduces to finding non- n -local quantum behaviors $\mathbf{P} \in \mathcal{Q}^{Net}$ that violate a particular Bell inequality. Fortunately, many Bell inequalities have been derived that tightly bound the n -local correlations of important network topologies including star [41], chain [42],

and tree [43], [44], [45] topologies. Thus, non- n -locality can be studied broadly without deriving new n -local bounds.

B. NOISE ROBUSTNESS OF QUANTUM NON- n -LOCALITY

Noise causes the set of quantum behaviors $\mathcal{Q}^{Net}$ to deform. That is, $\mathcal{Q}^{Net} = \mathcal{Q}^{Net}(\text{id}^{Net}) \supset \mathcal{Q}^{Net}(\mathcal{N}^{Net})$, where $\mathcal{Q}_{id}^{Net}$ is the set of noiseless quantum network behaviors and $\mathcal{N}^{Net}$ is some noise model. Furthermore, if a sufficient amount of noise is present, then $\mathcal{Q}^{Net}(\mathcal{N}^{Net}) \subseteq \mathcal{L}^{Net}$ and the non- n -locality of the network is said to be broken. We extend the concept non- n -locality breaking from references [63], [64], which introduce the general concept of nonlocality breaking channels.

The amount of noise that can be tolerated by a quantum network before its behaviors become n -local is the robustness of the network's non- n -locality. To quantify noise, a quantum channel $\mathcal{N}_{\gamma}^{Net}$ or classical postprocessing map E_{γ} is parameterized by the noise parameter $\gamma \in [0, 1]$, where $\gamma = 0$ corresponds to the noiseless case. The robustness of quantum non- n -locality can be quantified by the critical noise parameter γ_0 at which non- n -locality is broken, that is,

$$\gamma_0 = \sup\{\gamma \in [0, 1] : \mathcal{Q}^{Net}(\mathcal{N}_{\gamma}^{Net}) \subseteq \mathcal{L}^{Net}\}. \quad (22)$$

where $\mathcal{Q}^{Net}(\mathcal{N}_{\gamma}^{Net})$ is the set of quantum network behaviors having noise model $\mathcal{N}_{\gamma}^{Net}$.

In general, (22) is challenging to solve because proving that $\mathcal{Q}^{Net}(\mathcal{N}^{Net}) \subseteq \mathcal{L}^{Net}$ in the worst case requires every Bell inequality bounding $\mathcal{L}^{Net}$ to be derived and checked for violation. The practical approach taken in [63], [64] is to obtain the critical noise parameter γ_0 at which the non- n -locality of the network is broken with respect to a particular Bell inequality S_{Bell}

$$\gamma_0 = \sup\{\gamma \in [0, 1] : \max_{\mathbf{P} \in \mathcal{Q}^{Net}(\mathcal{N}_{\gamma}^{Net})} S_{Bell}(\mathbf{P}) \leq \beta\}. \quad (23)$$

However, the challenge still remains of finding the maximal Bell score for a given noisy network.

C. VQO OF QUANTUM NON- n -LOCALITY

We use VQO to maximize the non- n -locality in quantum networks with respect to a particular Bell inequality $S_{Bell}(\mathbf{P})$ and noise model $\mathcal{N}^{Net}$

$$\max_{\mathbf{P} \in \mathcal{Q}^{Net}(\mathcal{N}^{Net})} S_{Bell}(\mathbf{P}). \quad (24)$$

This optimization problem is important for understanding the noise robustness of nonlocality in quantum networks as shown in (23). Furthermore, optimizing nonlocality in noisy quantum networks is a practical example for which to showcase the quantum network VQO techniques developed in Section III.

The maximization of non- n -locality in noisy quantum networks in (24) is well suited for our VQO framework. The quantum network ansatz is constructed similarly to Fig. 2, where the state preparations and measurements are modeled by (11) and (12), respectively. Furthermore, the network

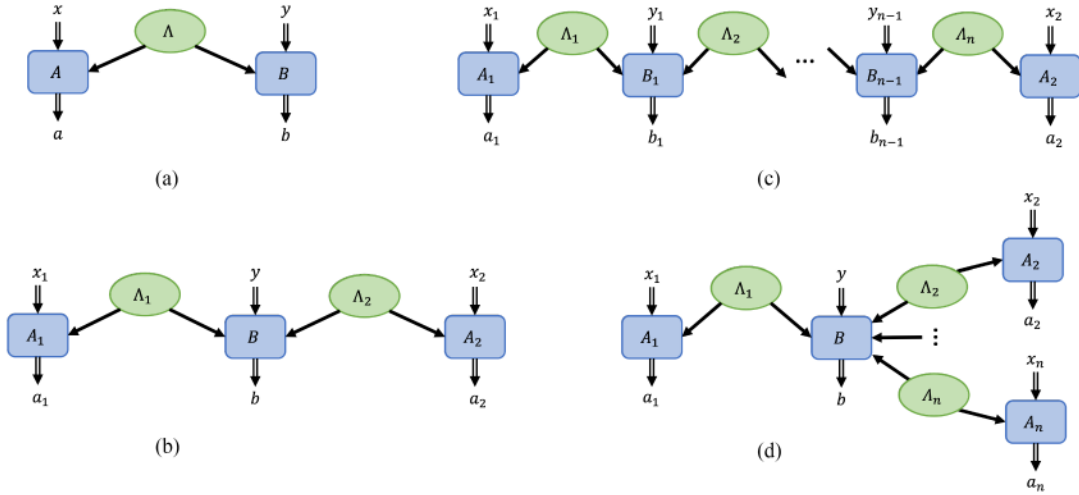


FIGURE 5. Quantum networks. (a) Local network. (b) Bilocal network. (c) n -local chain. (d) n -local star. Sources are depicted as green ovals and nodes are blue rectangles. The labels A_j and B_j distinguish between exterior and interior nodes, respectively.

noise $\mathcal{N}^{\text{Net}}$ can be modeled as described in Section III-A2. The cost function is expressed as

$$\text{Cost}(\mathbf{P}_{\text{Net}}(\Theta)) = -S_{\text{Bell}}(\mathbf{P}_{\text{Net}}(\Theta)) \quad (25)$$

where the minus sign transforms the minimization of the cost into the maximization of the Bell score. It is then a simple matter of using software such as qNetVO [69] to perform the VQO algorithm shown in Fig. 4. Furthermore, the network settings Θ parameterize the state preparations $\rho^{\text{Net}} = \bigotimes_{i=1}^n \rho^{\Lambda_i}$ and measurements $O_{\vec{x}}^{\text{Net}} = \bigotimes_{j=1}^m O_{x_j}^{A_j}$.

We now formalize some notation used throughout the work. First, let $S_{\text{Bell}}(O_{\vec{x}}^{\text{Net}}, \mathcal{N}^{\text{Net}}, \rho^{\text{Net}})$ be the Bell score for a fixed network state preparation ρ^{Net} , noise model $\mathcal{N}^{\text{Net}}$, and observables $O_{\vec{x}}^{\text{Net}}$. We define the maximal Bell score for fixed-state preparations as

$$S_{\text{Bell}}^*(\rho^{\text{Net}}) = \max_{O_{\vec{x}}^{\text{Net}}} S_{\text{Bell}}(O_{\vec{x}}^{\text{Net}}, \text{id}^{\text{Net}}, \rho^{\text{Net}}) \quad (26)$$

where the optimization is over all network observables and id^{Net} denotes a noiseless channel on all $N = 2n$ qubits. Furthermore, for a static noise model, we define

$$\tilde{S}_{\text{Bell}}^*(\mathcal{N}^{\text{Net}}) = \max_{\rho^{\text{Net}} \in D(\mathcal{H}^{\text{Net}})} S_{\text{Bell}}^*(\mathcal{N}^{\text{Net}}, \rho^{\text{Net}}) \quad (27)$$

where the optimization is over all state preparations.

D. BELL INEQUALITIES FOR N -LOCAL NETWORKS

This work considers general n -local star [41] and chain [42] networks (see Fig. 5). Star networks are important because they can perform generalized entanglement swapping protocols [30], [31], [32], [41] where the exterior nodes in the star become arbitrarily entangled based upon the measurement applied in the central node. Similarly, chain networks are important because they model quantum repeater chains that enable long-distance quantum communication via relay nodes that perform entanglement swapping [33],

[34]. In both cases, non- n -locality is important for device-independent certification of the entanglement sources and quantum measurements required to implement entanglement swapping protocols and network security protocols [13], [15], [60], [61].

To study quantum nonlocality, we quantify the correlation across multiple measurement nodes using the formalism of Hermitian observables. We focus on networks that output a binary value $a_j \in \{\pm 1\}$ at each measurement node. Hence, the observable at the j th node is then expressed as $O_{x_j}^{A_j} = \sum_{a_j \in \{\pm 1\}} a_j \Pi_{a_j|x_j}^{A_j}$, where $\Pi_{+|x_j}^{A_j}$ and $\Pi_{-|x_j}^{A_j}$ constitute a PVM. In the qubit case, the observable is expressed in the Pauli basis as $O_{x_j}^{A_j} = \vec{s} \cdot \vec{\sigma}$, where the Bloch vector $\vec{s} \in \mathbb{R}^3$ has elements $s_j = \text{Tr}[O_{x_j}^{A_j} \sigma_j]$ and the Pauli vector is $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$.

For a fixed state ρ^{A_j} , the expectation of the observable is

$$\langle O_{x_j}^{A_j} \rangle_{\rho^{A_j}} = \text{Tr}[O_{x_j}^{A_j} \rho^{A_j}] = \sum_{a_j \in \pm 1} a_j P(a_j|x_j). \quad (28)$$

The correlation between network nodes is quantified by the expectation of the m -partite correlator

$$\langle O_{x_1}^{A_1} \dots O_{x_m}^{A_m} \rangle_{\rho^{\text{Net}}} = \text{Tr}\left[\left(\bigotimes_{j=1}^m O_{x_j}^{A_j}\right) \rho^{\text{Net}}\right]. \quad (29)$$

Since the considered observables $O_{x_j}^{A_j}$ are dichotomic, (29) corresponds to the expectation value of the parity of the bit string $\vec{a}$ output from the network.

We now introduce the four quantum networks on which this work focuses its investigation, local, bilocal, chain, and star (see Fig. 5). When all measurement devices have binary inputs and outputs, a class of nonlinear n -local Bell inequalities is known to bound these quantum networks. For all considered networks, the optimal quantum violation is achieved

in a noiseless setting using maximally entangled state preparations $|\psi^{\Lambda_i}\rangle = |\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ and qubit measurements with bases aligned in the xz -plane of Bloch sphere.

1) CHSH NETWORKS

Consider the local $n = 1$ case depicted in Fig. 5(a) having one source and two nonsignaling measurement devices. This simple two-qubit network is the fundamental example of nonlocality and is commonly referred to as the Clauser–Horne–Shimony–Holt (CHSH) scenario [37]. The set of classical behaviors for this network $\mathcal{L}_{\text{CHSH}}$ is bound by the CHSH inequality [81]

$$S_{\text{CHSH}} = \left| \sum_{x,y=0}^1 (-1)^{xy} \langle O_x^A O_y^B \rangle \right| \leq 2. \quad (30)$$

The maximal violation of the CHSH inequality is $S_{\text{CHSH}}^* = 2\sqrt{2} \not\leq 2$ [82].

2) BILOCAL NETWORKS

Consider the bilocal $n = 2$ network depicted in Fig. 5(b) having two sources and three measurement devices. The two sources each hold two-qubits, $\Lambda_1 = (q_1, q_2)$ and $\Lambda_2 = (q_3, q_4)$. The two exterior measurement node $A_1 = (q_1)$ and $A_2 = (q_4)$ measure the first and last qubit, respectively. The central node $B = (q_2, q_3)$ jointly measures the qubits received from sources Λ_1 and Λ_2 . The set of classical bilocal behaviors $\mathcal{L}_{\text{Bilocal}}$ is bounded by the bilocal Bell inequality [39], [40]

$$S_{\text{Bilocal}} = \sqrt{|I_{2,y=0}|} + \sqrt{|I_{2,y=1}|} \leq 1 \quad (31)$$

where the quantity

$$I_{2,y} = \frac{1}{4} \sum_{x_1, x_2} (-1)^{y(x_1+x_2)} \langle O_{x_1}^{A_1} O_y^B O_{x_2}^{A_2} \rangle_{\rho_{\text{Bilocal}}} \quad (32)$$

is a linear combination of tripartite correlators on the network state $\rho^{\text{Bilocal}} = \rho^{\Lambda_1} \otimes \rho^{\Lambda_2}$. The maximal violation is $S_{\text{Bilocal}}^* = \sqrt{2} \not\leq 1$ [40].

3) N-LOCAL CHAIN NETWORKS

As depicted in Fig. 5(c), the n -local chain network is an important extension of the bilocal network. In this network, the n sources $\{\Lambda_i = (q_{2i-1}, q_{2i})\}_{i=1}^n$ connect a collection of $m = n + 1$ nonsignaling measurement devices in a chain structure. The exterior nodes $A_1 = (q_1)$ and $A_2 = (q_{2n})$ measure the first and last qubits, respectively, while the interior nodes $\{B_j = (q_{2j-2}, q_{2j-1})\}_{j=2}^n$ each measure two qubits. The set of n -local chain behaviors $\mathcal{L}_{n\text{-Chain}}$ is bounded by the inequality [42]

$$S_{n\text{-Chain}} = \sqrt{|J_{n,y=0}|} + \sqrt{|J_{n,y=1}|} \leq 1 \quad (33)$$

where $J_{n,y} = I_{2,y}$ as defined in (32) when the observable O_y^B takes the form $O_y^B = \bigotimes_{j=2}^n O_{y_j=y}^{B_j}$. Here, $O_{y_j=y}^{B_j} = O_y^{q_1} \otimes O_y^{q_2}$ is the two-qubit observable applied at each interior node

in the chain while the observables $\{O_y^{B_j}\}_{j=2}^n$ jointly conditioned on the input $y \in \{0, 1\}$. The n -local chain Bell inequality is identical to bilocal inequality S_{Bilocal} described by (31) and the maximal quantum violation of (33) is $S_{n\text{-Chain}}^* = \sqrt{2} \not\leq 1$ [42].

4) N-LOCAL STAR NETWORKS

The n -local star network depicted in Fig. 5(d) consists of n , two-qubit sources $\{\Lambda_i = (q_i, q_{n+i})\}_{i=1}^n$ and a collection of $m = n + 1$ measurement nodes arranged in a star formation. In this formation, the set of n exterior nodes $\{A_j = (q_j)\}_{j=1}^n$ each measure one qubit and serve as points of the star. Each exterior node A_j is connected to a single central node $B = (q_{n+j})_{j=1}^n$ using the entanglement source $\Lambda_{i=j}$, where central node jointly measures its local n -qubit state. The set of classical n -local star network behaviors $\mathcal{L}_{n\text{-Star}}$ is bounded by the n -local star Bell inequality defined as [41]

$$S_{n\text{-Star}} = |I_{n,0}|^{1/n} + |I_{n,1}|^{1/n} \leq 1 \quad (34)$$

where the quantity

$$I_{n,y} = \frac{1}{2^n} \sum_{x_1, \dots, x_n} (-1)^{y(\sum_j x_j)} \langle O_{x_1}^{A_1} \dots O_{x_n}^{A_n} O_y^B \rangle_{\rho^{\text{Net}}} \quad (35)$$

is an n -partite generalization of (32). The observables $O_{x_j}^{A_j}$ and O_y^B are all assumed to be dichotomic. The maximal quantum violation of the star inequality is $S_{n\text{-Star}}^* = \sqrt{2} \not\leq 1$ [41]. In the $n = 1$ and $n = 2$ cases, the n -local star network reduces to the CHSH and bilocal networks, respectively. That is, it can be shown that

$$S_{1\text{-Star}} = \frac{1}{2} S_{\text{CHSH}} \quad S_{2\text{-Star}} = S_{\text{Bilocal}}. \quad (36)$$

Hence, the star network, like the chain network, generalizes the CHSH and bilocal networks.

E. MAXIMAL N-LOCAL VIOLATIONS

In Section IV-D, the optimal quantum strategies for non-locality are described in the noiseless case. In the presence of noise, these strategies are not guaranteed to be optimal. When a network measures local qubit observables, $O_{x_j}^{A_j} = \bigotimes_{k=1}^{|A_j|} O_{x_j}^{q_k}$ at all A_j , the maximal n -local violation for star and chain networks is [83]

$$S_{n\text{-Star}}^* (\rho^{\text{Net}}) = \frac{1}{2} \prod_{i=1}^n S_{\text{CHSH}}^* (\rho^{\Lambda_i})^{\frac{1}{n}} \quad (37)$$

$$S_{n\text{-Chain}}^* (\rho^{\text{Net}}) = S_{2\text{-Star}}^* (\rho^{\Lambda_1} \otimes \rho^{\Lambda_n}) \prod_{i=2}^{n-1} \tau_{i,0} \quad (38)$$

where the maximal CHSH score is [84]

$$S_{\text{CHSH}}^* (\rho^{\Lambda_i}) = 2\sqrt{\tau_{i,0}^2 + \tau_{i,1}^2}. \quad (39)$$

In these equations, $1 \geq \tau_{i,0} \geq \tau_{i,1} \geq 0$ are the two largest singular values of the two-qubit correlation matrix $T_{\rho^{\Lambda_i}} \in \mathbb{R}^{3 \times 3}$

having elements

$$T_{\rho^{\Lambda_i}}^{(j,k)} = \text{Tr} \left[\rho^{\Lambda_i} \sigma_j^{q_1} \otimes \sigma_k^{q_2} \right] \quad (40)$$

where $j, k \in \{x, y, z\}$ index the Pauli matrices.

The maximal Bell scores in (37) and (38) do not account for entangled measurements or positive operator-valued measures (POVMs). In general, the maximal n -local star score is the geometric mean of independent CHSH violations, as shown in (37), and is achieved using measurements separable across qubit systems. Thus, we do not expect entangled measurements to improve the maximal n -local scores beyond (37) and (38). For a general discussion regarding the maximal qubit violation of n -locality in star and chain networks and the corresponding measurement strategies, please to [83]. In Appendix C, we give concrete examples relating the maximal n -local scores in (37) and (38) to the maximal n -local scores derived in prior works [85], [86], [87].

V. USING VQO TO INVESTIGATE THE NOISE ROBUSTNESS OF QUANTUM NON- n -LOCALITY

In this section, we demonstrate on a classical simulator that VQO can reproduce the maximal n -local scores derived in Appendix D. We begin with an overview of our use of VQO to investigate the noise robustness of quantum non- n -locality. Then, for unital and nonunital channel examples, we derive the theoretical n -local scores and show that our VQO methods can obtain these values. We find that network noise models consisting of general unital qubit channels have their non- n -locality maximized by maximally entangled states are optimal. On the contrary, we show that when nonunital qubit amplitude damping channels are considered, there exist non-maximally entangled state preparations that outperform maximally entangled preparations.

A. VQO OF THE NOISE ROBUSTNESS OF NON- n -LOCALITY

Our objectives are to verify that our VQO framework can reproduce known and derived noise robustness results. Our investigative approach to noise robustness is distinct from previous works [63], [64] that evaluate the precise noise parameters at which nonlocality is broken. Instead, we use VQO to find maximal violations of a Bell inequality given a static noise model $\mathcal{N}_{\vec{\gamma}}^{\text{Net}}$. By scanning through the noise parameters, we create a picture of how the non- n -locality deteriorates as the amount of noise increases. Hence, we are able to easily compare the relative noise robustness across different quantum network topologies.

