# Probing Many-Body Bell Correlation Depth with Superconducting Qubits

Ke Wang,[1,*] Weikang Li,[2,3,4,*] Shibo Xu,[1,*] Mengyao Hu,[3,4] Jiachen Chen,[1] Yaozu Wu,[1] Chuanyu Zhang,[1] Feitong Jin,[1] Xuhao Zhu,[1] Yu Gao,[1] Ziqi Tan,[1] Zhengyi Cui,[1] Aosai Zhang,[1] Ning Wang,[1] Yiren Zou,[1] Tingting Li,[1] Fanhao Shen,[1] Jiarun Zhong,[1] Zehang Bao,[1] Zitian Zhu,[1] Zixuan Song,[1] Jinfeng Deng,[1] Hang Dong,[1] Xu Zhang,[1] Pengfei Zhang,[1] Wenjie Jiang,[2] Zhide Lu,[2,5] Zheng-Zhi Sun,[2] Hekang Li,[1] Qiujiang Guo,[1,6] Zhen Wang,[1,6] Patrick Emonts,[3,4] Jordi Tura,[3,4] Chao Song,[1,6,†] H. Wang,[1,6,‡] and Dong-Ling Deng[2,5,6,§]

[1]*School of Physics, ZJU-Hangzhou Global Scientific and Technological Innovation Center, and Zhejiang Key Laboratory of Micro-nano Quantum Chips and Quantum Control, Zhejiang University, Hangzhou 310027, China*
[2]*Center for Quantum Information, IIIS, Tsinghua University, Beijing 100084, China*
[3]*Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands*
[4]*⟨aQa^L⟩ Applied Quantum Algorithms Leiden, The Netherlands*
[5]*Shanghai Qi Zhi Institute, Shanghai 200232, China*
[6]*Hefei National Laboratory, Hefei 230088, China*

(Received 20 August 2024; revised 13 January 2025; accepted 10 March 2025; published 22 April 2025)

Quantum nonlocality describes a stronger form of quantum correlation than that of entanglement. It refutes Einstein's belief of local realism and is among the most distinctive and enigmatic features of quantum mechanics. It is a crucial resource for achieving quantum advantages in a variety of practical applications, ranging from cryptography and certified random number generation via self-testing to machine learning. Nevertheless, the detection of nonlocality, especially in quantum many-body systems, is notoriously challenging. Here, we report an experimental certification of genuine multipartite Bell-operator correlations, which signal nonlocality in quantum many-body systems, up to 24 qubits with a fully programmable superconducting quantum processor. In particular, we employ energy as a Bell-operator correlation witness and variationally decrease the energy of a many-body system across a hierarchy of thresholds, below which an increasing Bell-operator correlation depth can be certified from experimental data. We variationally prepare the low-energy state of a two-dimensional honeycomb model with 73 qubits and certify its Bell-operator correlations by measuring an energy that surpasses the corresponding classical bound with up to 48 standard deviations. In addition, we variationally prepare a sequence of low-energy states and certify their genuine multipartite Bell-operator correlations up to 24 qubits via energies measured efficiently by parity oscillation and multiple quantum coherence techniques. Our results establish a viable approach for preparing and certifying multipartite Bell-operator correlations, which provide not only a finer benchmark beyond entanglement for quantum devices, but also a valuable guide toward exploiting multipartite Bell correlations in a wide spectrum of practical applications.

DOI: 10.1103/PhysRevX.15.021024 Subject Areas: Quantum Physics, Quantum Information

## I. INTRODUCTION

Quantum and classical physics differ fundamentally [1]. Particles of a composite quantum system can exhibit correlations that are stronger than any classical theory permits. Several types of such correlations have been identified and experimentally verified: entanglement [2], Einstein-Podolsky-Rosen (EPR) steering [3], and Bell nonlocality [4]. They serve as inequivalent resources for quantum technologies and form a hierarchical structure [Fig. 1(a)], wherein the former is necessary but not enough to certify the latter [5]. In particular, any quantum state that manifests Bell nonlocality, which can be experimentally validated by demonstrating violations of Bell inequalities [6,7], is necessarily entangled, but the opposite is not true: There exist quantum states that are entangled but admit a local-hidden-variable description, and hence do not violate any Bell inequality and cannot show Bell nonlocality properties [8,9]. Apart from this fundamental difference, Bell nonlocality (rather than entanglement) has been shown to be an