We consider noise applied during the preparation, communication, and measurement stages of a quantum network in the n -local setting (see Fig. 6). Source noise occurs during the state preparation at each source and is modeled as $\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{N}_{\gamma_i}^{\Lambda_i}$. Communication noise occurs during the transmission of quantum states and is modeled as

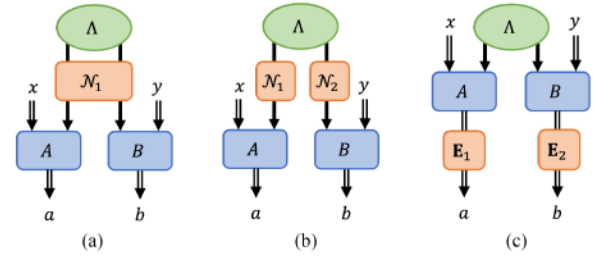


FIGURE 6. Noise models in quantum networks. (a) Source noise: A quantum channel is applied to all qubits at a given source. (b) Communication noise: A quantum channel is applied independently to each qubit. (c) Detector noise: Classical postprocessing is applied to the classical data output from the network.

$\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{k=1}^l \mathcal{N}_{\gamma_k}^{\Lambda_k}$. Detector noise occurs during measurement and is modeled as $\mathbf{E}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{j=1}^m \mathbf{E}_{\gamma_j}^{\Lambda_j}$. Alternatively, detector noise can be modeled as an adjoint channel applied to the measurement $\mathcal{N}_{\vec{\gamma}}^{\text{Net}\dagger}(\Pi_{a|\vec{x}}^{\text{Net}})$.

Our VQO approach can easily be extended to nonuniform noise models as it introduces no additional computational overhead in comparison to the uniform noise model. To simplify our investigation, we characterize the network noise using one parameter. First, we consider an ideal quantum network that has a single faulty component such as a noisy source, link, or measurement device. In this case, the noise parameters take the form $\gamma_v = (\gamma, 0, \dots, 0)$. Second, we consider quantum networks having noise applied uniformly to all sources, links, or measurement devices. In this case, the noise parameters are $\gamma_v = (\gamma, \dots, \gamma)$.

To evaluate the noise robustness, we begin with a static noise model $\mathcal{N}_{\vec{\gamma}}^{\text{Net}}$. To create a high-level overview of the noise robustness, we scan through the noise parameter $\gamma \in [0, 1]$ using an interval of 0.05. For each γ , we use VQO to find the optimal state preparations and measurements that maximize non- n -locality with respect to the Bell inequality S_{Bell} . We repeat this procedure for all considered network topologies depicted in Fig. 5 and compare their relative noise robustness. Furthermore, we compare the optimized results with theoretical bounds on the max violation. For some channels, we derive max violation directly and show that our optimizations reproduce the expected results. In other cases, we use (37), (38), and (39) to derive the maximal violation for a fixed state such as the Bell state $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$.

We organize our investigation into two broad classes of noise, unital and nonunital. In each case, we consider noise applied to sources, qubit communication, and detectors. For each noise model, we compare the $n \leq 4$ cases for the chain and star n -local Bell inequalities expressed in (33) and (34), respectively. We note that we label the $n = 1$ and $n = 2$ cases as the CHSH and bilocal Bell inequalities due to the relations expressed in (36). Finally, we explore a wide range of state preparation and measurement ansatzes to broadly investigate the relation between entanglement and noise robustness. For details, refer to Table 1 in Appendix B.

B. UNITAL CHANNELS

Unital channels model common types of hardware noise such as depolarizing and dephasing. A quantum channel $\mathcal{U}$ is unital if and only if it satisfies $\mathcal{U}(\mathbb{I}) = \mathbb{I}$. Consequently, unital noise cannot improve the purity of a general quantum state ρ , that is, $\text{Tr}[\mathcal{U}(\rho)^2] \leq \text{Tr}[\rho^2]$. For more details on the theory of unital channels, refer to Appendix D.

In the following sections, we demonstrate on a classical simulator that our VQO framework can find the theoretically maximal n -local violations in the presence of unital channels such as qubit depolarizing noise, source depolarizing noise, detector white noise, and qubit dephasing noise. We show that network ansatzes allowing maximally entangled state preparation and local qubit measurements can achieve the theoretical maximum. Furthermore, we find no example of entangled measurements or nonmaximally entangled states achieving a larger Bell score. These numerical results provide evidence that maximally entangled state preparations and local qubit measurements are optimal for general unital noise models on sources and measurements. Indeed, the numerical VQO results in this section led to the derivation of (37) and (38) in [83]; thus, our VQO framework can be used to obtain novel theoretical insights.

1) QUBIT DEPOLARIZING NOISE ROBUSTNESS

A qubit depolarizing channel mixes white noise with the input qubit state as

$$\mathcal{D}_v(\rho) = v\rho + \frac{(1-v)}{2}\mathbb{I}_2\text{Tr}[\rho] \quad (41)$$

where v is a parameter commonly referred to as the *visibility*. The visibility relates to the noise parameter as

$$\gamma = \frac{3}{4}(1-v). \quad (42)$$

The qubit depolarizing channel's Kraus operators are

$$\begin{aligned} K_0 &= \sqrt{1-\gamma}\mathbb{I}_2 & K_1 &= \sqrt{\frac{\gamma}{3}}\sigma_x \\ K_2 &= \sqrt{\frac{\gamma}{3}}\sigma_y & K_3 &= \sqrt{\frac{\gamma}{3}}\sigma_z. \end{aligned} \quad (43)$$

In Proposition 1 of Appendix D1, we show that in the presence of qubit depolarizing noise, $\mathcal{D}_{\vec{v}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{D}_{v_i}^{A_i} \otimes \mathcal{D}_{v_i}^{B_i}$ the maximal n -local star and chain scores are

$$\tilde{S}_{n\text{-Star}}^*(\mathcal{D}_{\vec{v}}^{\text{Net}}) = \sqrt{2} \left(\prod_{i=1}^n v_i^{A_i} v_i^{B_i} \right)^{\frac{1}{n}} \quad (44)$$

$$\tilde{S}_{n\text{-Chain}}^*(\mathcal{D}_{\vec{v}}^{\text{Net}}) = \sqrt{2} \left(\prod_{i=1}^n v_i^{A_i} v_i^{B_i} \right)^{\frac{1}{2}}. \quad (45)$$

In Fig. 7, we find a close correspondence between the theoretical maximal Bell scores given by (44) and (45) and the maxima obtained using our VQO framework. In our optimizations, we consider arbitrary preparation and measurement ansatzes (see Table 2 in Appendix B). We find that

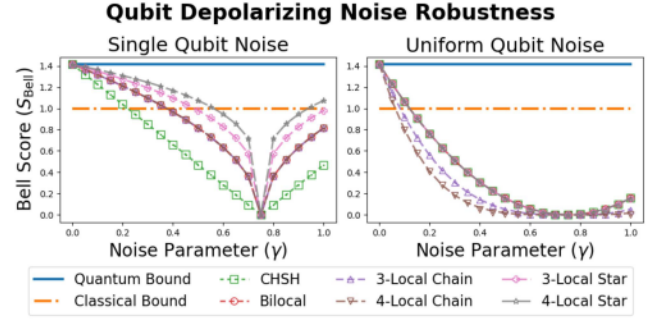


FIGURE 7. Qubit depolarizing noise robustness of non- n -locality. (Left) Qubit depolarizing noise is applied to qubit q_1 . (Right) Qubit depolarizing noise is applied uniformly to all qubits. The markers show the maximal Bell score achieved using VQO. The dashed lines show the theoretical maximal given by (44) and (45), where the relation between the visibility v and the noise parameter γ is given by (42).

maximal violations can be achieved using the Bell state $|\Phi^+\rangle$ and local qubit measurements.

2) SOURCE DEPOLARIZING NOISE ROBUSTNESS

Depolarizing noise on two-qubit sources is expressed as

$$\mathcal{D}_v(\rho) = v\rho + \frac{(1-v)}{4}\mathbb{I}_4\text{Tr}[\rho] \quad (46)$$

where the visibility v relates to the noise parameter as

$$\gamma = \frac{15}{16}(1-v). \quad (47)$$

The Kraus operators for a two-qubit depolarizing channel are expressed as

$$K_{0,0} = \sqrt{1-\gamma}\mathbb{I}_4, \quad \left\{ K_{i,j} = \sqrt{\frac{\gamma}{15}}\sigma_i \otimes \sigma_j \right\}_{i,j=0}^3 \quad (48)$$

where all i and j are considered except $i = j = 0$. The maximal Bell scores have previously been derived for n -local star and chain networks having source depolarizing noise. Namely, the maximal noisy Bell score for the star network is $\tilde{S}_{n\text{-Star}}^*(\bigotimes_{i=1}^n \mathcal{D}_{v_i}^{A_i}) = \sqrt{2}(\prod_{i=1}^n v_i)^{1/n}$ [40], [41], and the maximal noisy Bell score for the chain network is $\tilde{S}_{n\text{-Chain}}^*(\bigotimes_{i=1}^n \mathcal{D}_{v_i}^{A_i}) = \sqrt{2}\sqrt{\prod_{i=1}^n v_i}$ [42]. We use (47) to redefine the maximal violations in terms of the noise parameter as

$$\tilde{S}_{n\text{-Star}}^*(\bigotimes_{i=1}^n \mathcal{D}_{v_i}^{A_i}) = \sqrt{2} \left(\prod_{i=1}^n \left| 1 - \frac{16}{15}\gamma_i \right| \right)^{1/n} \quad (49)$$

$$\tilde{S}_{n\text{-Chain}}^*(\bigotimes_{i=1}^n \mathcal{D}_{v_i}^{A_i}) = \sqrt{2} \sqrt{\prod_{i=1}^n \left| 1 - \frac{16}{15}\gamma_i \right|}. \quad (50)$$

In Fig. 8, we show that there is a close correspondence between the theoretical maximal Bell scores given by (49) and (50) and the maxima obtained using VQO optimization. In our optimization, we consider source depolarizing noise and a broad range of preparation and measurement ansatzes (see Table 2 in Appendix B). We find that the state preparation

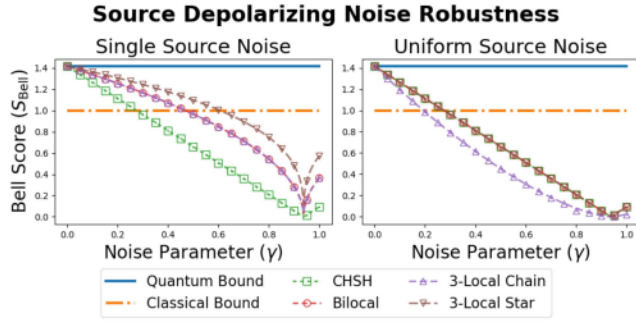


FIGURE 8. Source depolarizing noise robustness of non- n -locality. (Left) Depolarizing noise is applied to source A_1 . (Right) Depolarizing noise is applied uniformly to all sources. The markers show the maximal Bell score achieved using VQO. The dashed lines show the theoretical maximal score from (49) and (50).

$|\Phi^+\rangle$ and local qubit measurements optimized over rotations about the y -axis are sufficient for maximal violations in all cases.

3) WHITE NOISE DETECTOR ERRORS

For a detector with binary outputs, we define a white noise error as the classical postprocessing map

$$W_\gamma = (1 - \gamma)\mathbb{I}_2 + \frac{\gamma}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (51)$$

where with probability γ , the detector outputs a binary value drawn from a uniform random distribution. Noting that (51) is a convex combination of noiseless and white noise measurements, we can equivalently express W_γ as the POVM

$$W_{\pm|x,\gamma} = (1 - \gamma)\Pi_{\pm|x} + W_{\pm|x} \quad (52)$$

where $W_{\pm|x} = \frac{1}{2}\mathbb{I}_M$ are white noise POVM elements and $\Pi_{\pm|x}$ are projectors onto even (+) and odd (−) parity subspaces that satisfy $\text{Tr}[\Pi_{\pm|x}] = 2^{M-1}$. The projector onto even and odd parity subspaces corresponds to the fact that we use parity to coarse grain a multibit output into a single bit.

In Proposition 3 of Appendix D3, we show that detector white noise postprocessing map W_γ is a unital process equivalent to a depolarizing channel on the detector's M -qubits

$$\mathcal{D}_{(1-\gamma)}^{A_j}(X) = (1 - \gamma)X + \frac{\gamma}{2M}\mathbb{I}_{2M}\text{Tr}[X]. \quad (53)$$

Then, in Proposition 4 of Appendix D3, we find that given the network noise model $\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{j=1}^m \mathcal{D}_{(1-\gamma_j)}^{A_j}$, the maximal n -local star and chain scores are

$$\tilde{S}_{n\text{-Star}}^* \left(\bigotimes_{j=1}^m \mathcal{D}_{(1-\gamma_j)}^{A_j} \right) = \sqrt{2} \left(\prod_{j=1}^{n+1} (1 - \gamma_j) \right)^{1/n} \quad (54)$$

$$\tilde{S}_{n\text{-Chain}}^* \left(\bigotimes_{j=1}^m \mathcal{D}_{(1-\gamma_j)}^{A_j} \right) = \sqrt{2} \sqrt[n+1]{\prod_{j=1}^{n+1} (1 - \gamma_j)}. \quad (55)$$

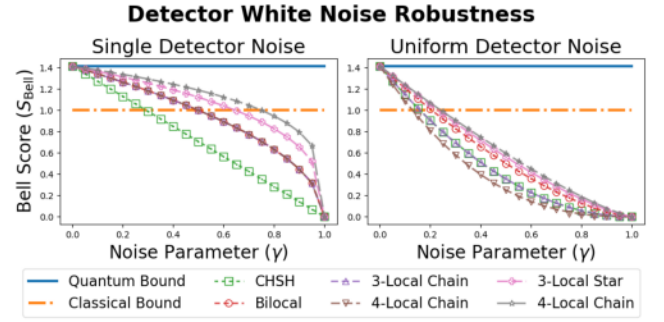


FIGURE 9. Detector white noise robustness of non- n -locality. (Left) Detector white noise is applied to measurement device A_1 . (Right) Detector white noise is applied uniformly to all measurement devices. The markers show the maximal Bell score achieved using VQO. The dashed lines show the theoretical maximal score from (54) and (55).

In Fig. 9, we show that there is a close correspondence between the theoretical maximal n -local violations given by (54) and (55) and the maxima obtained using VQO. In our optimizations, we consider a broad range of preparation and measurement ansatzes (see Table 2 in Appendix B). For all networks, we find that the state preparation $|\Phi^+\rangle$ and local qubit measurements optimized over rotations about the y -axis are sufficient to achieve the theoretical maximum.

4) QUBIT DEPHASING NOISE

Qubit dephasing noise is a unital channel that describes the decoherence process as

$$\mathcal{P}_\gamma(\rho) = \frac{1 + \sqrt{1-\gamma}}{2}\rho + \frac{1 - \sqrt{1-\gamma}}{2}\sigma_z\rho\sigma_z \quad (56)$$

where the off-diagonal terms go to zero as the noise parameter γ increases. The Kraus operators for the dephasing channel are

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad K_1 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{pmatrix}. \quad (57)$$

In Proposition 2 of Appendix D2, we show that, given a network noise model $\mathcal{P}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{P}_{\gamma^{A_i}}^{A_i} \otimes \mathcal{P}_{\gamma^{B_i}}^{B_i}$, the maximal n -local star and chain scores are

$$\tilde{S}_{n\text{-Star}}^* \left(\mathcal{P}_{\vec{\gamma}}^{\text{Net}} \right) = \prod_{i=1}^n (1 + (1 - \gamma^{A_i})(1 - \gamma^{B_i}))^{\frac{1}{2n}} \quad (58)$$

$$\tilde{S}_{n\text{-Chain}}^* \left(\mathcal{P}_{\vec{\gamma}}^{\text{Net}} \right) = \tilde{S}_{2\text{-Star}}^* \left(\mathcal{P}_{\gamma^{A_1}}^{A_1} \otimes \mathcal{P}_{\gamma^{B_1}}^{B_1} \otimes \mathcal{P}_{\gamma^{A_n}}^{A_n} \otimes \mathcal{P}_{\gamma^{B_n}}^{B_n} \right). \quad (59)$$

In Fig. 10, we show that there is a close correspondence between the theoretical maximal n -local violations given by (58) and (59) and the maxima obtained using VQO. In our optimizations, we consider a broad range of preparation and measurement ansatzes (see Table 2 in Appendix B). We find that maximally entangled states and local qubit measurements are sufficient for maximal n -local violation.

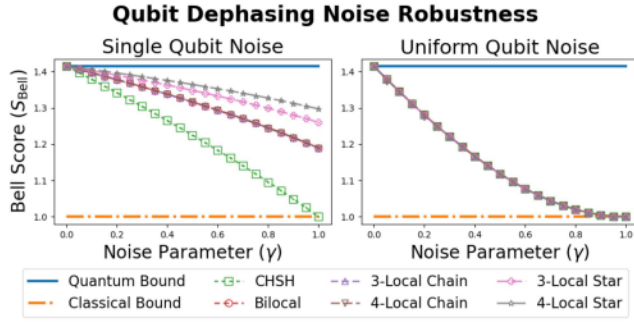


FIGURE 10. Qubit dephasing noise robustness of non- n -locality. (Left) Dephasing noise is applied to qubit q_1 . (Right) Dephasing noise is applied uniformly to all qubits. The markers show the maximal Bell score achieved using VQO. The dashed lines show the maximal scores predicted by (58).

C. NONUNITAL NOISE

In this section, we investigate the noise robustness of quantum non- n -locality with respect to nonunital noise. A quantum channel is nonunital if and only if it does not preserve the identity $\mathcal{N}(\mathbb{I}) \neq \mathbb{I}$, where important examples include qubit amplitude damping, fixed errors on measurements, and two-qubit colored noise. As opposed to unital channels, we show that maximally entangled state preparations are not necessarily optimal in the presence of nonunital noise. We verify that VQO can be used to find the optimal state preparations and measurements for maximal n -local violation results by comparing our numerical results with the theoretical maximal noisy n -local scores derived in Appendix E.

1) QUBIT AMPLITUDE DAMPING NOISE

The qubit amplitude damping channel $\mathcal{A}_\gamma$ describes the energy dissipation process where the high-energy $|1\rangle\langle 1|$ state transitions into the low-energy $|0\rangle\langle 0|$ state. The Kraus operators are defined as

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} \quad K_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} \quad (60)$$

where the effect on a qubit density matrix is

$$\mathcal{A}_\gamma(\rho) = \begin{pmatrix} \gamma + (1-\gamma)\rho_{00} & \sqrt{1-\gamma}\rho_{01} \\ \sqrt{1-\gamma}\rho_{10} & (1-\gamma)\rho_{11} \end{pmatrix}. \quad (61)$$

In star and chain networks, we compare the noise robustness of n -local violations for nonmaximally entangled states, $|\psi_\lambda\rangle = \sqrt{\lambda}|00\rangle + \sqrt{1-\lambda}|11\rangle$ for $\lambda \in [0, 1]$, and maximally entangled states $|\psi_{\lambda=1/2}\rangle$. We express the noisy maximally entangled states as $\tilde{\rho}_{\frac{1}{2},\gamma}^A = \mathcal{A}_\gamma^A \otimes \mathcal{A}_\gamma^B(\rho_{\lambda=1/2})$, and noisy nonmaximally entangled states as $\tilde{\rho}_{\lambda,\gamma}^A = \mathcal{A}_\gamma^A \otimes \mathcal{A}_\gamma^B(\rho_\lambda)$. We now prove through example that nonmaximally entangled states achieve larger Bell scores than maximally entangled states. Then, we verify that our VQO software can reproduce our theoretical results.