*These authors contributed equally to this work.
†Contact author: chaosong@zju.edu.cn
‡Contact author: hhwang@zju.edu.cn
§Contact author: dldeng@tsinghua.edu.cn

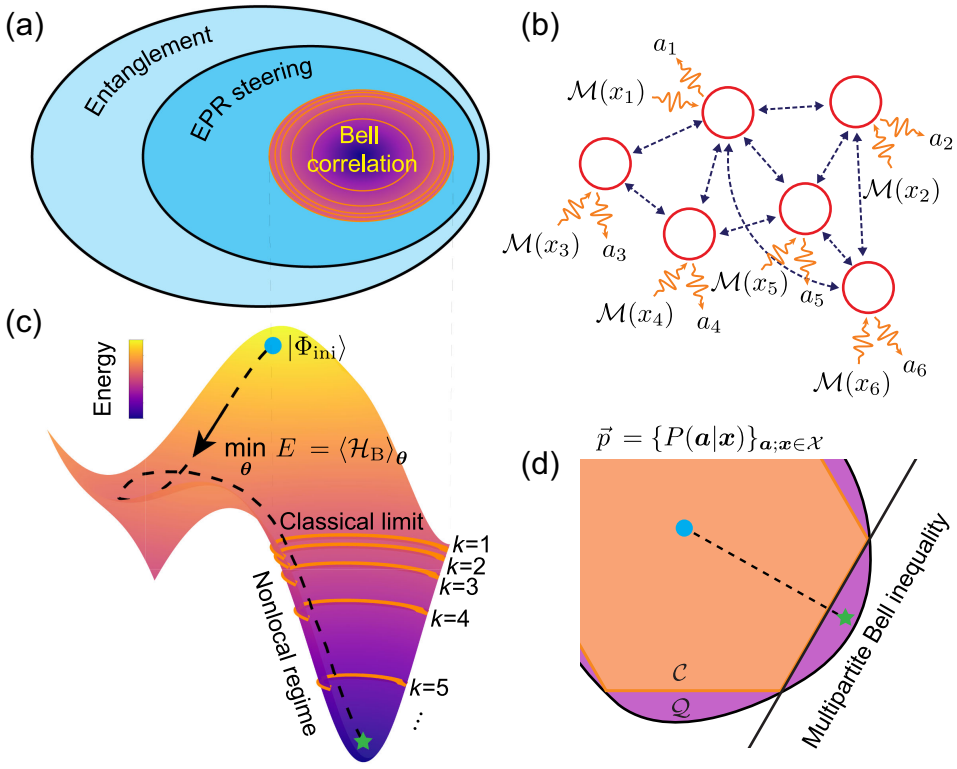


FIG. 1. Many-body Bell correlation and the variational detection approach. (a) The hierarchy of quantum correlations, starting from entanglement and Einstein-Podolsky-Rosen (EPR) steering, and finally to Bell correlation. Bell correlation is the strongest quantum correlation, in the sense that whenever it is detected then the other two types of correlation are guaranteed. In quantum many-body systems, Bell correlation can be further characterized by its depth (indicated by yellow contours), which quantifies the minimal number of particles sharing genuine nonlocal correlations. (b) A generic multipartite quantum system for detecting many-body Bell correlation, where site-resolved measurements $\mathcal{M}(\mathbf{x})$ ($\mathbf{x} \coloneqq x_1, \ldots, x_N$ with $x_i$ being the direction of measurement performed on the $i$th party) are performed on all parties and the corresponding outputs $\mathbf{a} \coloneqq a_1, \ldots, a_N$ are collected. The multipartite Bell inequality can be transformed into a many-body Hamiltonian by assigning each party the corresponding quantum observables. The detection of Bell correlation then becomes equivalent to finding low-energy states of the system with the energy exceeding the classical bound. (c) A schematic illustration of the essential idea of probing the many-body Bell-operator correlation with a variational quantum circuit (VQC). Starting from the initial state $|\Phi_{\rm ini}\rangle$, the VQC iteratively updates variational parameters $\boldsymbol{\theta}$ to minimize the energy of the many-body system $E = \langle \mathcal{H}_{\rm B} \rangle_{\boldsymbol{\theta}}$, leading to violations of the corresponding Bell inequality and detection of the correlation depth. (d) A sketch of the sets of correlations. For a given measurement configuration set $\mathcal{X} \coloneqq \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K\}$, the probability distribution of all possible outcomes, denoted collectively by $\vec{p} = \{P(\mathbf{a}|\mathbf{x})\}_{\mathbf{a};\mathbf{x}\in\mathcal{X}}$, is confined in the polytope (light orange), which denotes the set of all possible classical correlations assuming local realism ($\mathcal{C}$). A facet of the polytope denotes a tight multipartite Bell inequality (black solid line), outside which certifies Bell correlations ($\mathcal{Q}$).

indispensable resource for a wide range of practical applications, such as device-independent quantum key distribution [10–12], benchmarking of quantum devices [13], certified random number generation via self-testing [14], and unconditional quantum computational advantages [15–18].

Given its fundamental and practical importance, Bell nonlocality has been tested and confirmed repeatedly in various systems over the past 50 years, ranging from optical photons [19–25] and spins in nitrogen-vacancy centers [26] to neutral atoms [27] and superconducting qubits [28]. The Nobel Prize in Physics 2022 was awarded jointly to Aspect, Clauser, and Zeilinger for their pioneering works in this direction [29]. So far, most existing experiments on Bell nonlocality have been focused on two- or few-body systems [19–28]. For the observation of genuine multipartite nonlocality, experiments have also been accomplished with a handful of qubits [30,31]. While entanglement has been extensively studied [2,32] and hierarchies of genuine multipartite entanglement [33] have been reported with different experimental platforms, Bell nonlocality for quantum many-body systems remains much less explored. Noteworthy previous experiments have demonstrated the existence of so-called Bell-type correlations through measuring certain collective observables in Bose-Einstein condensates [34] and spin-squeezed states with up to 500 000 rubidium atoms [35]. These experiments involve measurements on specific many-body operators derived from Bell correlators. Here, the validity of quantum mechanics is assumed and we leave, strictly speaking, the device-independent paradigm [10] where loophole-free Bell experiments are necessary.

With recent progress in quantum computing, noisy intermediate-scale quantum devices with dozens to hundreds of individually addressable qubits are now available in laboratories, allowing more flexible Hamiltonian engineering and model design [36]. In this work, we consider a scenario in between the device-independent and the Bell-correlation setting in terms of assumptions. We do not require the assumption of collective observables nor the permutational invariance of the Bell inequalities used in Refs. [34,35], but meanwhile the scenario is not loophole-free. To make our setting clear and distinguish it from other results [19–28,30,31], we provide a detailed analysis and comparison in Sec. II, and throughout this paper we use the term *Bell-operator correlations* to describe the studied correlations.

Here, we report for the first time an experimental certification of genuine multipartite Bell-operator correlations up to 24 particles on a fully programmable superconducting quantum processor. Our superconducting processor consists of up to 73 qubits arranged on a two-dimensional (2D) lattice, with site-resolved controllability over all qubits and their nearest neighbor couplings [37]. By optimizing the control and readout procedures, we achieve parallel single-qubit (two-qubit) gates and simultaneous site-resolved qubit-state measurements with high precision, and the medians of the characteristic gate or measurement fidelities are benchmarked to all be above 99%. This enables us to demonstrate Bell-operator correlations in a fully controllable quantum many-body system with unprecedented size and accuracy. The implementation represents a step forward in tackling the challenge of scaling up the demonstration of genuine Bell-operator correlations in addressable many-body systems. This not only yields a stronger quantumness benchmark beyond entanglement for these devices, but also lays down the foundation for their practical applications based on multipartite Bell-operator correlations.

The rest of the paper is structured as follows. In Sec. II, we introduce the general framework and experimental setups for this work. To illustrate the validity and effectiveness of our approach, we conduct three experiments demonstrating Bell-operator correlations by variationally preparing low-energy states of three spin Hamiltonians. In Sec. III, we present a pedagogical example of showing Bell-operator correlations in a one-dimensional (1D) XXZ-type Hamiltonian on 21 qubits with a violation up to 7 standard deviations. In Sec. IV, we certify Bell-operator correlations for an XZ-type Hamiltonian on a 2D honeycomb lattice with 73 qubits by measuring an energy that surpasses the classical bound with up to 48 standard deviations. Furthermore, in Sec. V, to detect genuine multipartite Bell-operator correlations between more than two parties, we prepare a sequence of low-energy states and certify their Bell-operator correlation depth. This experiment reveals the minimal number of parties sharing genuine nonlocal correlations in a composite system by observing violations of the Svetlichny inequality [38,39] on up to 24 qubits. Finally, we conclude in Sec. VI.