First, consider the CHSH scenario with uniform qubit amplitude damping noise where the noise parameters are $\gamma^A = \gamma^B = \gamma$. For maximally entangled states, we use Proposition 6 in Appendix E1 to find that the maximal CHSH score is

$$S_{\text{CHSH}}^*(\tilde{\rho}_{\frac{1}{2},\gamma}) = \begin{cases} 2\sqrt{2(1-\gamma)^2}, & \gamma \in [0, \frac{1}{2}] \\ 2\sqrt{(1-\gamma)^2 + (2\gamma^2 - 2\gamma + 1)^2}, & \gamma \in [\frac{1}{2}, 1] \end{cases} \quad (62)$$

where $\gamma_{c,\frac{1}{2}} = 1/2$ is the noise parameter at which the crossover occurs between the two curves in (62). Furthermore, we see from (62) that the CHSH nonlocality is broken for all maximally entangled states $\rho_{\lambda=1/2}$ at the critical noise parameter $\gamma_{0,\frac{1}{2}} = (1 - \frac{1}{\sqrt{2}})$.

Next, in Proposition 7 of Appendix E1, we show that for noisy nonmaximally entangled states, the maximal CHSH score is

$$S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda^*,\gamma}) = \begin{cases} 2\sqrt{2(1-\gamma)^2}, & \gamma \in [0, \gamma_c] \\ 2\sqrt{(1-\gamma)^2 - \frac{(2\gamma^2 - 2\gamma + 1)^2}{4\gamma^2 - 1}}, & \gamma \in [\gamma_c, \frac{1}{3}] \\ 2, & \gamma \in [\frac{1}{3}, 1] \end{cases} \quad (63)$$

where the optimal entanglement parameter is

$$\lambda^* = \begin{cases} \frac{1}{2}, & \gamma \in [0, \gamma_c] \\ \frac{1}{2} \left(1 - \frac{\gamma^2 + (1-\gamma)^2(2\gamma(1-\gamma))}{(2\gamma(1-\gamma))^2 - (1-\gamma)^2} \right), & \gamma \in [\gamma_c, \frac{1}{3}] \\ 1, & \gamma \in [\frac{1}{3}, 1] \end{cases} \quad (64)$$

the crossover noise parameter is

$$\gamma_c = \frac{1}{6} \left(4 + \frac{(3\sqrt{114} - 32)^{\frac{1}{3}}}{2^{\frac{2}{3}}} - \frac{1}{(6\sqrt{114} - 64)^{\frac{1}{3}}} \right) \quad (65)$$

and the nonlocality is broken with respect to the CHSH inequality when $\gamma = \gamma_0 = \frac{1}{3}$. Therefore, in the presence of uniform qubit amplitude damping noise, maximally entangled states maximize the CHSH score in the range $\gamma \in [0, \gamma_c]$, nonmaximally entangled states maximize the CHSH score in the range $\gamma \in (\gamma_c, \frac{1}{3})$, and the classical state $|00\rangle\langle 00|$ maximizes the CHSH score in the range $\gamma \in [\frac{1}{3}, 1]$. The maximal separation between the CHSH scores of maximally and nonmaximally entangled states occurs at $\gamma = (1 - \frac{1}{\sqrt{2}})$, where the nonmaximally entangled state achieves

$S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda^*}) = 2\sqrt{\frac{1}{14}(3 + 8\sqrt{2})} \approx 2.0222836 > 2$. We plot this data in Fig. 11, where we show that VQO successfully finds the maximal CHSH score for both maximally entangled and nonmaximally entangled state preparations.

We now generalize our uniform qubit amplitude damping results to n -local star and chain networks. Using (37), we can write

$$S_{n\text{-Star}}^* \left(\bigotimes_{i=1}^n \tilde{\rho}_{\lambda=\frac{1}{2},\gamma} \right) = \frac{1}{2} S_{\text{CHSH}}^* (\tilde{\rho}_{\lambda^*,\gamma}). \quad (66)$$

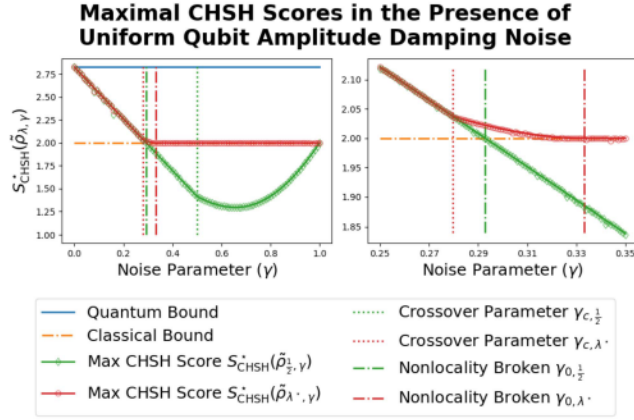


FIGURE 11. Uniform qubit amplitude damping in the CHSH scenario. We plot the maximal CHSH scores obtained when uniform qubit amplitude damping noise is applied to maximally entangled states (green) and nonmaximally entangled states (red). (Left) Noise parameter $\gamma \in [0, 1]$ is scanned over an interval of 0.01. (Right) Noise parameter $\gamma \in [0.25, 0.35]$ is scanned over an interval of 0.001. The solid green line plots the maximal CHSH score for maximally entangled states, as given by (62), while the diamond markers plot numerical results obtained using VQO. The solid red line plots the maximal CHSH score for nonmaximally entangled states, as given by (63). The green and red dotted vertical lines plot the crossover parameters at which the cases switch in (62) and (63), respectively. Likewise, the green and red dash-dotted vertical lines plot the critical parameters at which nonlocality is broken for maximally entangled states and nonmaximally entangled states, respectively.

For the chain network, we note that $\mathcal{A}_\gamma \otimes \mathcal{A}_\gamma(|00\rangle\langle 00|) = |00\rangle\langle 00|$, that is, the $|00\rangle\langle 00|$ state experiences no noise from amplitude damping. Since for $\rho = |00\rangle\langle 00|$, the largest singular value of T_ρ is $\tau_0 = 1$, then (38) can be used to obtain

$$S_{n\text{-Chain}}^* \left(\bigotimes_{i=1}^n \tilde{\rho}_{\lambda=\frac{1}{2}, \gamma} \right) = \frac{1}{2} S_{\text{CHSH}}^* (\tilde{\rho}_{\lambda^*, \gamma}). \quad (67)$$

However, in the case of maximally entangled state preparations, the maximal singular value is $\tau_{0, \gamma} = \max\{1 - \gamma, \gamma^2 + (1 - \gamma)^2\}$; thus, by (38), we have

$$S_{n\text{-Chain}}^* \left(\bigotimes_{i=1}^n \tilde{\rho}_{\lambda=\frac{1}{2}, \gamma} \right) = \frac{1}{2} S_{\text{CHSH}}^* (\tilde{\rho}_{\lambda=\frac{1}{2}, \gamma}) \prod_{i=2}^{n-1} (\tau_{0, \gamma})^{\frac{1}{2}}. \quad (68)$$

We now consider the case where amplitude damping noise is applied to a single qubit in the network. Like in the uniform noise case, we start with the CHSH scenario and maximally entangled states. Using Proposition 6 of Appendix E1 and setting $\gamma^B = 0$ and $\gamma^A = \gamma \in [0, 1]$, we find that $\tau_0 = \tau_1 = \sqrt{1 - \gamma} \geq \tau_2 = 1 - \gamma$. Thus

$$S_{\text{CHSH}}^* (\tilde{\rho}_{\lambda=\frac{1}{2}, \gamma}) = 2\sqrt{2(1 - \gamma)} \quad (69)$$

and by (37), we see that

$$S_{n\text{-Star}}^* \left(\tilde{\rho}_{\lambda=\frac{1}{2}, \gamma} \bigotimes_{i=2}^n |\Phi^+\rangle\langle\Phi^+| \right) = \left((\sqrt{2})^{n-1} \sqrt{2(1 - \gamma)} \right)^{\frac{1}{n}}. \quad (70)$$

Note that the n -local chain score is equivalent to (70) because the interior chain nodes are not affected by the amplitude damping noise.

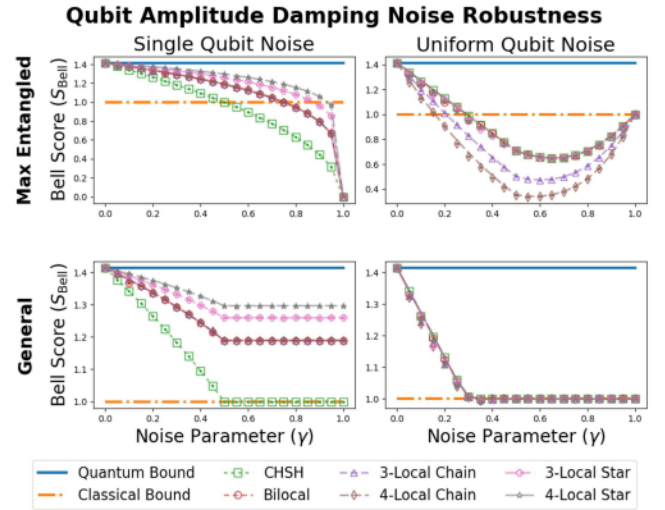


FIGURE 12. Qubit amplitude damping noise robustness. Amplitude damping noise is applied to a single qubit (left column) and uniformly to all qubits (right column). The markers show the maximal Bell score achieved using VQO while the dashed lines show the theoretical maximum. The top row plots the maximal n -local scores for maximally entangled state preparations and local qubit measurements while the bottom row plots the maximal n -local scores over arbitrary state preparations and measurements. The theoretical scores are given by (70) (top right), (66), and (68) for the n -local star and chain networks, respectively (top left), (72) (bottom left), and (63) (bottom right).

When nonmaximally entangled states are considered we use Proposition 8 of Appendix E1 to find that

$$S_{\text{CHSH}}^* (\tilde{\rho}_{\lambda^*, \gamma}) = \begin{cases} 2\sqrt{2(1 - \gamma)}, & \gamma \in [0, \frac{1}{2}] \\ 2, & \gamma \in [\frac{1}{2}, 1] \end{cases} \quad (71)$$

where $\lambda^* = \frac{1}{2}$ when $\gamma \in [0, \frac{1}{2}]$ and $\lambda^* = 1$ when $\gamma \in [\frac{1}{2}, 1]$. Thus, when amplitude damping is applied to a single qubit, maximally entangled state preparations are optimal on the range $\gamma \in [0, \frac{1}{2}]$ and the classical state preparation is optimal for $\gamma \in [\frac{1}{2}, 1]$. Then, by (37), we find

$$S_{n\text{-Star}}^* \left(\tilde{\rho}_{\lambda=\frac{1}{2}, \gamma} \bigotimes_{i=2}^n |\Phi^+\rangle\langle\Phi^+| \right) = \max \left\{ \sqrt{2}^{\frac{n-1}{n}}, \left(\sqrt{2}^{n-1} \sqrt{2(1 - \gamma)} \right)^{\frac{1}{n}} \right\}. \quad (72)$$

In Figs. 11 and 12, we see a close correspondence between theoretical results and those found using VQO. In our optimizations, we consider a range of preparation and measurement ansatzes (see Table 2 in Appendix B). We find that nonmaximally entangled state preparations and local qubit measurements are sufficient for maximal n -local violation.

2) SOURCE COLORED NOISE

Colored noise is typically found on the singlet state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ produced by parametric down conversion [62]. Colored noise is a two-qubit noise model representing depolarization on a preferred axis of the state

$$C_\gamma(\rho) = (1 - \gamma)\rho + \frac{\gamma}{2} (|\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|) \quad (73)$$

where $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. The Kraus operators for colored noise are

$$\begin{aligned} K_0 &= \sqrt{1-\gamma} \mathbb{I} \\ K_{\Psi^\pm, \Phi^\pm} &= \sqrt{\gamma/2} |\Psi^\pm\rangle\langle\Phi^\pm| \\ K_{\Psi^\pm, \Psi^\pm} &= \sqrt{\gamma/2} |\Psi^\pm\rangle\langle\Psi^\pm| \end{aligned} \quad (74)$$

where $|\Phi^\pm\rangle$ and $|\Psi^\pm\rangle$ constitute the Bell basis. Without loss of generality, we focus colored noise on $|\Psi^+\rangle$ and $|\Phi^+\rangle$ Bell states.

We begin by evaluating the CHSH score of the state $\tilde{\rho}_\gamma^\Psi = \mathcal{C}_\gamma(|\Psi^+\rangle\langle\Psi^+|)$. It follows that the correlation matrix is $T_{\tilde{\rho}_\gamma^\Psi} = \text{diag}(1-\gamma, 1-\gamma, -1)$. Then, we use (39) to obtain $S_{\text{CHSH}}^*(\tilde{\rho}_\gamma^\Psi) = 2\sqrt{1+(1-\gamma)^2}$. We note that this CHSH score is equivalent to (58), the score obtained for uniform qubit dephasing noise; hence, there is a direct correspondence between uniform qubit dephasing and colored noise applied to the $|\Psi^+\rangle\langle\Psi^+|$ state.

For uniform-source colored noise, the maximal n -local star score is given by (37) as

$$S_{n\text{-Star}}^*\left(\bigotimes_{i=1}^n \tilde{\rho}_\gamma^\Psi\right) = \sqrt{1+(1-\gamma)^2}. \quad (75)$$

Since the maximal singular value of $T_{\tilde{\rho}_\gamma^\Psi}$ is $\tau_0 = 1$, $S_{n\text{-Chain}}^*\left(\bigotimes_{i=1}^n \tilde{\rho}_\gamma^\Psi\right) = S_{2\text{-Star}}^*\left(\bigotimes_{i=1}^2 \tilde{\rho}_\gamma^\Psi\right)$. When colored noise is applied to a single source, we find that

$$S_{n\text{-Star}}^*\left(\tilde{\rho}_\gamma^\Psi \otimes_{i=2}^n |\Psi^+\rangle\langle\Psi^+|\right) = \left(\sqrt{2}^{n-1} \sqrt{1+(1-\gamma)^2}\right)^{\frac{1}{n}}. \quad (76)$$

Next, we evaluate the CHSH score of the state $\tilde{\rho}_\gamma^\Phi = \mathcal{C}_\gamma(|\Phi^+\rangle\langle\Phi^+|)$. It follows that the correlation matrix is $T_{\tilde{\rho}_\gamma^\Phi} = \text{diag}(1-\gamma, 1-2\gamma, -1)$. Then, we use (39) to obtain

$$S_{\text{CHSH}}^*(\tilde{\rho}_\gamma^\Phi) = \begin{cases} 2\sqrt{2(1-\gamma)^2}, & \gamma \in [0, \frac{2}{3}] \\ 2\sqrt{(1-\gamma)^2 + (1-2\gamma)^2}, & \gamma \in [\frac{2}{3}, 1]. \end{cases} \quad (77)$$

When colored noise is applied uniformly to $|\Phi^+\rangle\langle\Phi^+|$ sources, it follows from (37) that

$$S_{n\text{-Star}}^*\left(\bigotimes_{i=1}^n \tilde{\rho}_\gamma^\Phi\right) = \frac{1}{2} S_{\text{CHSH}}^*(\tilde{\rho}_\gamma^\Phi). \quad (78)$$

Furthermore, using (38) in the presence of uniform colored noise, we find that the n -local chain score is

$$S_{n\text{-Chain}}^*\left(\bigotimes_{i=1}^n \tilde{\rho}_\gamma^\Phi\right) = \frac{1}{2} S_{\text{CHSH}}^*(\tilde{\rho}_\gamma^\Phi) \prod_{i=2}^{n-1} \sqrt{\tau_0} \quad (79)$$

where $\tau_0 = 1-\gamma$. When colored noise is applied to a single source in a star, we find that

$$S_{n\text{-Star}}^*\left(\tilde{\rho}_\gamma^\Phi \otimes_{i=2}^n |\Psi^+\rangle\langle\Psi^+|\right) = \left(\sqrt{2}^{n-1} \frac{1}{2} S_{\text{CHSH}}^*(\tilde{\rho}_\gamma^\Phi)\right)^{\frac{1}{n}} \quad (80)$$

where the n -local chain score is given by the bilocal score.

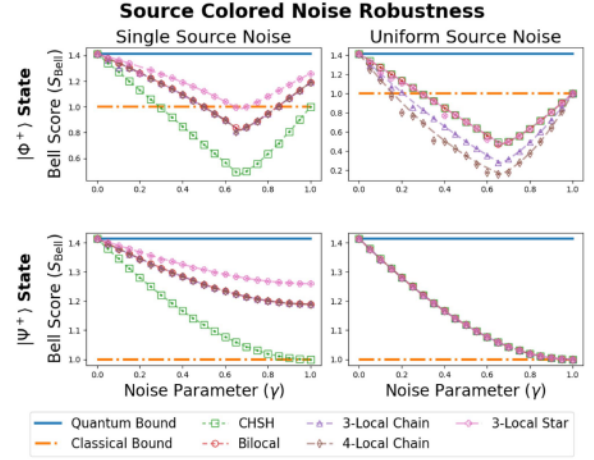


FIGURE 13. Colored noise is applied to a single source (left column) and uniformly to all sources (right column). The markers show the maximal Bell score achieved using VQO over $|\Phi^+\rangle$ state preparations and local qubit measurements (top row) and over $|\Psi^+\rangle$ state preparations and arbitrary measurements (bottom row). The dashed lines show the theoretically maximal score. (Top) Theoretical lines are given by (80) for single-source noise and by (78) and (79) for uniform-source noise. (Bottom) Theoretical lines are given by (76) and (75) for single-source and uniform-source noise.

In Fig. 13, we show that there is a close correspondence between the theoretical maximal n -local violations in the presence of colored noise and the maxima obtained using VQO. We consider a range of preparation and measurement ansatzes (see Table 2 in Appendix B). We find that $|\Psi^+\rangle$ is the optimal state preparation and that local qubit measurements are sufficient to achieve the maximal n -local violation in all cases.

3) BIASED DETECTOR ERRORS

We define a biased detector error as the classical postprocessing map

$$\mathbf{R}_\gamma = (1-\gamma)\mathbb{I}_2 + \gamma \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad (81)$$

where the fixed classical value of +1 is output whenever an error occurs. The biased detector error can be described by the POVM with elements

$$\hat{\Pi}_{+|x} = (1-\gamma)\Pi_{+|x} + \gamma\mathbb{I}_{2^M} \quad (82)$$

$$\hat{\Pi}_{-|x} = (1-\gamma)\Pi_{-|x} \quad (83)$$

where $\Pi_{+|x}$ and $\Pi_{-|x}$ constitute an M -qubit PVM. In Proposition 9 of Appendix E2, we show that the postprocessing map $\mathbf{R}_\gamma$ in (81) is nonunital and equivalent to a partial replacer channel

$$\mathcal{R}_{\gamma,x}(\rho^{A_j}) = (1-\gamma)\rho^{A_j} + \gamma\rho'_x \text{Tr}[\rho^{A_j}] \quad (84)$$

applied to the local quantum state ρ^{A_j} , where the replacer state ρ'_x is a density operator contained by the projective subspace of $\Pi_{+|x}$.