## II. GENERAL RECIPE AND EXPERIMENTAL SETUP

Although a number of Bell inequalities for quantum many-body systems have been discovered [4,38–44], the experimental demonstration of Bell-operator correlations by violating an inequality is exceedingly challenging, especially for the case of certifying genuine multipartite correlations.

Our general recipe is illustrated in Fig. 1. For a given Bell inequality written in the generic form $\mathcal{I} \geq \beta_C$, where $\mathcal{I}$ is a linear function of correlators and $\beta_C$ denotes the classical bound, we assign appropriate quantum measurements to each party involved in the inequality [Fig. 1(b)]. After choosing a set of measurements, we can map $\mathcal{I}$ to a Bell operator by replacing the terms in the inequality with the corresponding measurement observables. Note that by measuring the Bell operator as a proxy for the Bell inequality, we are already leaving the device-independent formalism. This implies, for instance, an underlying Hilbert space structure of a certain dimension in which such operators live. Only by assuming the validity of quantum mechanics, it makes sense to talk about a Bell operator. A crucial idea for our framework is the reinterpretation of this Bell operator as a many-body Hamiltonian $\mathcal{H}_B$ [45]. In this formalism, the certification of Bell-operator correlations is reduced to the task of finding quantum states with energies lower than the corresponding classical bounds $\beta_C$. Because of the individual addressability of our system, the programmable quantum processor allows us to flexibly fabricate Hamiltonians associated to a variety of desired Bell inequalities. To contextualize our work with existing literature, we note that (a) our work studies a weaker form of correlations than traditional loophole-free Bell experiments, as ours assumes the validity of quantum mechanics [23–26,28]. Hence, it does not provide the same level of experimental evidence against local realism as a traditional loophole-free Bell test. (b) Our work studies a stronger form of correlations than experiments based on atomic ensembles from the interpretation perspective [34,35]: Many-body nonlocality demands the separation of individual parties. Although the particles in our experiments are in the same light cone, i.e., not spacelike separated, the individually addressable qubits take us one step closer to the original Bell scenario.

When measuring many-body Bell-operator correlations, there are at least two pronounced difficulties: (i) preparing highly entangled quantum many-body states that violate these Bell inequalities with suitably chosen measurement settings, despite inevitable experimental noise in current quantum devices, and (ii) some of these inequalities are constructed from correlation functions involving all parties and encompass an exponential number of correlators [46]. To address the first difficulty, we employ energy as a Bell-operator correlation witness [45] and introduce an approach

based on variational quantum circuits (VQCs) [47]. We start with an initialized variational quantum state $|\Phi(\boldsymbol{\theta})\rangle$, where $\boldsymbol{\theta}$ denotes collectively the variational parameters, and then iteratively update $\boldsymbol{\theta}$ to minimize the energy expectation $\langle\Phi(\boldsymbol{\theta})|\mathcal{H}_{\mathrm{B}}|\Phi(\boldsymbol{\theta})\rangle$. The gradients required in updating $\boldsymbol{\theta}$ are obtained by exploiting the parameter shift rule [48] and measuring the difference between energies at two shifted parameters directly on the device (see Appendix B for details). In general, the initial states would not show a violation of the corresponding Bell inequality. As we update $\boldsymbol{\theta}$ step by step, the energy decreases and eventually becomes lower than the classical bound, indicating the presence of Bell correlations [Figs. 1(c) and 1(d)]. Updating $\boldsymbol{\theta}$ further may decrease the energy further, leading to the certification of a larger correlation depth [Fig. 1(c)]. To overcome the second difficulty, on the one hand, we judiciously choose measurement settings to reduce the number of terms involved in $\mathcal{H}_{\mathrm{B}}$ and, on the other hand, upgrade the experimental setup substantially. In particular, the measurement fidelity was improved from 94% [37] to 99%.

Our experiments are performed on a superconducting processor with up to 73 frequency-tunable transmon qubits arranged in a square lattice [Fig. 3(a)]. Each qubit has an on-chip control line for individual XY control (rotating around axes in the $xy$ plane of the Bloch sphere) and Z control (rotating around $z$ axis), which enable arbitrary single-qubit gates (see Appendix D for detailed experimental gate design). Each qubit is also dispersively coupled to its own readout resonator to realize individual projective measurements. The nearest neighboring qubits are linked through a tunable coupler, which is also a transmon qubit with an individual on-chip control line, enabling the *in situ* tunability of the coupling strength. Through optimizing the control and readout procedures, we push the median simultaneous single-qubit (two-qubit) gate fidelity above 99.95% (99.4%) and the median measurement fidelity above 99.1% (averaged for each qubit of readout error $|0\rangle$ and $|1\rangle$) on 73 qubits. During the implementation of VQCs, the couplings are turned off for parallel single-qubit gates and readout, and activated for applying two-qubit gates, which is essential for realizing high-fidelity quantum operations and measurements required in the detection of genuine multipartite Bell-operator correlations. The single-qubit measurement takes about 2 μs in our experiments. Since this is longer than it takes for light to traverse the distance between two qubits, the individual qubits are in the same light cone. Hence, our experiment can resolve only Bell-operator correlations.

## III. BELL-OPERATOR CORRELATIONS IN A 1D CHAIN

We start with an illustrative example in a single spatial dimension. To measure Bell-operator correlations, we start with a one-dimensional XXZ model with open boundary conditions and staggered weights,

$$\mathcal{H}_{\mathrm{B}} = \sum_{i=1}^{N-1} J_i(\epsilon)(\sigma_x^{(i)}\sigma_x^{(i+1)} + \sigma_y^{(i)}\sigma_y^{(i+1)} + \Delta\sigma_z^{(i)}\sigma_z^{(i+1)}), \quad (1)$$

where $J_i(\epsilon) \coloneqq (4/\sqrt{3})[1-(-1)^i\epsilon]$ denotes the coupling strength and $\sigma_a^{(i)}$ is the corresponding Pauli matrix $a \in \{x, y, z\}$ acting on site $i$.

The derivation for this Hamiltonian starts from Gisin's elegant inequality [49], which involves two parties $A$ and $B$ with four and three dichotomic measurements, respectively. These measurements produce outcomes $\pm 1$. Here, we use a modified version of Gisin's elegant inequality,

$$\begin{aligned}\mathcal{I}_\Delta &= \langle A_0B_0\rangle + \langle A_1B_0\rangle - \langle A_2B_0\rangle - \langle A_3B_0\rangle \\ &\quad + \langle A_0B_1\rangle - \langle A_1B_1\rangle + \langle A_2B_1\rangle - \langle A_3B_1\rangle \\ &\quad + \Delta(\langle A_0B_2\rangle - \langle A_1B_2\rangle - \langle A_2B_2\rangle + \langle A_3B_2\rangle) \\ &\geq -(2|\Delta| + |\Delta+2| + |\Delta-2|), \end{aligned} \quad (2)$$

where $\Delta$ is a tunable parameter [45].

Given an odd number $N \geq 3$, we define the 1D-chain model involving $N$ parties given in Eq. (1). For each pair of neighboring parties, we assign the inequality described in Eq. (2), where the odd (even) sites play the role of party $A$ ($B$). The overall inequality results from a weighted sum of two-party inequalities $\mathcal{I}_{\mathrm{G}} \coloneqq \sum_{i=1}^{N-1}[1-(-1)^i\epsilon]\mathcal{I}_\Delta^{(i,i+1)}$, where $\mathcal{I}_\Delta^{(i,i+1)}$ denotes the inequality $\mathcal{I}_\Delta$ acting on sites indexed by $i$ and $i+1$. The classical bound $\beta_{\mathrm{C}}$ is $\beta_{\mathrm{C}} = -(N-1)(2|\Delta| + |\Delta+2| + |\Delta-2|)$. The mapping from this inequality to the Hamiltonian in Eq. (1) is achieved by assigning specific measurement settings (see Appendix A 2) [45,49].

For the experiment, we set $N = 21$, $\Delta = 2$, and $\epsilon = 0.95$. Given the mapping from an inequality to a many-body Hamiltonian, the quantum violation of inequality Eq. (2) corresponds to a system energy $\langle\mathcal{H}_{\mathrm{B}}\rangle$ that is lower than the classical bound $\beta_{\mathrm{C}}$. In particular, the maximal violation corresponds to the ground-state energy of Eq. (1). In the 1D case, we could obtain a good approximation of the ground-state energy with classical means. Yet, we use the system here as an introductory example. In our approach, we exploit the VQC ansatz to tackle this challenge and detect many-body Bell-operator correlations. As shown in Fig. 2(a), we design a parametrized quantum circuit ansatz which consists of single-qubit rotation gates and two-qubit controlled-Z (CZ) gates. During the optimization process, we use the parameter shift rule [48] to obtain the gradients (i.e., the derivatives of the model Hamiltonian's energy with respect to the variational parameters) by measuring the corresponding observables.