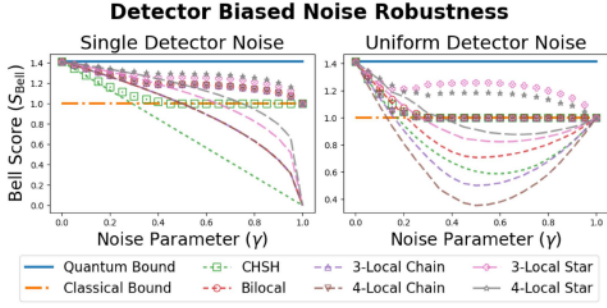


FIGURE 14. Biased detector noise robustness. (Left) Biased detector noise is applied to a single detector. (Right) Biased detector noise is applied uniformly to all detectors. The markers show the maximal Bell score achieved using VQO when using arbitrary state preparations. The dashed lines show the maximal Bell score achieved using VQO with respect to maximally entangled states.

In Fig. 14, we compare the VQO results of maximally entangled state preparations and arbitrary state preparations. We find significant improvements when arbitrary state preparations are considered, implying that nonmaximally entangled states are optimal for n -local violations in the presence of biased detector noise.

VI. VQO OF NON- N -LOCALITY ON QUANTUM HARDWARE

In this section, we demonstrate that our VQO techniques can maximize non- n -locality on IBM quantum computers, we discuss how to scale our methods, and we discuss the advantages of deploying our VQO methods on quantum network hardware.

A. VQO OF NON- N -LOCALITY ON NOISY IBM QUANTUM COMPUTERS

Using IBM quantum computers, we implement the VQO algorithm depicted in Fig. 4. Using (24), we maximize the CHSH score S_{CHSH} , the bilocal score $S_{2\text{-Star}}$, the trilocal chain score $S_{3\text{-Chain}}$, and the trilocal star score $S_{3\text{-Star}}$. Each optimization uses the quantum network ansatz depicted in Fig. 15 having the Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ prepared at each source and having local qubit measurements whose bases are free to rotate in the xz -plane of the Bloch sphere. This network ansatz reflects the capabilities of experimental setups of quantum networks [46], [47], [48], [49], [50] where the free parameters are rotations applied to each photon detection apparatus.

Our VQO results are shown in Fig. 16. In our optimizations, each circuit is evaluated using 6000 shots. Additionally, the initial settings used in each optimization run are selected from a uniform random distribution. The CHSH case was optimized using the five-qubit `ibmq_belem` device while the bilocal, trilocal chain, and trilocal star networks were optimized using the seven-qubit `ibmq_casablanca` and `ibmq_jakarta` devices. The CHSH plot shown in the upper-left of Fig. 16 aggregates data from 11 separate optimizations using a step size of $\eta = 0.12$. The bilocal plot shown in the upper-right plot of

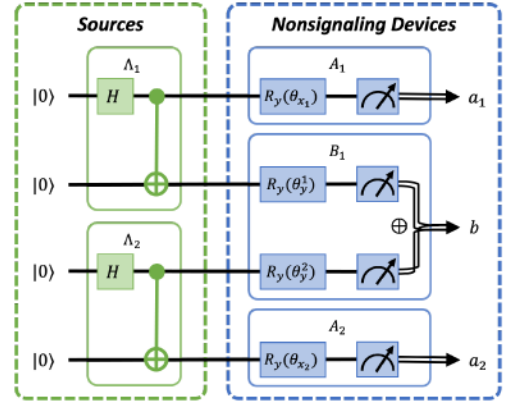


FIGURE 15. VQO hardware ansatz for bilocal network. When applying VQO on hardware, we consider a simple ansatz. Each source (green) prepares the Bell state $|\Phi^+\rangle$ using a Hadamard and CNOT gate. Each nonsignaling device (blue) applies a local rotation about the y -axis to each qubit before measurement in the computational basis. When a measurement device contains more than one local qubit, the XOR is taken to convert the bit string into a binary output.

Fig. 16 aggregates data from five optimizations and using a step size ranging from $\eta = 1.4$ to $\eta = 1.5$. The trilocal chain plot shown in the bottom-left plot of Fig. 16 aggregates data from six optimization using a step size ranging from $\eta = 1.6$ to $\eta = 2$. The trilocal star plot shown in the bottom-right plot of Fig. 16 aggregates data from five optimizations using a step size ranging from $\eta = 1.6$ to $\eta = 1.8$.

As a baseline measure of noise on the IBM quantum computers, we calculate the theoretical score using the optimal settings in the noiseless case described in Section IV-D. The noise on quantum hardware deteriorates the non- n -local correlations causing a separation from the noiseless quantum bound. In particular, the bottom-left plot shows the theoretical quantum violation of the trilocal chain inequality to be close the classical bound, especially when compared with the trilocal star network (bottom-right plot). This difference is due to the fact that $S_{n\text{-Star}}$ is more robust to noise than $S_{n\text{-Chain}}$ [41], [42]. Finally, the noise on the quantum computers is not constant; it can fluctuate throughout the day [88]. hence, the optimal settings in the noiseless case do not always produce the same value.

All plots in Fig. 16 show that the mean optimized score exceeds the classical bound. Thus, VQO finds non- n -local settings in all studied cases. In most cases, the error bars shrink as the optimization step increases, indicating convergence to an optimum. In the bilocal network optimization, the significant error bars on the final step are likely a result of the step size being too large. In the trilocal chain and star optimizations, the mean optimization score converges to the mean theoretical score showing that the optimization consistently finds the theoretical maximum. In the bilocal and CHSH optimizations, the mean optimization score does not reach the mean theoretical score. This is the result of some optimizations finding the local optimum of the classical bound.

In all plots, the max optimized score converges to a value consistent with the max theoretical score. In all cases, except

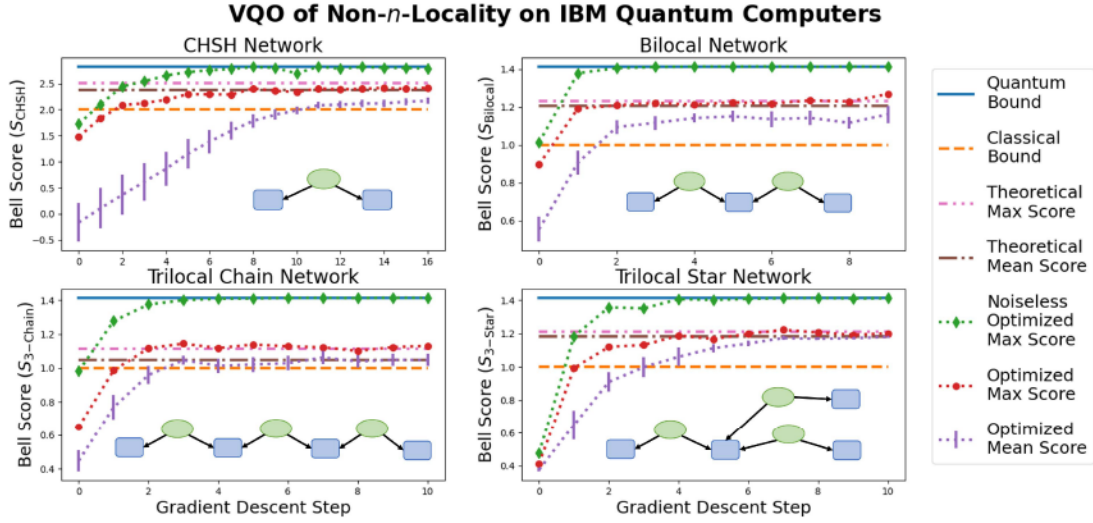


FIGURE 16. VQO of non- n -locality on IBM quantum computers. Non- n -locality is optimized in four different quantum networks. The x -axis shows the step of the gradient descent optimization while the y -axis shows the Bell score. The noiseless quantum bound is shown by the solid blue line and the classical bound is shown by the dashed orange line. For each Bell inequality, we aggregate data across several different optimization runs. At the end of each optimization, the theoretically optimal score is evaluated on noisy hardware to serve as a baseline. The mean theoretical score is shown by the dash-dotted brown line and the max theoretical score is shown by the dash-dot-dotted pink line. In each step, the max score across all optimization is shown by the dotted red line with circle markers while the mean score across all optimizations is shown by the dotted purple line with error bars showing the standard error. The dotted green line with diamond markers shows the noiseless optimal score obtained by running the settings for the maximal score in each step on a noiseless classical simulator.

for the CHSH optimization, an optimized score is found that exceeds the max theoretical score. While there is a statistical chance that the optimized score may be larger than the max theoretical score, it is also possible that our VQO framework may be finding optimal settings tailored for the quantum hardware. That is, there may be biases in the gate operations and measurement bases that VQO may be optimizing against. On the other hand, the theoretical settings are naïve to any such hardware biases.

Finally, we find that the settings optimized on the IBM hardware correspond to the optimal settings in the noiseless case. In reference to Fig. 16, when we take the settings for the maximal score in each optimization step and run them on a noiseless classical simulator, they maximally violate the Bell inequality in question. While this feature is remarkable, its value is questionable. We may be able to obtain the optimal settings on a large NISQ device, but the quantum computer will still output a nonoptimal answer that cannot be checked on a classical computer due to computing constraints. Further study of this optimization feature is important as there may exist an application where the optimal settings are valuable to know on their own.

B. SCALING VQO FOR PRACTICAL ADVANTAGES ON QUANTUM COMPUTERS

Scaling our VQO framework is largely limited by available quantum technology. As circuit network parameters are added, the number of circuit executions required by the parameter-shift rule grows polynomially [68], [78]. Furthermore, cost functions may require large numbers of circuit evaluations for large networks. For example, the n -local star

Bell inequality in (34) contains 2^{n+1} correlator terms that each require a quantum circuit to be run or differentiated. Unfortunately, additional circuit executions are a significant overhead due to the latency of remote execution, queue wait times, and serial circuit execution on a quantum computer. We mitigate this overhead by reserving quantum hardware and batching circuit executions; however, more can be done to scale the optimizations.

First, we need wide parallelization across quantum computers. In the worst case, a network cost function can require $|\mathcal{X}| = \prod_{j=1}^m |\mathcal{X}_j|$ unique circuit evaluations to construct the network behavior $\mathbf{P}_{\text{Net}}(\Theta)$ for all inputs $\vec{x} \in \mathcal{X}$. Furthermore, the parameter-shift rule scales the number of circuit executions polynomially. Fortunately, these circuits are independent and can be run in parallel. Thus, if we wish to scale VQO techniques, it will be exceedingly important to parallelize circuit executions across many NISQ devices.

Second, to overcome the latency of remote execution, we need classical and quantum hardware to be run in close proximity. The quantum computing industry is taking steps in this direction, for example, the Qiskit Runtime environment.

Third, larger NISQ devices will allow larger networks to be simulated. In this regime, we may find practical simulation advantages. However, NISQ devices have yet to show a simulation advantage [89], [90], [91], [92], [93], [94]. Nevertheless, NISQ devices are predicted to provide simulation advantages in the near-term [27], [95]. As an aside, we note that larger networks can be simulated using smaller quantum devices where an exponential increase in the number of circuit evaluations is accrued [96], [97]. Such methods are only feasible if wide parallelization across

NISQ devices can offset the exponential increase in circuit evaluations.

C. ADAPTING VQO TO QUANTUM NETWORK HARDWARE

In principle, the VQO framework depicted in Fig. 4 can be run on a quantum network rather than a quantum computer. The key requirement is that the network devices have free parameters that can be tuned continuously, and are, therefore, differentiable. In fact, hybrid optimization techniques on photonic systems have previously demonstrated the ability to maximize the violation of the CHSH inequality [28], [29]. Hence, our VQO framework could be extended to similar photonic implementations of quantum networks [46], [47], [48], [49], [50].

A key advantage of extending our VQO framework to quantum network hardware is that network protocols can be optimized against the noise inherent to the quantum network hardware. That is, the noise and biases on quantum computers may not accurately reflect the noise and biases on quantum network hardware. Hence, communications protocols optimized on quantum network hardware will be more robust because they are tailored specially for the hardware they run on. Therefore, the cost of noise tomography [18], [19], [20] can be avoided altogether.

Deploying VQO on quantum networks may provide a means of automating device integration and maintenance in quantum networks. Using VQO, quantum network devices may be able to automatically align photon polarization bases, maintain the communication capacity of channels, or set up device-independent protocols. Such self-organization amongst network devices may significantly reduce the manual work and expertise needed to build, scale, and maintain quantum networks.

VII. DISCUSSION

In this work, we introduce a hybrid, VQO framework for noisy quantum communication networks. We implement our optimization framework in the qNetVO Python package [69]. To showcase our optimization techniques, we maximize non- n -locality in noisy quantum networks simulated on both classical and quantum hardware.

In Section V, we demonstrate on a classical simulator that our VQO methods can find the maximal n -local violations in the presence of a wide range of noise models. For each noise model, we show a close correspondence between the theoretically maximal n -local violations derived in Appendix D and the maximal n -local violations obtained via VQO. That is, our VQO software leads to accurate theoretical results.

In Section VI, we demonstrate on a quantum computer that our VQO techniques can maximize non- n -locality in a quantum circuit run on noisy hardware. We show a close correspondence between the amount of non- n -locality expressed by the optimized circuit and the circuit designed using theoretically optimal parameters. Furthermore, we discuss how parallelization and larger quantum computers can be used to broadly improve the performance of VQO methods. In

principle, our methods can be deployed on quantum network hardware and used to optimize network protocols against the inherent hardware noise. Such engineering paradigms could be valuable for automating the setup and maintenance of quantum networking applications.

In practice, our VQO techniques are both convenient to apply and useful for gaining theoretical insights. For example, the numerical results posted in a preprint of this work led to the derivation of the maximal n -local star and chain scores in (37) and (38) [83]. Therefore, our VQO methods may be useful for other optimization problems in quantum networking. As noisy quantum computers scale and improve, our VQO techniques show promise of practical advantages on quantum computers. Thus, it is conceivable that VQO and other hybrid computing technologies will be essential to the design and development of quantum networking protocols.

A. FUTURE DIRECTIONS

Our framework has proven itself useful to the study of nonlocality in quantum networks having fewer than 20 qubits. We expect that our VQO framework should also be able to obtain interesting results when applied to optimization problems of maximizing entropic quantities, winning probabilities of multipartite games, or fidelity of quantum protocols. Additionally, our framework could be extended to allow local operations and classical communication protocols such as teleportation, superdense coding, and entanglement swapping, which would increase its applicability immensely.

While we obtain the most valuable results from a classical simulator, practical advantage on NISQ devices may be achievable if our methods are broadly parallelized, closely integrated with quantum hardware, and scaled to the largest available devices. On the noisy IBM quantum hardware, we found an interesting feature where the settings optimized on noisy hardware were optimal in the noiseless case. Investigating the extent to which this feature holds may prove to be important in the practical application of VQO of NISQ devices. Finally, one advantage of our VQO framework is that it can be implemented on network hardware to optimize network protocols against uncharacterized hardware noise. It is important to verify that VQO can, indeed, be deployed on quantum network hardware.

Finally, the optimal state preparations for nonlocality in the presence of unital and nonunital noise could be studied more broadly by considering additional Bell inequalities and noise models.

A. CODE AND DATA AVAILABILITY

Our VQO framework is released as a public Python package called qNetVO [69]. All numerics and data are found in a supplementary codebase on GitHub [70]. All optimizations were run within a few minutes to hours on a 2018 MacBook Pro having a 2.6-GHz 6-Core Intel Core i7 processor and 16-GB 2400-MHz DDR4 memory.

APPENDIX

A. NOTATION GLOSSARY

TABLE 1 Glossary of Mathematical Notation Used Throughout This Work

Symbol	Terminology	Definition
Λ_i	Source device	Labels the i^{th} source device that emits an M -qubit state. We treat Λ_i as the set of M qubits $\Lambda_i \subseteq [N]$ emitted from the source where $M = \Lambda_i $. Typically, we consider two-qubit states where $M = 2$ in this work.
A_j	Measurement device	Labels the j^{th} measurement device. We treat A_j as the set of M qubits $A_j \subseteq [N]$ measured by the device where $ A_j = M$.
L_k	Quantum link	Labels the k^{th} link between a source and measurement device. We treat L_k as the set of M qubits $L_k \subseteq [N]$ that are sent over the link where $ L_k = M$.
$\mathcal{X}$	Set of classical inputs for the network	For a network with m measurement devices that each have classical input, the input alphabet is $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_m$. All devices considered in this work have binary inputs such that $ \mathcal{X}_j = 2$ and $ \mathcal{X} = 2^m$. A classical input to the network is then $\vec{x} \in \mathcal{X}$.
$\mathcal{A}$	Set of classical outputs for the network	For a network with m measurement devices that each have a classical output, the output alphabet is $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_m$. All devices considered in this work have binary outputs such that $ \mathcal{A}_j = 2$ and $ \mathcal{A} = 2^m$. A classical value output from the network is then $\vec{a} \in \mathcal{A}$.
$ \psi^{\text{Net}}\rangle$	Network pure state	In total, a quantum network prepares the pure state $ \psi^{\text{Net}}\rangle = \bigotimes_{i=1}^n \psi^{\Lambda_i}\rangle$, where the state output from each source is defined on its own Hilbert space, $ \psi^{\Lambda_i}\rangle \in \mathcal{H}^{\Lambda_i}$.
ρ^{Net}	Network mixed state	In the noisy setting, a quantum network prepares a mixed state $\rho^{\text{Net}} = \bigotimes_{i=1}^n \rho^{\Lambda_i}$, where each source emits the state defined by the density operator $\rho^{\Lambda_i} \in D(\mathcal{H}^{\Lambda_i})$.
$\mathcal{N}_{\vec{\gamma}}^{\text{Net}}$	Network noise model	In total, the network's noise is modeled by the CPTP map $\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{k=1}^l \mathcal{N}_{\gamma_k}^{L_k}$. We also consider noise on sources as $\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{N}_{\gamma_i}^{\Lambda_i}$ and measurement devices as $\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{j=1}^m \mathcal{N}_{\gamma_j}^{A_j}$.
$\Pi_{\vec{a} \vec{x}}^{\text{Net}}$	Network measurement operator	In total, the network measures projector-valued measure $\Pi_{\vec{a} \vec{x}}^{\text{Net}} = \bigotimes_{j=1}^m \Pi_{a_j x_j}^{A_j}$, where $\sum_{\vec{a} \in \mathcal{A}} \Pi_{\vec{a} \vec{x}}^{\text{Net}} = \mathbb{I}_{2^N}$ for a network containing N qubits.
$\mathbf{P}_{\text{Net}}$	Network behavior	A column stochastic matrix, $\mathbf{P}_{\text{Net}} = \sum_{\vec{a} \in \mathcal{A}} \sum_{\vec{x} \in \mathcal{X}} P(\vec{a} \vec{x}) \vec{a}\rangle\langle\vec{x} $. For a quantum network, the conditional probabilities are calculated as $P(\vec{a} \vec{x}) = \text{Tr} \left[\Pi_{\vec{a} \vec{x}}^{\text{Net}} \mathcal{N}_{\vec{\gamma}}^{\text{Net}}(\rho^{\text{Net}}) \right]$.
$\mathbf{E}_{\vec{\gamma}}^{\text{Net}}$	Detector error model	A column stochastic post-processing map that adds error to a network behavior as $\mathbf{P}'_{\text{Net}} = \mathbf{E}_{\vec{\gamma}}^{\text{Net}} \mathbf{P}_{\text{Net}}$.
$\mathcal{Z}$	Set of qubit Measurement outcomes	For a network simulated on N qubits, the set of all N -bit measurement outcomes is $\mathcal{Z}$ where $ \mathcal{Z} = 2^N$. The output N -bit string is then $\vec{z} \in \mathcal{Z}$.
$U_{\vec{x}}^{\text{Net}}$	Network ansatz circuit	The unitary operator that simulates a quantum network's conditional probabilities $P(\vec{z} \vec{x}) = \langle \vec{z} U^{\text{Net}} 0 \dots 0 \rangle ^2$ where $U^{\text{Net}} = U_{\text{Meas}} U_{\mathcal{N}} U_{\text{Prep}}$ decomposes into preparation, noise, and measurement layers.
$\text{Cost}(\mathbf{P}_{\text{Net}})$	Network Cost Function	A generic function on a network behavior that is minimized during optimization. The cost function can be tailored to a wide range of optimization problems. Throughout this work, we minimize the cost function $\text{Cost}(\mathbf{P}_{\text{Net}}) = -S_{\text{Bell}}(\mathbf{P}_{\text{Net}})$ to maximize the Bell score.
$\mathcal{L}^{\text{Net}}$	Set of n -local network behaviors	The set of network behaviors whose probabilities $P(\vec{a} \vec{x})$ all can decompose as Eq. (3).
$\mathcal{Q}^{\text{Net}}(\mathcal{N}^{\text{Net}})$	Set of quantum network behaviors	The set of quantum network behaviors whose probabilities all can decompose as Eq. (4) for a fixed noise model $\mathcal{N}^{\text{Net}}$. In the noiseless case, we use $\mathcal{Q}^{\text{Net}} = \mathcal{Q}^{\text{Net}}(\text{id}^{\text{Net}})$.
$O_{\vec{x}}^{\text{Net}}$	Network observable	In total, the network measures the Hermitian observable $O_{\vec{x}}^{\text{Net}} = \bigotimes_{j=1}^m O_{x_j}^{A_j}$.
$S_{\text{Bell}}(\mathbf{P}) \leq \beta$	Bell inequality	An inequality that tightly bounds the n -local set. The quantity, $S_{\text{Bell}}(\mathbf{P})$ is referred to as the Bell score and β is the n -local bound. For convenience, we use $S(O_{\vec{x}}^{\text{Net}}, \mathcal{N}^{\text{Net}}, \rho^{\text{Net}})$ to denote the Bell score for a fixed observables, noise model, and state preparation.
$S_{\text{Bell}}^*(\rho^{\text{Net}})$	Maximal Bell score	The maximal Bell score that can be achieved for a fixed state preparation, $\max_{O_{\vec{x}}^{\text{Net}}} S_{\text{Bell}}(O_{\vec{x}}^{\text{Net}}, \text{id}^{\text{Net}}, \rho^{\text{Net}})$.
$\tilde{S}_{\text{Bell}}^*(\mathcal{N}^{\text{Net}})$	Maximal noisy Bell score	The maximal Bell score that can be achieved in the presence of a fixed noise model, $\max_{O_{\vec{x}}^{\text{Net}}, \rho^{\text{Net}}} S(O_{\vec{x}}^{\text{Net}}, \mathcal{N}^{\text{Net}}, \rho^{\text{Net}})$.
T_{ρ}	Two-qubit correlation matrix	The nonlocal content of a two-qubit state ρ is contained by its correlation matrix $T_{\rho} \in \mathbb{R}^{3 \times 3}$ that has elements $T_{\rho}^{(j,k)} = \text{Tr}[\sigma_j \otimes \sigma_k \rho]$ where $j, k \in \{x, y, z\}$.

B. PREPARE AND MEASURE ANSATZES

TABLE 2 Various Preparation and Measurement Ansatzes Used Throughout This Work

Ansatz name	Ansatz circuit	Number of parameters	Ansatz Description
Arbitrary M -Qubit Pure State Preparation		$ \vec{\phi} = 2^{M+1} - 2$	Parameterizes all M -qubit pure state preparations. See the <code>ArbitraryStatePreparation</code> method in PennyLane for implementation details.
$ \Phi^+\rangle$ State preparation		None	Prepares the state $ \Phi^+\rangle = (00\rangle + 11\rangle)/\sqrt{2}$.
$ \Psi^+\rangle$ State preparation		None	Prepares the state $ \Psi^+\rangle = (01\rangle + 10\rangle)/\sqrt{2}$.
Maximally Entangled State Preparation		$ \vec{\phi} = 3$	Parameterizes all 2-qubit maximally entangled state preparations using an arbitrary qubit rotation $Rot(\vec{\phi}) = R_z(\phi_1)R_y(\phi_2)R_z(\phi_3)$.
Nonmaximally Entangled State Preparation		$ \vec{\phi} = 2$	Parameterizes a family of separable through maximally entangled states as $ \psi\rangle = \cos(\phi_1/2) 00\rangle + \sin(\phi_1/2)e^{i\phi_2} 11\rangle$.
Arbitrary M -Qubit Projective Measurement		$ \vec{\theta} = 2^{2M} - 1$	Parameterizes all M -qubit projective measurements. See the <code>ArbitraryUnitary</code> method in PennyLane for implementation details.
M -Qubit local R_y measurement		$ \vec{\theta} = M$	Parameterizes all M -qubit projective measurements that decompose as $\Pi_{\vec{z} \vec{x}}^A = \bigotimes_{i=1}^M R_y(\phi_i) z^{q_i}\rangle\langle z^{q_i} $.
Arbitrary local qubit measurement		$ \vec{\theta} = 3M$	Parameterizes all M -qubit projective measurements that decompose as $\Pi_{\vec{z} \vec{x}}^A = \bigotimes_{i=1}^M Rot(\vec{\phi}_i) z^{q_i}\rangle\langle z^{q_i} $ where $Rot(\vec{\phi}) = R_z(\phi_1)R_y(\phi_2)R_z(\phi_3)$.

Notes: For each ansatz, we provide the circuit diagram, the number of tunable parameters, and a short description of the parameterization.

C. N-LOCAL VIOLATIONS IN STAR AND CHAIN NETWORKS WITH K CLASSICAL SOURCES

It is important to note that the maximal n -local violations in (37) and (38) are larger than those reported in previous works [85], [86], [87]

$$\hat{S}_{n\text{-Star}}^*(\rho^{\text{Net}}) = \sqrt{\prod_{i=1}^n \tau_{i,0}^{2/n} + \prod_{i=1}^n \tau_{i,1}^{2/n}} \quad (85)$$

$$\hat{S}_{n\text{-Chain}}^*(\rho^{\text{Net}}) = \sqrt{\prod_{i=1}^n \tau_{i,0} + \prod_{i=1}^n \tau_{i,1}} \quad (86)$$

that is, $\hat{S}_{n\text{-Star}}^*(\rho^{\text{Net}}) \geq \hat{S}_{n\text{-Star}}^*(\rho^{\text{Net}})$ and $\hat{S}_{n\text{-Chain}}^*(\rho^{\text{Net}}) \geq \hat{S}_{n\text{-Chain}}^*(\rho^{\text{Net}})$. The key distinction these two maximal violations is that (85) and (86) assume that multiqubit measurements are assumed to take the form $O_0^{A_j} = \bigotimes_{k=1}^{|A_j|} \sigma_z$ and $O_1^{A_j} = \bigotimes_{k=1}^{|A_j|} \sigma_x$ whereas (37) and (38) assume general local qubit observables $O_{x_j}^{A_j} = \bigotimes_{k=1}^{|A_j|} O_{x_j}^{q_k}$.

We now outline two extreme cases where (37) and (38) lead to n -local violations that are significantly larger than those of (85) and (86). First, we consider networks that have $k < n$ sources that emit the classical state $|00\rangle\langle 00|$. In this setting, we discuss two extreme examples where there is a large separation between the maximal n -local violations given by (37) and (38) and those given by (85) and (86).

Prediction 1: Consider an n -local star network that has k classical sources $\{\rho^{\Lambda_i} = |00\rangle\langle 00|\}_{i=1}^k$ and $n - k$ quantum sources that each produce the Bell state $\{\rho^{\Lambda_i} = |\Phi^+\rangle\langle \Phi^+|\}_{i=k+1}^n$. Since the correlation matrix $T_{|00\rangle\langle 00|} = \text{diag}(1, 0, 0)$, (85) predicts that $\hat{S}_{n\text{-Star}}^*(\{\rho^{\Lambda_i}\}_{i=1}^n) = 1$. Thus, if the n -local star network contains one or more classical sources, then its non- n -locality is broken.

Counterexample 1: Suppose that for inputs $x_i, y \in \{0, 1\}$, the observables $O_{x_i}^{A_i} = (1 - x_i)\sigma_z + x_i\sigma_x$ and $O_y^{B_i} = \sigma_z$ are applied to each classical source $\rho^{\Lambda_i} = |00\rangle\langle 00|$. The resulting two-body correlator is then $\langle O_{x_i}^{A_i} O_y^{B_i} \rangle_{\rho^{\Lambda_i}} = (1 - x_i)$. This measurement strategy is equivalent to the classical protocol where the interior node B_i outputs 1 with certainty while the exterior node A_i outputs 1 if $x_i = 0$, and otherwise, outputs ± 1 with equal probability. The remaining $n - k$ quantum sources implement the optimal measurements for the violation of the n -local star inequality [41]. Then, using (35) and the fact that all correlator terms with $x_i = 1$ vanish due to the uniformly random ± 1 output from A_i , we find that

$$I_{n,y} = \frac{1}{2^k} \left(\frac{1}{\sqrt{2}} \right)^{n-k} \quad (87)$$

from which it follows that

$$S_{n\text{-Star}} = 2 \left(\frac{1}{2^k} \left(\frac{1}{\sqrt{2}} \right)^{n-k} \right)^{\frac{1}{n}} = \left(\sqrt{2} \right)^{\frac{n-k}{n}} \geq 1. \quad (88)$$

Thus, we have described a partially classical strategy that outperforms the “maximal” score predicted in Prediction 1 by using (85).

Prediction 2: Consider an n -local chain network having exterior sources that prepare Bell states

$\rho^{\Lambda_1} = \rho^{\Lambda_n} = |\Phi^+\rangle\langle \Phi^+|$ and interior sources that prepare the classical states $\{\rho^{\Lambda_i} = |00\rangle\langle 00|\}_{i=2}^{n-1}$. Since the correlation matrix $T_{|00\rangle\langle 00|} = \text{diag}(1, 0, 0)$, (86) predicts that $\hat{S}_{n\text{-Chain}}^*(\{\rho^{\Lambda_i}\}_{i=1}^n) = 1$. Thus, we predict that if the n -local chain network contains one or more classical sources, it cannot generate non- n -local correlations.

Counterexample 2: Consider an n -local chain network having $N = 2n$ qubits where sources $\Lambda_1 = (q_1, q_2)$ and $\Lambda_n = (q_{N-1}, q_N)$ are quantum and the remaining sources $\Lambda_i = (q_{2i}, q_{2i+1})$. If the Pauli observable $O_y^{q_j} = \sigma_z$ is applied to each qubit prepared at a classical source, then

$$\begin{aligned} \langle O_{\vec{x},y}^{\text{Net}} \rangle_{\rho^{\text{Net}}} &= \langle O_{x_1}^{q_1} O_y^{q_2} \rangle_{|\Phi^+\rangle\langle \Phi^+|} \left(\prod_{j=3}^{N-2} \langle \sigma_z^{q_j} \rangle_{|0\rangle\langle 0|} \right) \\ &\quad \times \langle O_y^{q_{N-1}} O_{x_N}^{q_N} \rangle_{|\Phi^+\rangle\langle \Phi^+|} \end{aligned} \quad (89)$$

$$= \langle O_{x_1}^{q_1} O_y^{q_2} O_y^{q_{N-1}} O_{x_N}^{q_N} \rangle_{|\Phi^+\rangle\langle \Phi^+|} \quad (90)$$

$$= \langle O_{x_1}^{A_1} O_y^{B_1} O_y^{B_n} O_{x_n}^{A_n} \rangle_{|\Phi^+\rangle\langle \Phi^+|} \quad (91)$$

where the inputs and observables are relabeled in the last line to yield the Bilocal network correlator. It follows that if the optimal Bilocal measurement strategy is applied on qubits (q_1, q_2, q_{N-1}, q_N) , then the maximal bilocal score is $\hat{S}_{n\text{-Chain}}^* = \sqrt{2}$. Thus, we have described a partially classical strategy that outperforms the “maximal” n -local score predicted in Prediction 2 by using (86). Furthermore, this example is not limited to the classical state $|00\rangle\langle 00|$, but can be extended to any state ρ^{Λ_i} that satisfies $\text{Rank}(T_{\rho^{\Lambda_i}})$ where the important requirement for this protocol is that expectation of the parity $\prod_{j=3}^{2n-2} \langle O^{q_j} \rangle_{\rho^{\Lambda_i}} = \pm 1$ is constant.

D. MAXIMAL N-LOCAL VIOLATIONS IN THE PRESENCE OF UNITAL NOISE

In this section, we provide a theoretical analysis of the maximal n -local violations in star and chain networks that have unital noise. In the following sections, we derive the maximal n -local star and chain scores $\hat{S}_{n\text{-Star}}^*(\mathcal{N}^{\text{Net}})$ and $\hat{S}_{n\text{-Chain}}^*(\mathcal{N}^{\text{Net}})$ that can be obtained in the presence of unital channels including qubit depolarizing, qubit dephasing, source depolarizing, and detector white noise.

In the qubit case, the theory of unital channels is well known [98]. A unital qubit channel $\mathcal{U}$ can be expressed in Pauli basis as a 3×3 matrix

$$M_{\mathcal{U}} = \sum_{i,j \in \{x,y,z\}} \text{Tr}[\sigma_j \mathcal{U}(\sigma_i)] |j\rangle\langle i| \quad (92)$$

where the operator $M_{\mathcal{U}}$ is applied to the Bloch vector $\vec{s}$ of the input state ρ as $\vec{r} = M_{\mathcal{U}} \vec{s}$ to produce $\vec{r}$ the Bloch vector of the output states $\mathcal{U}(\rho)$. An important property of unital qubit channels is that the matrix $M_{\mathcal{U}}$ can always be diagonalized using rotations $R_{\text{in}}, R_{\text{out}} \in SO(3)$ such that

$$R_{\text{out}} M_{\mathcal{U}} R_{\text{in}} = \text{diag}(u_0, u_1, u_2) = M_{\mathcal{U}}^* \quad (93)$$

where the singular values are bound as $1 \geq u_x, u_y, u_z \geq 0$ and can be permuted along the diagonal of $M_{\mathcal{U}^A}$ without loss of generality.

By the isometry between $SO(3)$ and $SU(2)$, the rotation operators R_{in} and R_{out} correspond directly to unitaries V_{in} and $V_{\text{out}} \in SU(2)$ applied to $\mathcal{U}$ as

$$\mathcal{U}^*(\rho) = V_{\text{out}} \mathcal{U}(V_{\text{in}} \rho V_{\text{in}}^\dagger) V_{\text{out}}^\dagger \quad (94)$$

where $u_i \sigma_i = \mathcal{U}(\sigma_i)$ is an *eigenoperator* of the unital channel $\mathcal{U}$. We use this property in the following theorem to derive the maximal CHSH score that can be obtained in the presence of unital qubit noise and projective measurements.

Lemma 1: Consider a CHSH scenario having a unital qubit noise model $\mathcal{N}^{\text{Net}} = \mathcal{U}^A \otimes \mathcal{U}^B$. The maximal noisy CHSH score is achieved using a maximally entangled state preparation $|\psi\rangle^\Lambda = V^A \otimes V^B |\Phi^+\rangle$, that is,

$$\tilde{S}_{\text{CHSH}}(\mathcal{U}^A \otimes \mathcal{U}^B) = S_{\text{CHSH}}^*(\mathcal{U}^A \otimes \mathcal{U}^B(|\psi\rangle\langle\psi|^\Lambda)). \quad (95)$$

Proof: Consider a two-qubit mixed state $\rho \in D(\mathcal{H}^A \otimes \mathcal{H}^B)$ and local qubit PVM observables, $O_x^A = \vec{a}_x \cdot \vec{\sigma}$ and $O_y^B = \vec{b}_y \cdot \vec{\sigma}$, where $\vec{a}_x$ are $\vec{b}_y$ Bloch vectors of unit length $|\vec{a}_x| = |\vec{b}_y| = 1$. The CHSH score is $S_{\text{CHSH}} = \sum_{y=0}^1 \langle O_y^{AB} \rangle_\rho$, where

$$O_y^{AB} = (O_0^A + (-1)^y O_1^A) \otimes O_y^B \quad (96)$$

$$= (\vec{a}_0 + (-1)^y \vec{a}_1) \cdot \vec{\sigma} \otimes \vec{b}_y \cdot \vec{\sigma}. \quad (97)$$

Let $\tilde{\rho}^\Lambda = \mathcal{U}^A \otimes \mathcal{U}^B(\rho^\Lambda)$, then

$$\langle O_y^{AB} \rangle_{\tilde{\rho}^\Lambda} = \text{Tr} [O_y^{AB} \mathcal{U}^A \otimes \mathcal{U}^B(\rho^\Lambda)] \quad (98)$$

$$= \text{Tr} [\mathcal{U}^{\dagger A} \otimes \mathcal{U}^{\dagger B} (O_y^{AB}) \rho^\Lambda] \quad (99)$$

$$= \text{Tr} [\tilde{O}_y^{AB} \rho^\Lambda] \quad (100)$$

where the adjoint of a unital channel is unital. The unital qubit channels $\mathcal{U}^A$ and $\mathcal{U}^B$ can be diagonalized as in (94) such that $M_{\mathcal{U}^A} = \text{diag}(u_x^A, u_y^A, u_z^A)$, where $u_x^A \geq u_y^A \geq u_z^A$ and similarly for $M_{\mathcal{U}^B}$. Then, the correlation operator for observable O_y^{AB} is expressed in terms of Pauli vectors as $T_{O_y^{AB}} = |\vec{a}_0 + (-1)^y \vec{a}_1\rangle\langle\vec{b}_y|$ and the correlation operator of the noisy observable $\tilde{O}_y^{AB}$ is

$$T_{\tilde{O}_y^{AB}} = M_{\mathcal{U}^A} |\vec{a}_0 + (-1)^y \vec{a}_1\rangle\langle\vec{b}_y| M_{\mathcal{U}^B} \quad (101)$$

$$= |\vec{a}_0^i + (-1)^y \vec{a}_1^i\rangle\langle\vec{b}_y^i| \quad (102)$$

where $\vec{a}_x = (u_x^A a_x, u_y^A a_y, u_z^A a_z)$ and $\vec{b}_y = (u_x^B b_x, u_y^B b_y, u_z^B b_z)$. We make the important observation that $\text{Rank}(T_{\tilde{O}_y^{AB}}) \leq 1$, where the rank is 0 in only the case where the unital channel is fully depolarizing $u_j = 0$ for all $j \in \{x, y, z\}$. Furthermore, since $\text{Tr}[\tilde{O}_{y=0}^{AB} \tilde{O}_{y=1}^{AB}] = 0$, the operators are linearly independent, which implies that $\text{Rank}(T_{\tilde{O}_{y=0}^{AB}} + T_{\tilde{O}_{y=1}^{AB}}) \leq 2$.

Next, Proposition 1 from Zhang et al. [64] states that for any 4×4 matrix $M = \sum_{i,j=1}^3 t_{i,j} \sigma_i \otimes \sigma_j$ whose correlation

operator $T_M = \sum_{i,j=1}^3 t_{i,j} |i\rangle\langle j|$ satisfies $\text{Rank}(T_M) \leq 2$, the quantity $\text{Tr}[M\rho]$ is maximized by a maximally entangled state $\rho = |\psi\rangle\langle\psi|$, where $|\psi\rangle = V^A \otimes V^B |\Phi^+\rangle$. It follows from Proposition 1 that there exists a maximally entangled state $\rho = |\psi\rangle\langle\psi|$ that simultaneously maximizes $\text{Tr}[\tilde{O}_{y=0}^{AB}\rho]$, $\text{Tr}[\tilde{O}_{y=1}^{AB}\rho]$, and $\text{Tr}[(O_{y=0}^{AB} + O_{y=1}^{AB})\rho]$ because the correlation operators $T_{\tilde{O}_{y=0}^{AB}}$ and $T_{\tilde{O}_{y=1}^{AB}}$ are orthogonal and each rank-one. That is, the correlation operator for a maximally entangled state is diagonalized as $T_\rho = \text{diag}(1, -1, 1)$; hence, ρ can simultaneously maximize two observables if their correlation operators are orthogonal as $\text{Tr}[T_{\tilde{O}_{y=0}^{AB}} T_{\tilde{O}_{y=1}^{AB}}] = 0$. Since there exists a maximally entangled state, $|\psi\rangle\langle\psi|^\Lambda$, that maximizes the expectation of $\langle O_y^{AB} \rangle_{\rho^\Lambda}$ for both $y = 0, 1$, maximally entangled states are optimal for achieving the maximal CHSH score in the presence of unital qubit noise. ■

Lemma 1 shows that maximally entangled states achieve the maximal CHSH violation in the presence of unital qubit noise. However, maximally entangled states are not guaranteed to be optimal when multiqubit unital channels are considered at sources or measurement as in Fig. 6(a) and (c). Since the maximal n -local star and chain scores in (37) and (38) factor as the maximal CHSH score on each independent source, Lemma 1 can be extended to star and chain networks, allowing us to derive their respective maximal n -local scores in the presence of unital qubit noise.

Theorem 1: Consider an n -local star network having local qubit PVM measurements and a unital qubit noise model on each source $\mathcal{N}^{\Lambda_i} = \mathcal{U}^{\Lambda_i} \otimes \mathcal{U}^{B_i}$. The maximal noisy n -local star and chain scores are

$$\tilde{S}_{n\text{-Star}}^*(\mathcal{N}^{\text{Net}}) = \frac{1}{2} \prod_{i=1}^n \tilde{S}_{\text{CHSH}}^*(\mathcal{U}^{\Lambda_i} \otimes \mathcal{U}^{B_i})^{\frac{1}{n}} \quad (103)$$

$$\tilde{S}_{n\text{-Chain}}^*(\mathcal{N}^{\text{Net}}) = \tilde{S}_{2\text{-Star}}^*(\mathcal{N}^{\Lambda_1} \otimes \mathcal{N}^{\Lambda_n}) \prod_{i=2}^{n-1} \sqrt{u_0^{\Lambda_i} u_0^{B_i}} \quad (104)$$

where $\tilde{S}_{\text{CHSH}}^*(\mathcal{U}^{\Lambda_i} \otimes \mathcal{U}^{B_i}) = 2\sqrt{(u_0^{\Lambda_i} u_0^{B_i})^2 + (u_1^{\Lambda_i} u_1^{B_i})^2}$ and $u_0 \geq u_1$ are the two largest singular values of matrix $M_{\mathcal{U}}$ as in (92).

Proof: We begin by deriving the maximal noisy CHSH score $\tilde{S}_{\text{CHSH}}^*(\mathcal{U}^A \otimes \mathcal{U}^B)$ for a single source. By Lemma 1, the optimal state preparation for CHSH violation is a maximally entangled state $|\psi\rangle = V^A \otimes V^B |\Phi^+\rangle$, where the Bell state's correlation operator is $T_{|\Phi^+\rangle\langle\Phi^+|} = \text{diag}(1, -1, 1)$. Since we allow qubit unitary freedom on the state preparation and measurement, we can diagonalize the unital qubit channels in the Pauli basis as $M_{\mathcal{U}^A} = \text{diag}(u_1^A, u_2^A, u_0^A)$ and $M_{\mathcal{U}^B} = \text{diag}(u_1^B, u_2^B, u_0^B)$, where $u_0 \geq u_1 \geq u_2$. Then, $\rho^\Lambda = \mathcal{U}^A \otimes \mathcal{U}^B(|\Phi^+\rangle\langle\Phi^+|)$ can be expressed as

$$T_{\rho^\Lambda} = M_{\mathcal{U}^A} T_{|\Phi^+\rangle\langle\Phi^+|} M_{\mathcal{U}^B}^T \quad (105)$$

$$= \text{diag}(u_1^A u_1^B, -u_2^A u_2^B, u_0^A u_0^B). \quad (106)$$

Noting that $\tau_j = u_j^A u_j^B$, we use (39) to find

$$S_{\text{CHSH}}^*(\mathcal{U}^A \otimes \mathcal{U}^B) = 2\sqrt{(u_0^A u_0^B)^2 + (u_1^A u_1^B)^2}. \quad (107)$$

Then, inserting $S_{\text{CHSH}}^*(\mathcal{U}^A \otimes \mathcal{U}^B)$ into (37) and (38), we recover (103) and (104), respectively. Finally, since $T_{|\Phi^+\rangle\langle\Phi^+|}$, $M_{\mathcal{U}}^A$, and $M_{\mathcal{U}}^B$ are diagonal, the singular values of T_{ρ^A} and the resulting Bell scores are maximized. ■

1) MAXIMAL N-LOCAL VIOLATIONS FOR QUBIT DEPOLARIZING NOISE

Proposition 1: Consider a star or chain quantum network having a qubit depolarizing noise model, $\mathcal{D}_{\vec{v}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{D}_{v_i^{A_i}}^{A_i} \otimes \mathcal{D}_{v_i^{B_i}}^{B_i}$. The maximal n -local star and chain scores are

$$\tilde{S}_{n\text{-Star}}^*(\mathcal{D}_{\vec{v}}^{\text{Net}}) = \sqrt{2} \left(\prod_{i=1}^n v_i^{A_i} v_i^{B_i} \right)^{\frac{1}{n}} \quad (108)$$

$$\tilde{S}_{n\text{-Chain}}^*(\mathcal{D}_{\vec{v}}^{\text{Net}}) = \sqrt{2} \left(\prod_{i=1}^n v_i^{A_i} v_i^{B_i} \right)^{\frac{1}{2}}. \quad (109)$$

Proof: A qubit depolarizing channel is diagonalized in the Pauli basis as $M_{\mathcal{D}_v} = \text{diag}(v, v, v)$. (108) and (109) follow directly from Theorem 1. ■

2) MAXIMAL N-LOCAL VIOLATIONS FOR QUBIT DEPHASING NOISE

Proposition 2: Consider a star or chain quantum network having a qubit dephasing noise model $\mathcal{P}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{P}_{\gamma_i^{A_i}}^{A_i} \otimes \mathcal{P}_{\gamma_i^{B_i}}^{B_i}$. The maximal n -local star and chain scores are

$$\tilde{S}_{n\text{-Star}}^*(\mathcal{P}_{\vec{\gamma}}^{\text{Net}}) = \prod_{i=1}^n (1 + (1 - \gamma_i^{A_i})(1 - \gamma_i^{B_i}))^{\frac{1}{2n}} \quad (110)$$

$$\tilde{S}_{n\text{-Chain}}^*(\mathcal{P}_{\vec{\gamma}}^{\text{Net}}) = \tilde{S}_{2\text{-Star}}^*(\mathcal{P}_{\gamma_1^{A_1}}^{A_1} \otimes \mathcal{P}_{\gamma_1^{B_1}}^{B_1} \otimes \mathcal{P}_{\gamma_n^{A_n}}^{A_n} \otimes \mathcal{P}_{\gamma_n^{B_n}}^{B_n}). \quad (111)$$

Proof: A qubit dephasing channel is diagonalized in the Pauli basis as $M_{\mathcal{P}_\gamma} = \text{diag}(\sqrt{1-\gamma}, \sqrt{1-\gamma}, 1)$. Equations (110) and (111) follow directly from Theorem 1. ■

3) MAXIMAL N-LOCAL VIOLATIONS FOR DETECTOR WHITE NOISE

Proposition 3: The postprocessing map modeling white noise detector errors

$$\mathbf{W}_\gamma = (1 - \gamma)\mathbb{I} + \frac{\gamma}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (112)$$

is a unital channel because it is equivalent to the M -qubit depolarizing channel

$$\mathcal{D}_{(1-\gamma)}(X) = (1 - \gamma)X + \frac{\gamma}{2^M} \mathbb{I}_{2^M} \text{Tr}[X] \quad (113)$$

acting upon the detector's local projective measurement $\Pi_{\pm|x}$ as

$$\mathbf{W}_{\pm|x,\gamma} = \mathcal{D}_{(1-\gamma)}^\dagger(\Pi_{\pm|x}) = (1 - \gamma)\Pi_{\pm|x} + \frac{1}{2} \mathbb{I}_M. \quad (114)$$

Proof: The M -qubit depolarizing channel $\mathcal{D}_{(1-\gamma)}(X) = (1 - \gamma)X + \frac{\gamma}{2^M} \mathbb{I}_M \text{Tr}[X]$ has Hermitian Kraus operators $K_i = K_i^\dagger$; thus, $\mathcal{D}_{(1-\gamma)}(X) = \mathcal{D}_{(1-\gamma)}^\dagger(X)$. Then

$$\mathcal{D}_{(1-\gamma)}^\dagger(\Pi_{\pm|x}) = (1 - \gamma)\Pi_{\pm|x} + \frac{\gamma}{2^M} \mathbb{I}_M \text{Tr}[\Pi_{\pm|x}] \quad (115)$$

$$= (1 - \gamma)\Pi_{\pm|x} + \frac{1}{2} \mathbb{I}_M \quad (116)$$

$$= \mathbf{W}_{\pm|x,\gamma}. \quad (117)$$

Proposition 4: Consider a quantum network with a detector white noise model $\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{j=1}^m \mathcal{D}_{(1-\gamma_j)}^{A_j}$. The maximal noisy n -local scores for the star and chain networks are

$$\tilde{S}_{n\text{-Star}}^*(\mathcal{N}_{\vec{\gamma}}^{\text{Net}}) = \sqrt{2} \left(\prod_{j=1}^{n+1} (1 - \gamma_j) \right)^{1/n} \quad (118)$$

$$\tilde{S}_{n\text{-Chain}}^*(\mathcal{N}_{\vec{\gamma}}^{\text{Net}}) = \sqrt{2} \sqrt{\prod_{j=1}^{n+1} (1 - \gamma_j)}. \quad (119)$$

Proof: As given by (52), white noise on each detector is modeled by a POVM with elements $\mathbf{W}_{\pm|x,\gamma} = (1 - \gamma)P_{\pm|x} + \gamma \mathbf{W}_{\pm|x}$, where $P_{\pm|x}$ are projectors onto even and odd parity subspaces and $\mathbf{W}_{\pm|x}$ is a white noise POVM. The corresponding dichotomic observable having ± 1 eigenvalues is constructed as

$$\mathbf{W}_{x_j,\gamma_j}^{A_j} = \mathbf{W}_{+|x_j,\gamma_j} - \mathbf{W}_{-|x_j,\gamma_j} \quad (120)$$

$$= (1 - \gamma_j)(P_{+|x} - P_{-|x}) \quad (121)$$

$$= (1 - \gamma_j)O_{x_j}^{A_j} \quad (122)$$

where $O_{x_j}^{A_j}$ is the observable in the noiseless case. By (29), the m -partite correlators are written

$$\langle \mathbf{W}_{\vec{x},\vec{\gamma}}^{\text{Net}} \rangle_{\rho^{\text{Net}}} = \text{Tr} \left[\left(\bigotimes_{j=1}^m (1 - \gamma_j) O_{x_j}^{A_j} \right) \rho^{\text{Net}} \right] \quad (123)$$

$$= \prod_{j=1}^m (1 - \gamma_j) \langle O_{\vec{x}}^{\text{Net}} \rangle_{\rho^{\text{Net}}} \quad (124)$$

where $m = n + 1$ is the number of measurement nodes in the star and chain networks. Since $|\langle O^{\text{Net}} \rangle| \leq 1$, it must hold that $|\langle \mathbf{W}_{\vec{x},\vec{\gamma}}^{\text{Net}} \rangle| \leq \prod_{j=1}^m (1 - \gamma_j)$, where the upper bound is only achieved when the optimal strategy non- n -locality is used. For the noisy quantum network, the quantity $I_{n,y}$ defined in (35) is bound as

$$|I_{n,y}| \leq \frac{1}{\sqrt{2^n}} \prod_{j=1}^{n+1} (1 - \gamma_j) \frac{1}{2^n}. \quad (125)$$

Hence, the maximal noisy n -local star score is

$$\tilde{S}_{n\text{-Star}}^*(\mathcal{N}_{\vec{\gamma}}^{\text{Net}}) = \sqrt{2} \left(\prod_{j=1}^{n+1} (1 - \gamma_j) \right)^{1/n} \quad (126)$$

and achieved using the optimal noiseless strategy for the star network [41]. Likewise, if the optimal strategy for the chain inequality is used [42], then

$$|I_{2,y}^*| = \frac{1}{2} \prod_{j=1}^{n+1} (1 - \gamma_j) \quad (127)$$

and the maximal noisy n -local chain score is

$$\tilde{S}_{n\text{-Chain}}^*(\mathcal{N}_{\vec{\gamma}}^{\text{Net}}) = \sqrt{2} \sqrt{\prod_{j=1}^{n+1} (1 - \gamma_j)}. \quad (128)$$

■

E. MAXIMAL N -LOCAL VIOLATIONS IN THE PRESENCE OF NONUNITAL NOISE

In this section, we derive maximal violations of n -locality in the presence of nonunital noise models. We begin with a general discussion on nonunital qubit channels. It is useful to note that upon local unitary rotations, as in (94), a qubit nonunital channel can be represented in the Pauli basis as

$$\mathbb{T}_{\mathcal{N}_{\gamma}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & u_x & 0 & 0 \\ 0 & 0 & u_y & 0 \\ t_z & 0 & 0 & u_z \end{pmatrix} \quad (129)$$

where we write the 3×3 matrix block on the bottom right as $M_{\mathcal{N}_{\gamma}} = \text{diag}(u_x, u_y, u_z)$ similarly to the case of unital channels in (92).

Proposition 5: Consider two nonunital qubit channels $\mathcal{N}^A$ and $\mathcal{N}^B$ of the form in (129) acting upon an arbitrary nonmaximally entangled state prepared as $\rho_{\lambda} = |\psi_{\lambda}\rangle\langle\psi_{\lambda}|$, where $|\psi_{\lambda}\rangle = \sqrt{\lambda}|00\rangle + \sqrt{1-\lambda}|11\rangle$ and $\lambda \in [0, 1]$. Given the noisy two-qubit state $\tilde{\rho}_{\lambda} = \mathcal{N}^A \otimes \mathcal{N}^B(\rho_{\lambda})$, the maximal CHSH score is

$$S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda^*, \gamma}) = \begin{cases} 2\sqrt{(u_x^A u_x^B)^2 + (u_y^A u_y^B)^2}, & \gamma \in [0, \gamma_c] \\ 2 \left((u_x^A u_x^B)^2 + (t_z^A t_z^B + u_z^A u_z^B)^2 - \frac{(t_z^A t_z^B + u_z^A u_z^B)^2 (t_z^A u_z^B + u_z^A t_z^B)^2}{(t_z^A u_z^B + u_z^A t_z^B)^2 - (u_x^A u_x^B)^2} \right)^{\frac{1}{2}}, & e\gamma \in [\gamma_c, 1] \end{cases} \quad (130)$$

where $\lambda^* = \arg \max_{\lambda \in [0, 1]} S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda})$, γ_c is the crossover noise parameter where the two cases in (130) are equal, and $|u_x| \geq |u_y|$ without loss of generality. Furthermore, $\lambda^* = \frac{1}{2}$ when $\gamma \in [0, \gamma_c]$ and when $\gamma \in [\gamma_c, 1]$,

$$\lambda^* = \frac{1}{2} \left(1 - \frac{(t_z^A t_z^B + u_z^A u_z^B)(t_z^A u_z^B + u_z^A t_z^B)}{(t_z^A u_z^B + u_z^A t_z^B)^2 - (u_x^A u_x^B)^2} \right). \quad (131)$$

Proof: In the Pauli basis, $\rho_{\lambda} = |\psi_{\lambda}\rangle\langle\psi_{\lambda}|$ is expressed

$$T_{\rho_{\lambda}} = \begin{pmatrix} 1 & 0 & 0 & 2\lambda - 1 \\ 0 & 2\sqrt{\lambda}\sqrt{1-\lambda} & 0 & 0 \\ 0 & 0 & -2\sqrt{\lambda}\sqrt{1-\lambda} & 0 \\ 2\lambda - 1 & 0 & 0 & 1 \end{pmatrix}. \quad (132)$$

The qubit nonunital channels $\mathcal{N}^A$ and $\mathcal{N}^B$ are applied as $T_{\tilde{\rho}_{\lambda}} = \mathbb{T}_{\mathcal{N}^A} T_{\rho_{\lambda}} \mathbb{T}_{\mathcal{N}^B}$, where the singular values of $T_{\tilde{\rho}_{\lambda}}$ are

$$\tau_x = 2u_x^A u_x^B \sqrt{\lambda(1-\lambda)} \quad (133)$$

$$\tau_y = 2u_y^A u_y^B \sqrt{\lambda(1-\lambda)} \quad (134)$$

$$\tau_z = t_z^A t_z^B + u_z^A u_z^B + (2\lambda - 1)(t_z^A u_z^B + u_z^A t_z^B). \quad (135)$$

We find $\max_{\lambda \in [0, 1]} S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda})$ by checking the singular value pairs (τ_x, τ_y) and (τ_x, τ_z) where it is assumed without loss of generality that $|\tau_x| \geq |\tau_y|$. For the pair (τ_x, τ_y) , we find

$$\begin{aligned} \max_{\lambda \in [0, 1]} \tau_x^2 + \tau_y^2 &= \max_{\lambda \in [0, 1]} 4\lambda(1-\lambda) \left((u_x^A u_x^B)^2 + (u_y^A u_y^B)^2 \right) \\ &= (u_x^A u_x^B)^2 + (u_y^A u_y^B)^2 \end{aligned} \quad (136)$$

where $\lambda^* = \frac{1}{2}$ is the optimal entanglement parameter. Inserting (136) into (39) recovers the $\gamma \in [0, \gamma_c]$ case of (130). Next, we consider the singular value pair (τ_x, τ_z) , where

$$\begin{aligned} \max_{\lambda \in [0, 1]} \tau_x^2 + \tau_z^2 &= \max_{\lambda \in [0, 1]} 4\lambda(1-\lambda) (u_x^A u_x^B)^2 \\ &\quad + (t_z^A t_z^B + u_z^A u_z^B + (2\lambda - 1)(t_z^A u_z^B + u_z^A t_z^B))^2. \end{aligned} \quad (137)$$

To maximize, we take the derivative of the RHS with respect to λ and set the result equal to 0. We solve for the critical point λ^* and find it to be given by (131). Inserting λ^* into (137) and rearranging the terms recovers the $\gamma \in [\gamma_c, 1]$ case of (130). ■

1) MAXIMAL N -LOCAL VIOLATIONS FOR QUBIT AMPLITUDE DAMPING NOISE

Proposition 6: Consider an n -local star network having maximally entangled states $\rho^{\text{Net}} = \bigotimes_{i=1}^n |\psi\rangle\langle\psi|$, where $|\psi\rangle = U^A \otimes U^B |\Phi^+\rangle$, local qubit PVM measurements, and an amplitude damping channel applied to each qubit as $\mathcal{N}_{\vec{\gamma}}^{\text{Net}} = \bigotimes_{i=1}^n \mathcal{A}_{\gamma^{A_i}} \otimes \mathcal{A}_{\gamma^{B_i}}$. The star network inequality in (34) is violated if and only if

$$(1 - \gamma^{A_i})(1 - \gamma^{B_i}) > \frac{1}{2} \quad (138)$$

and for γ^{A_i} and $\gamma^{B_i} \in [0, \frac{1}{2}]$, the maximal n -local star score is

$$S_{n\text{-Star}}^*(\rho^{\text{Net}}) = \prod_{i=1}^n \left(\sqrt{2(1 - \gamma^{A_i})(1 - \gamma^{B_i})} \right)^{\frac{1}{n}}. \quad (139)$$

Proof: Let $\rho = |\Phi^+\rangle\langle\Phi^+|$ be a maximally entangled state preparation and $u_x = u_y = \sqrt{1 - \gamma}$, $u_z = (1 - \gamma)$, and $t_z = \gamma$ be the nonunital channel parameters for the amplitude

damping channel. Substituting these values into the singular value expressions listed in the proof of Proposition 5, we find

$$\tau_{i,x} = \tau_{i,y} = \sqrt{(1 - \gamma^{A_i})(1 - \gamma^{B_i})} \quad (140)$$

$$\tau_{i,z} = \gamma^{A_i} \gamma^{B_i} + (1 - \gamma^{A_i})(1 - \gamma^{B_i}). \quad (141)$$

In the domain $\gamma^A, \gamma^B \in [0, \frac{1}{2}]$, the maximal singular values are (τ_x, τ_y) . Therefore, (39) calculates the maximal CHSH score as

$$S_{\text{CHSH}}^*(\tilde{\rho}) = 2\sqrt{2(1 - \gamma^A)(1 - \gamma^B)} \quad (142)$$

for which the CHSH inequality is violated if and only if $(1 - \gamma^A)(1 - \gamma^B) > \frac{1}{2}$. We then verify that for $\gamma^A \in [\frac{1}{2}, 1]$ and $\gamma^B \in [0, 1]$, no violations occur because $\tau_x^2 + \tau_y^2 \leq 1$ and $\tau_x^2 + \tau_z^2 \leq 1$ (and similarly for the values $\gamma^A \in [0, 1]$ and $\gamma^B \in [\frac{1}{2}, 1]$). Thus, the CHSH inequality is violated only when $(1 - \gamma^A)(1 - \gamma^B) > \frac{1}{2}$. Since $\tau_x = \tau_y$ are the two maximal eigenvalues of R_ρ when $\gamma^A, \gamma^B \in [0, \frac{1}{2}]$, (37) can be applied to recover the n -local star score in (139). Finally, the Bell state $|\Phi^+\rangle\langle\Phi^+|$ and amplitude damping channel are both diagonal in the Pauli basis; therefore, the evaluated score is maximal. ■

Proposition 7: Consider the CHSH scenario with uniform qubit amplitude damping noise $\mathcal{A}_\gamma^A \otimes \mathcal{A}_\gamma^B$, where $\gamma^A = \gamma^B = \gamma$. For a noisy nonmaximally entangled state preparation $\tilde{\rho}_{\lambda,\gamma} = \mathcal{A}_\gamma^A \otimes \mathcal{A}_\gamma^B(|\psi_\lambda\rangle\langle\psi_\lambda|)$, where $|\psi_\lambda\rangle = \sqrt{\lambda}|00\rangle + \sqrt{1-\lambda}|11\rangle$, the maximal CHSH score is

$$S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda^*,\gamma}) = \begin{cases} 2\sqrt{2(1-\gamma)^2}, & \gamma \in [0, \gamma_c] \\ 2\sqrt{(1-\gamma)^2 - \frac{(2\gamma^2-2\gamma+1)^2}{4\gamma^2-1}}, & \gamma \in [\gamma_c, \frac{1}{3}] \\ 2, & \gamma \in [\frac{1}{3}, 1] \end{cases} \quad (143)$$

where the optimal entanglement parameter is

$$\lambda^* = \begin{cases} \frac{1}{2}, & \gamma \in [0, \gamma_c] \\ \frac{1}{2} \left(1 - \frac{(\gamma^2 + (1-\gamma)^2)(2\gamma(1-\gamma))}{(2\gamma(1-\gamma))^2 - (1-\gamma)^2} \right), & \gamma \in [\gamma_c, \frac{1}{3}] \\ 1, & \gamma \in [\frac{1}{3}, 1] \end{cases} \quad (144)$$

the crossover noise parameter is

$$\gamma_c = \frac{1}{6} \left(4 + \frac{(3\sqrt{114} - 32)^{\frac{1}{3}}}{2^{\frac{2}{3}}} - \frac{1}{(6\sqrt{114} - 64)^{\frac{1}{3}}} \right) \quad (145)$$

and the nonlocality is broken with respect to the CHSH inequality when $\gamma = \gamma_0 = \frac{1}{3}$.

Proof: Using Proposition 5, we first substitute the values $u_x = u_y = \sqrt{1-\gamma}$, $u_z = (1-\gamma)$, and $t_z = \gamma$ into (130) and (131) to obtain (143) and (144). Note that when $\gamma \in [\frac{1}{3}, 1]$, then $\lambda^* = 1$ and the optimal state preparation corresponds to the classical state $|00\rangle\langle 00|$. Setting $\lambda = 1$ in (137) yields $S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda=1}) = 2$, which implies that the CHSH nonlocality is broken for all nonmaximally entangled states when $\gamma = \frac{1}{3}$. For $\gamma \geq \frac{1}{3}$, the classical state preparation $|00\rangle\langle 00|$ is optimal and achieves the classical bound.

We now solve for the crossover noise parameter γ_c at which the optimal state preparation switches from being maximally entangled to nonmaximally entangled. We thus set the two cases in (63) to be equal and solve for γ to find (145). ■

Proposition 8: Consider the CHSH scenario with single-qubit amplitude damping noise $\mathcal{A}_\gamma^A \otimes \text{id}^B$, where $\gamma^A = \gamma$. For a noisy nonmaximally entangled state preparation, $\tilde{\rho}_{\lambda,\gamma} = \mathcal{A}_\gamma^A \otimes \text{id}^B(|\psi_\lambda\rangle\langle\psi_\lambda|)$, where $|\psi_\lambda\rangle = \sqrt{\lambda}|00\rangle + \sqrt{1-\lambda}|11\rangle$, the maximal CHSH score is

$$S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda^*}) = \begin{cases} 2\sqrt{2(1-\gamma)}, & \gamma \in [0, \frac{1}{2}] \\ 2, & \gamma \in [\frac{1}{2}, 1] \end{cases} \quad (146)$$

where $\lambda^* = \frac{1}{2}$ when $\gamma \in [0, \frac{1}{2}]$ and $\lambda^* = 1$ when $\gamma \in [\frac{1}{2}, 1]$.

Proof: When nonmaximally entangled states are considered, we use Proposition 5 to find that

$$\lambda^* = \frac{1}{2} \left(1 - \frac{(1-\gamma)\gamma}{\gamma^2 - (1-\gamma)} \right) \quad (147)$$

which yields

$$S_{\text{CHSH}}^*(\tilde{\rho}_{\lambda^*}) = 2\sqrt{(1-\gamma) + (1-\gamma)^2 - \frac{(1-\gamma)^2\gamma^2}{\gamma^2 - (1-\gamma)}}. \quad (148)$$

However, $\lambda^* = 1$ when $\gamma = \frac{1}{2}$ and within the domain $\gamma \in [0, \frac{1}{2}]$, it can be verified that

$$(1-\gamma) \geq (1-\gamma)^2 - \frac{(1-\gamma)^2\gamma^2}{\gamma^2 - (1-\gamma)}. \quad (149)$$

Thus, when amplitude damping is applied to a single qubit, maximally entangled state preparations are optimal on the range $\gamma \in [0, \frac{1}{2}]$ and the classical state preparation is optimal for $\gamma \in [\frac{1}{2}, 1]$. ■

2) MAXIMAL N -LOCAL VIOLATIONS FOR BIASED DETECTOR NOISE

Proposition 9: The biased detector error postprocessing map

$$\mathbf{R}_\gamma = (1-\gamma)\mathbb{I}_2 + \gamma \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad (150)$$

is nonunital and equivalent to a partial replacer channel

$$\mathcal{R}_{\gamma,x}(\rho^A) = (1-\gamma)\rho^A + \gamma\rho'_x \text{Tr}[\rho^A] \quad (151)$$

that is applied to the quantum state ρ^{A_j} , where the replacer state ρ'_x is a density operator contained by the projective subspace of $\Pi_{+|x}^{A_j}$.

Proof: Consider an M -qubit partial replacer channel having a pure replacer state $\rho'_x = |\psi'_x\rangle\langle\psi'_x|$. The Kraus operators are

$$K_i = \sqrt{\gamma} |\psi'_x\rangle\langle i| \quad \forall i \in [0, 2^M), \quad K_{2^M} = \sqrt{1-\gamma} \mathbb{I}_{2^M} \quad (152)$$

where $\{|i\rangle\}_{i=0}^{2^M-1}$ form an orthonormal basis. Then, for a measurement operator $\Pi_{\pm|x}$, we find

$$\text{Tr}[\Pi_{\pm|x}\mathcal{R}_\gamma(\rho)] = \sum_i \text{Tr}[\Pi_{\pm|x}K_i\rho K_i^\dagger] \quad (153)$$

$$= \sum_i \text{Tr}[K_i^\dagger \Pi_{\pm|x} K_i \rho] \quad (154)$$

$$= \text{Tr}[\mathcal{R}_\gamma^\dagger(\Pi_{\pm|x})\rho]. \quad (155)$$

If the replacer state $|\psi'_x\rangle$ lies within the projective subspace of $\Pi_{+|x}$ such that $\langle\psi'_x|\Pi_{+|x}|\psi'_x\rangle = 1$, thus

$$\mathcal{R}_{\gamma,x}^\dagger(\Pi_{+|x}) = \sum_i K_i^\dagger \Pi_{+|x} K_i \quad (156)$$

$$= (1-\gamma)\Pi_{+|x} + \gamma \sum_{i=0}^{2^M-1} |i\rangle\langle\psi'_x| \Pi_{+|x} |\psi'_x\rangle\langle i| \quad (157)$$

$$= (1-\gamma)\Pi_{+|x} + \gamma\mathbb{I}_{2^M}. \quad (158)$$

However, repeating the procedure for $\Pi_{-|x}$ yields $\mathcal{R}_{\gamma,x}^\dagger(\Pi_{-|x}) = (1-\gamma)\Pi_{-|x}$ because $\langle\psi'_x|\Pi_{-|x}|\psi'_x\rangle = 0$. Thus, we recover the POVM in (82) from considered replacer channel $\mathcal{R}_{\gamma,x}$. ■

ACKNOWLEDGMENT

The authors acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

REFERENCES

- [1] H. J. Kimble, "The quantum internet," *Nature*, vol. 453, no. 7198, pp. 1023–1030, Jun. 2008, doi: [10.1038/nature07127](#).
- [2] C. Simon, "Towards a global quantum network," *Nature Photon.*, vol. 11, no. 11, pp. 678–680, Oct. 2017, doi: [10.1038/s41566-017-0032-0](#).
- [3] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: A vision for the road ahead," *Science*, vol. 362, 2018, Art. no. 6412, doi: [10.1126/science.aam9288](#).
- [4] W. Kozłowski and S. Wehner, "Towards large-scale quantum networks," in *Proc. 6th Annu. ACM Int. Conf. Nanoscale Comput. Commun.*, 2019, pp. 1–7, doi: [10.1145/3345312.3345497](#).
- [5] P. Kómár et al., "A quantum network of clocks," *Nature Phys.*, vol. 10, no. 8, pp. 582–587, Jun. 2014, doi: [10.1038/nphys3000](#).
- [6] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, "Optical interferometry with quantum networks," *Phys. Rev. Lett.*, vol. 123, Art. no. 070504, Aug. 2019, doi: [10.1103/PhysRevLett.123.070504](#).
- [7] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, "Quantum-assisted telescope arrays," *Phys. Rev. A*, vol. 100, Art. no. 022316, Aug. 2019, doi: [10.1103/PhysRevA.100.022316](#).
- [8] Q.-C. Sun et al., "Quantum teleportation with independent sources and prior entanglement distribution over a network," *Nature Photon.*, vol. 10, no. 10, pp. 671–675, Sep. 2016, doi: [10.1038/nphoton.2016.179](#).
- [9] D. Cazzolino, B. D. Lio, D. Bacco, and L. K. Oxenløwe, "High-dimensional quantum communication: Benefits, progress, and future challenges," *Adv. Quantum Technol.*, vol. 2, no. 12, Oct. 2019, Art. no. 1900038, doi: [10.1002/qute.201900038](#).
- [10] S. Herbert, "Increasing the classical data throughput in quantum networks by combining quantum linear network coding with superdense coding," *Phys. Rev. A*, vol. 101, Jun. 2020, Art. no. 062332, doi: [10.1103/PhysRevA.101.062332](#).
- [11] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Rev. Mod. Phys.*, vol. 81, pp. 1301–1350, Sep. 2009, doi: [10.1103/RevModPhys.81.1301](#).
- [12] A. Broadbent, J. Fitzsimons, and E. Kashefi, "Universal blind quantum computation," in *Proc. IEEE 50th Annu. Symp. Found. Comput. Sci.*, Oct. 2009, pp. 517–526, doi: [10.1109/focs.2009.36](#).
- [13] C. M. Lee and M. J. Hoban, "Towards device-independent information processing on general quantum networks," *Phys. Rev. Lett.*, vol. 120, Jan. 2018, Art. no. 020504, doi: [10.1103/PhysRevLett.120.020504](#).
- [14] S. Pirandola et al., "Advances in quantum cryptography," *Adv. Opt. Photon.*, vol. 12, no. 4, pp. 1012–1236, Dec. 2020, doi: [10.1364/AOP.361502](#).
- [15] M.-X. Luo, "Fully device-independent model on quantum networks," *Phys. Rev. Res.*, vol. 4, Mar. 2022, Art. no. 013203, doi: [10.1103/PhysRevResearch.4.013203](#).
- [16] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, "Nonlocality and communication complexity," *Rev. Mod. Phys.*, vol. 82, pp. 665–698, Mar. 2010, doi: [10.1103/RevModPhys.82.665](#).
- [17] D. Cuomo, M. Caleffi, and A. S. Cacciapuoti, "Towards a distributed quantum computing ecosystem," *IET Quantum Commun.*, vol. 1, no. 1, pp. 3–8, Jul. 2020, doi: [10.1049/iet-qtc.2020.0002](#).
- [18] I. L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box," *J. Modern Opt.*, vol. 44, no. 11–12, pp. 2455–2467, 1997, doi: [10.1080/09500349708231894](#).
- [19] R. Harper, S. T. Flammia, and J. J. Wallman, "Efficient learning of quantum noise," *Nature Phys.*, vol. 16, no. 12, pp. 1184–1188, 2020, doi: [10.1038/s41567-020-0992-8](#).
- [20] E. Onorati, T. Kohler, and T. Cubitt, "Fitting quantum noise models to tomography data," 2021, *arXiv:2103.17243*, doi: [10.48550/arXiv.2103.17243](#).
- [21] R. Orús, "Tensor networks for complex quantum systems," *Nature Rev. Phys.*, vol. 1, no. 9, pp. 538–550, Aug. 2019, doi: [10.1038/s42254-019-0086-7](#).
- [22] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Phys. Rev. A*, vol. 98, Sep. 2018, Art. no. 032309, doi: [10.1103/PhysRevA.98.032309](#).
- [23] N. Moll et al., "Quantum optimization using variational algorithms on near-term quantum devices," *Quantum Sci. Technol.*, vol. 3, no. 3, Jun. 2018, Art. no. 030503, doi: [10.1088/2058-9565/aab822](#).
- [24] M. Cerezo et al., "Variational quantum algorithms," *Nature Rev. Phys.*, vol. 3, no. 9, pp. 625–644, Aug. 2021, doi: [10.1038/s42254-021-00348-9](#).
- [25] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, "Theory of variational quantum simulation," *Quantum*, vol. 3, p. 191, Oct. 2019, doi: [10.22331/q-2019-10-07-191](#).
- [26] S. Endo, J. Sun, Y. Li, S. C. Benjamin, and X. Yuan, "Variational quantum simulation of general processes," *Phys. Rev. Lett.*, vol. 125, Jun. 2020, Art. no. 010501, doi: [10.1103/PhysRevLett.125.010501](#).
- [27] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018, doi: [10.22331/q-2018-08-06-79](#).
- [28] A. Suprano et al., "Dynamical learning of a photonics quantum-state engineering process," *Adv. Photon.*, vol. 3, no. 6, Dec. 2021, Art. no. 066002, doi: [10.1117/1.ap.3.6.066002](#).
- [29] D. Poderini, E. Polino, G. Rodari, A. Suprano, R. Chaves, and F. Sciarrino, "Ab initio experimental violation of bell inequalities," *Phys. Rev. Res.*, vol. 4, Feb. 2022, Art. no. 013159, doi: [10.1103/PhysRevResearch.4.013159](#).
- [30] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, Mar. 1993, doi: [10.1103/PhysRevLett.70.1895](#).
- [31] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, "Event-ready-detectors' Bell experiment via entanglement swapping," *Phys. Rev. Lett.*, vol. 71, pp. 4287–4290, Dec. 1993, doi: [10.1103/PhysRevLett.71.4287](#).
- [32] S. Bose, V. Vedral, and P. L. Knight, "Multiparticle generalization of entanglement swapping," *Phys. Rev. A*, vol. 57, pp. 822–829, Feb. 1998, doi: [10.1103/PhysRevA.57.822](#).
- [33] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, "Quantum repeaters: The role of imperfect local operations in quantum communication," *Phys. Rev. Lett.*, vol. 81, pp. 5932–5935, Dec. 1998, doi: [10.1103/PhysRevLett.81.5932](#).
- [34] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, "Quantum repeaters based on atomic ensembles and linear optics," *Rev. Mod. Phys.*, vol. 83, pp. 33–80, Mar. 2011, doi: [10.1103/RevModPhys.83.33](#).

- [35] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Phys. Phys. Fizika*, vol. 1, no. 3, pp. 195–200, 1964, doi: [10.1103/PhysicsPhysique-Fizika.1.195](https://doi.org/10.1103/PhysicsPhysique-Fizika.1.195).
- [36] A. Aspect, P. Grangier, and G. Roger, "Experimental tests of realistic local theories via Bell's theorem," *Phys. Rev. Lett.*, vol. 47, pp. 460–463, Aug. 1981, doi: [10.1103/PhysRevLett.47.460](https://doi.org/10.1103/PhysRevLett.47.460).
- [37] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," *Rev. Mod. Phys.*, vol. 86, pp. 419–478, Apr. 2014, doi: [10.1103/RevModPhys.86.419](https://doi.org/10.1103/RevModPhys.86.419).
- [38] A. Tavakoli, A. Pozas-Kerstjens, M.-X. Luo, and M.-O. Renou, "Bell nonlocality in networks," *Rep. Prog. Phys.*, vol. 85, no. 5, Mar. 2022, Art. no. 056001, doi: [10.1088/1361-6633/ac41bb](https://doi.org/10.1088/1361-6633/ac41bb).
- [39] C. Branciard, N. Gisin, and S. Pironio, "Characterizing the nonlocal correlations created via entanglement swapping," *Phys. Rev. Lett.*, vol. 104, Apr. 2010, Art. no. 170401, doi: [10.1103/PhysRevLett.104.170401](https://doi.org/10.1103/PhysRevLett.104.170401).
- [40] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, "Bilocal versus non-bilocal correlations in entanglement-swapping experiments," *Phys. Rev. A*, vol. 85, Mar. 2012, Art. no. 032119, doi: [10.1103/PhysRevA.85.032119](https://doi.org/10.1103/PhysRevA.85.032119).
- [41] A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acín, "Nonlocal correlations in the star-network configuration," *Phys. Rev. A*, vol. 90, Art. no. 062109, Dec. 2014, doi: [10.1103/PhysRevA.90.062109](https://doi.org/10.1103/PhysRevA.90.062109).
- [42] K. Mukherjee, B. Paul, and D. Sarkar, "Correlations in n-local scenario," *Quantum Inf. Process.*, vol. 14, no. 6, pp. 2025–2042, Apr. 2015, doi: [10.1007/s11128-015-0971-7](https://doi.org/10.1007/s11128-015-0971-7).
- [43] D. Rosset, C. Branciard, T. J. Barnea, G. Pütz, N. Brunner, and N. Gisin, "Nonlinear Bell inequalities tailored for quantum networks," *Phys. Rev. Lett.*, vol. 116, Jan. 2016, Art. no. 010403, doi: [10.1103/PhysRevLett.116.010403](https://doi.org/10.1103/PhysRevLett.116.010403).
- [44] A. Tavakoli, "Bell-type inequalities for arbitrary noncyclic networks," *Phys. Rev. A*, vol. 93, Mar. 2016, Art. no. 030101, doi: [10.1103/PhysRevA.93.030101](https://doi.org/10.1103/PhysRevA.93.030101).
- [45] L. Yang, X. Qi, and J. Hou, "Nonlocal correlations in the tree-tensor-network configuration," *Phys. Rev. A*, vol. 104, Oct. 2021, Art. no. 042405, doi: [10.1103/PhysRevA.104.042405](https://doi.org/10.1103/PhysRevA.104.042405).
- [46] G. Carvacho, F. Andreoli, L. Santodonato, M. Bentivegna, R. Chaves, and F. Sciarrino, "Experimental violation of local causality in a quantum network," *Nature Commun.*, vol. 8, no. 1, pp. 1–6, 2017, doi: [10.1038/ncomms14775](https://doi.org/10.1038/ncomms14775).
- [47] D. J. Saunders, A. J. Bennet, C. Branciard, and G. J. Pryde, "Experimental demonstration of nonbilocal quantum correlations," *Sci. Adv.*, vol. 3, no. 4, 2017, Art. no. e1602743, doi: [10.1126/sciadv.1602743](https://doi.org/10.1126/sciadv.1602743).
- [48] F. Andreoli, G. Carvacho, L. Santodonato, M. Bentivegna, R. Chaves, and F. Sciarrino, "Experimental bilocality violation without shared reference frames," *Phys. Rev. A*, vol. 95, Jun. 2017, Art. no. 062315, doi: [10.1103/PhysRevA.95.062315](https://doi.org/10.1103/PhysRevA.95.062315).
- [49] Q.-C. Sun et al., "Experimental demonstration of non-bilocality with truly independent sources and strict locality constraints," *Nature Photon.*, vol. 13, no. 10, pp. 687–691, 2019, doi: [10.1038/s41566-019-0502-7](https://doi.org/10.1038/s41566-019-0502-7).
- [50] D. Poderini et al., "Experimental violation of n-locality in a star quantum network," *Nature Commun.*, vol. 11, no. 1, pp. 1–8, 2020, doi: [10.1038/s41467-020-16189-6](https://doi.org/10.1038/s41467-020-16189-6).
- [51] R. Gallego, N. Brunner, C. Hadley, and A. Acín, "Device-independent tests of classical and quantum dimensions," *Phys. Rev. Lett.*, vol. 105, Nov. 2010, Art. no. 230501, doi: [10.1103/PhysRevLett.105.230501](https://doi.org/10.1103/PhysRevLett.105.230501).
- [52] M. O. Renou, J. m. k. Kaniewski, and N. Brunner, "Self-testing entangled measurements in quantum networks," *Phys. Rev. Lett.*, vol. 121, Dec. 2018, Art. no. 250507, doi: [10.1103/PhysRevLett.121.250507](https://doi.org/10.1103/PhysRevLett.121.250507).
- [53] J.-D. Bancal, N. Sangouard, and P. Sekatski, "Noise-resistant device-independent certification of Bell state measurements," *Phys. Rev. Lett.*, vol. 121, Dec. 2018, Art. no. 250506, doi: [10.1103/PhysRevLett.121.250506](https://doi.org/10.1103/PhysRevLett.121.250506).
- [54] I. Šupić and J. Bowles, "Self-testing of quantum systems: A review," *Quantum*, vol. 4, Sep. 2020, Art. no. 337, doi: [10.22331/q-2020-09-30-337](https://doi.org/10.22331/q-2020-09-30-337).
- [55] I. Šupić, J.-D. Bancal, Y. Cai, and N. Brunner, "Genuine network quantum nonlocality and self-testing," *Phys. Rev. A*, vol. 105, Feb. 2022, Art. no. 022206, doi: [10.1103/PhysRevA.105.022206](https://doi.org/10.1103/PhysRevA.105.022206).
- [56] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Phys. Rev. Lett.*, vol. 67, pp. 661–663, Aug. 1991, doi: [10.1103/PhysRevLett.67.661](https://doi.org/10.1103/PhysRevLett.67.661).
- [57] J. Barrett, L. Hardy, and A. Kent, "No signaling and quantum key distribution," *Phys. Rev. Lett.*, vol. 95, Jun. 2005, Art. no. 010503, doi: [10.1103/PhysRevLett.95.010503](https://doi.org/10.1103/PhysRevLett.95.010503).
- [58] A. Acín, N. Gisin, and L. Masanes, "From Bell's theorem to secure quantum key distribution," *Phys. Rev. Lett.*, vol. 97, Sep. 2006, Art. no. 120405, doi: [10.1103/PhysRevLett.97.120405](https://doi.org/10.1103/PhysRevLett.97.120405).
- [59] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, "Device-independent security of quantum cryptography against collective attacks," *Phys. Rev. Lett.*, vol. 98, Jun. 2007, Art. no. 230501, doi: [10.1103/PhysRevLett.98.230501](https://doi.org/10.1103/PhysRevLett.98.230501).
- [60] Y. Liu et al., "High-speed device-independent quantum random number generation without a detection loophole," *Phys. Rev. Lett.*, vol. 120, Jan. 2018, Art. no. 010503, doi: [10.1103/PhysRevLett.120.010503](https://doi.org/10.1103/PhysRevLett.120.010503).
- [61] U. Vazirani and T. Vidick, "Fully device independent quantum key distribution," *Commun. ACM*, vol. 62, no. 4, pp. 133–133, 2019, doi: [10.1145/3310974](https://doi.org/10.1145/3310974).
- [62] A. Cabello, A. Feito, and A. Lamas-Linares, "Bell's inequalities with realistic noise for polarization-entangled photons," *Phys. Rev. A*, vol. 72, Nov. 2005, Art. no. 052112, doi: [10.1103/PhysRevA.72.052112](https://doi.org/10.1103/PhysRevA.72.052112).
- [63] R. Pal and S. Ghosh, "Non-locality breaking qubit channels: The case for CHSH inequality," *J. Phys. A: Math. Theor.*, vol. 48, no. 15, Mar. 2015, Art. no. 155302, doi: [10.1088/1751-8113/48/15/155302](https://doi.org/10.1088/1751-8113/48/15/155302).
- [64] Y. Zhang, R. A. Bravo, V. O. Lorenz, and E. Chitambar, "Channel activation of CHSH nonlocality," *New J. Phys.*, vol. 22, no. 4, Apr. 2020, Art. no. 043003, doi: [10.1088/1367-2630/ab7bef](https://doi.org/10.1088/1367-2630/ab7bef).
- [65] D.-L. Deng, "Machine learning detection of Bell nonlocality in quantum many-body systems," *Phys. Rev. Lett.*, vol. 120, Jun. 2018, Art. no. 240402, doi: [10.1103/PhysRevLett.120.240402](https://doi.org/10.1103/PhysRevLett.120.240402).
- [66] A. Canabarro, S. Brito, and R. Chaves, "Machine learning nonlocal correlations," *Phys. Rev. Lett.*, vol. 122, May 2019, Art. no. 200401, doi: [10.1103/PhysRevLett.122.200401](https://doi.org/10.1103/PhysRevLett.122.200401).
- [67] K. Bharti, T. Haug, V. Vedral, and L.-C. Kwek, "How to teach AI to play Bell non-local games: Reinforcement learning," 2019, *arXiv:1912.10783*, doi: [10.48550/arXiv.1912.10783](https://doi.org/10.48550/arXiv.1912.10783).
- [68] V. Bergholm et al., "PennyLane: Automatic differentiation of hybrid quantum-classical computations," 2018, *arXiv:1811.04968*, doi: [10.48550/arXiv.1811.04968](https://doi.org/10.48550/arXiv.1811.04968).
- [69] B. Doolittle and T. Bromley, "qNetVO: The quantum network variational optimizer, v0.1.3," Mar. 2022. [Online]. Available: <https://github.com/ChitambarLab/qNetVO>
- [70] B. Doolittle, "Supplemental code: Variational quantum optimization of nonlocality in noisy quantum networks, v0.1.0," May 2022. [Online]. Available: <https://github.com/ChitambarLab/vqo-nonlocality-noisy-quantum-networks>
- [71] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2009, doi: [10.1017/cbo9780511976667](https://doi.org/10.1017/cbo9780511976667).
- [72] W. F. Stinespring, "Positive functions on c^* -algebras," *Proc. Amer. Math. Soc.*, vol. 6, no. 2, pp. 211–216, 1955, doi: [10.2307/2032342](https://doi.org/10.2307/2032342).
- [73] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory. Ser., Lecture Notes in Physics*, vol. 190. Berlin, Heidelberg: Springer1983, doi: [10.1007/3-540-12732-1](https://doi.org/10.1007/3-540-12732-1).
- [74] S. Ruder, "An overview of gradient descent optimization algorithms," 2016, *arXiv:1609.04747*, doi: [10.48550/arXiv.1609.04747](https://doi.org/10.48550/arXiv.1609.04747).
- [75] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization methods for large-scale machine learning," *SIAM Rev.*, vol. 60, no. 2, pp. 223–311, 2018, doi: [10.1137/16M1080173](https://doi.org/10.1137/16M1080173).
- [76] A. Griewank et al., "On automatic differentiation," *Math. Program.: Recent Develop. Appl.*, vol. 6, no. 6, pp. 83–107, 1989, doi: [10.1590/0101-7438.2014.034.03.0621](https://doi.org/10.1590/0101-7438.2014.034.03.0621).
- [77] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, "Automatic differentiation in machine learning: A survey," *J. Mach. Learn. Res.*, vol. 18, no. 153, pp. 1–43, 2018, doi: [10.5555/3122009.3242010](https://doi.org/10.5555/3122009.3242010). [Online]. Available: <http://jmlr.org/papers/v18/baydin17-468.html>
- [78] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Phys. Rev. A*, vol. 99, Mar. 2019, Art. no. 032331, doi: [10.1103/PhysRevA.99.032331](https://doi.org/10.1103/PhysRevA.99.032331).
- [79] T. Fritz, "Beyond bells theorem: Correlation scenarios," *New J. Phys.*, vol. 14, no. 10, Oct. 2012, Art. no. 103001, doi: [10.1088/1367-2630/14/10/103001](https://doi.org/10.1088/1367-2630/14/10/103001).
- [80] R. Chaves, "Polynomial Bell inequalities," *Phys. Rev. Lett.*, vol. 116, Jan. 2016, Art. no. 010402, doi: [10.1103/PhysRevLett.116.010402](https://doi.org/10.1103/PhysRevLett.116.010402).
- [81] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, pp. 880–884, Oct. 1969, doi: [10.1103/PhysRevLett.23.880](https://doi.org/10.1103/PhysRevLett.23.880).

- [82] B. S. Cirelson, "Quantum generalizations of Bells inequality," *Lett. Math. Phys.*, vol. 4, no. 2, pp. 93–100, Mar. 1980, doi: [10.1007/bf00417500](https://doi.org/10.1007/bf00417500).
- [83] B. Doolittle and E. Chitambar, "Regarding the maximal qubit violations of n -locality in star and chain networks," 2022, *arXiv:2212.06915*, doi: [10.48550/arXiv.2212.06915](https://doi.org/10.48550/arXiv.2212.06915).
- [84] R. Horodecki, P. Horodecki, and M. Horodecki, "Violating Bell inequality by mixed spin-1/2 states: Necessary and sufficient condition," *Phys. Lett. A*, vol. 200, no. 5, pp. 340–344, 1995, doi: [10.1016/0375-9601\(95\)00214-N](https://doi.org/10.1016/0375-9601(95)00214-N).
- [85] N. Gisin, Q. Mei, A. Tavakoli, M. O. Renou, and N. Brunner, "All entangled pure quantum states violate the bilocality inequality," *Phys. Rev. A*, vol. 96, Aug. 2017, Art. no. 020304, doi: [10.1103/PhysRevA.96.020304](https://doi.org/10.1103/PhysRevA.96.020304).
- [86] F. Andreoli, G. Carvacho, L. Santodonato, R. Chaves, and F. Sciarrino, "Maximal qubit violation of n -locality inequalities in a star-shaped quantum network," *New J. Phys.*, vol. 19, no. 11, 2017, Art. no. 113020, doi: [10.1088/1367-2630/aa8b9b](https://doi.org/10.1088/1367-2630/aa8b9b).
- [87] A. Kundu, M. K. Molla, I. Chattopadhyay, and D. Sarkar, "Maximal qubit violation of n -local inequalities in a quantum network," *Phys. Rev. A*, vol. 102, Nov. 2020, Art. no. 052222, doi: [10.1103/PhysRevA.102.052222](https://doi.org/10.1103/PhysRevA.102.052222).
- [88] S. Dasgupta and T. S. Humble, "Stability of noisy quantum computing devices," 2021, *arXiv:2105.09472*, doi: [10.48550/arXiv.2105.09472](https://doi.org/10.48550/arXiv.2105.09472).
- [89] R. Gerritsma, G. Kirchmair, F. Zähringer, E. Solano, R. Blatt, and C. F. Roos, "Quantum simulation of the Dirac equation," *Nature*, vol. 463, no. 7277, pp. 68–71, Jan. 2010, doi: [10.1038/nature08688](https://doi.org/10.1038/nature08688).
- [90] B. P. Lanyon et al., "Universal digital quantum simulation with trapped ions," *Science*, vol. 334, no. 6052, pp. 57–61, Oct. 2011, doi: [10.1126/science.1208001](https://doi.org/10.1126/science.1208001).
- [91] R. Barends et al., "Digital quantum simulation of fermionic models with a superconducting circuit," *Nature Commun.*, vol. 6, no. 1, Jul. 2015, Art. no. 7654, doi: [10.1038/ncomms8654](https://doi.org/10.1038/ncomms8654).
- [92] E. A. Martinez et al., "Real-time dynamics of lattice gauge theories with a few-qubit quantum computer," *Nature*, vol. 534, no. 7608, pp. 516–519, Jun. 2016, doi: [10.1038/nature18318](https://doi.org/10.1038/nature18318).
- [93] B. Rost et al., "Simulation of thermal relaxation in spin chemistry systems on a quantum computer using inherent qubit decoherence," 2020, *arXiv:2001.00794*, doi: [10.48550/arXiv.2001.00794](https://doi.org/10.48550/arXiv.2001.00794).
- [94] B. Rost, L. Del Re, N. Earnest, A. F. Kemper, B. Jones, and J. K. Freericks, "Demonstrating robust simulation of driven-dissipative problems on near-term quantum computers," 2021, *arXiv:2108.01183*, doi: [10.48550/arXiv.2108.01183](https://doi.org/10.48550/arXiv.2108.01183).
- [95] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, "Toward the first quantum simulation with quantum speedup," *Proc. Nat. Acad. Sci.*, vol. 115, no. 38, pp. 9456–9461, Sep. 2018, doi: [10.1073/pnas.1801723115](https://doi.org/10.1073/pnas.1801723115).
- [96] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, "Simulating large quantum circuits on a small quantum computer," *Phys. Rev. Lett.*, vol. 125, Oct. 2020, Art. no. 150504, doi: [10.1103/PhysRevLett.125.150504](https://doi.org/10.1103/PhysRevLett.125.150504).
- [97] F. Barratt, J. Dborin, M. Bal, V. Stojevic, F. Pollmann, and A. G. Green, "Parallel quantum simulation of large systems on small NISQ computers," *NPJ Quantum Inf.*, vol. 7, no. 1, May 2021, Art. no. 79, doi: [10.1038/s41534-021-00420-3](https://doi.org/10.1038/s41534-021-00420-3).
- [98] M. Beth Ruskai, S. Szarek, and E. Werner, "An analysis of completely-positive trace-preserving maps on M_2 ," *Linear Algebra its Appl.*, vol. 347, no. 1, pp. 159–187, 2002, doi: [10.1016/S0024-3795\(01\)00547-X](https://doi.org/10.1016/S0024-3795(01)00547-X).