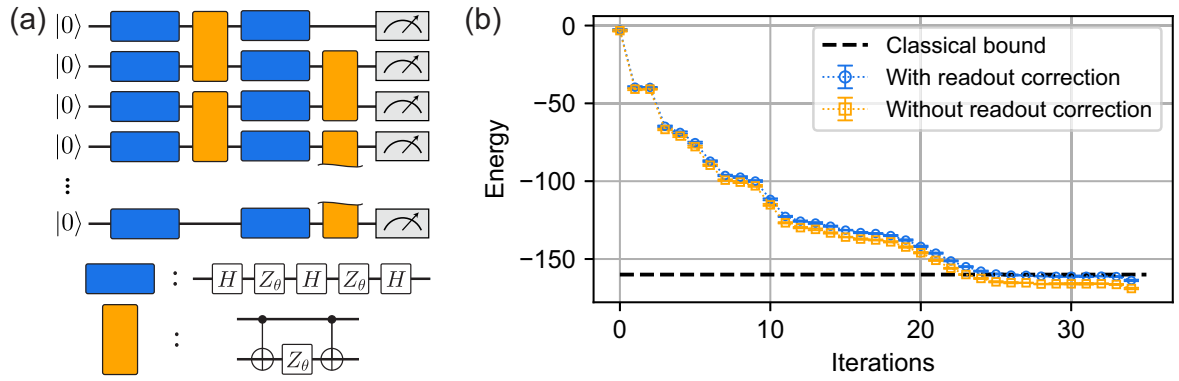


FIG. 2. Observation of Bell-operator correlations in a 1D chain. (a) The variational quantum circuit for detecting the Bell-operator correlations, where variational parameters are specified as $\theta$ in the virtual Z gate [50]. (b) The optimization trajectories of the XXZ-type Hamiltonian with hyperparameters $\Delta = 2$ and $\epsilon = 0.95$. The black dashed line denotes the classical bound, surpassing which certifies Bell-operator correlations.

We record the observed energy at each iteration during the training process, as plotted in Fig. 2(b). After 34 iterations of optimization, the energy of the prepared state reaches around $\langle \mathcal{H}_{\mathrm{B}} \rangle = -163.7$, achieving 7 standard deviations lower than the classical bound $-160$. After correcting readout errors, the energy becomes $-168.9$, which exceeds the classical bound by 17 standard deviations.

## IV. EXTENSION TO A HONEYCOMB LATTICE MODEL

Extending beyond the 1D structure, we consider a multipartite Bell inequality defined on a 2D honeycomb lattice. We label the two sublattices of the honeycomb lattice by $A$ and $B$, and the links between them by different colors $\tau \in \{\mathrm{r}, \mathrm{b}, \mathrm{g}\}$, as shown in Fig. 3(a). In comparison to the 1D experiment, we assign two choices of measurements $x, y \in \{0, 1\}$ with outcomes $a, b \in \{-1, +1\}$ for particles living on sublattices $A$ and $B$, respectively. The multipartite Bell inequality reads [51]

$$\mathcal{I} = \sum_{\tau\text{-link}} J_\tau(\epsilon) \sum_{x,y} (-1)^{x \cdot y} \langle A_x B_y \rangle_\tau \geq \beta_{\mathrm{C}}, \tag{3}$$

where $\beta_{\mathrm{C}} = -2 \sum_\tau J_\tau(\epsilon)$, $\langle A_x B_y \rangle = \sum_{a,b} abP(ab|xy)$, and $J_\tau(\epsilon) = 1 + \epsilon$ and $1 - \epsilon/2$ for $\tau = \mathrm{r}$ and $\tau = \mathrm{b}, \mathrm{g}$, respectively. By choosing proper measurement settings (see Appendix A 1 for details), the corresponding Bell operator reduces to an XZ-type Hamiltonian:

$$\mathcal{H}_{\mathrm{B}} = \sum_{\tau\text{-link}} J_\tau(\epsilon) (\sigma_x^A \sigma_x^B + \sigma_z^A \sigma_z^B)_\tau, \tag{4}$$

where $\sigma_x$ and $\sigma_z$ are Pauli operators acting on the two qubits at the ends of the link $\tau$ and the superscript denotes the sublattice they belong to. As in Sec. III, a measured system energy $\langle \mathcal{H}_{\mathrm{B}} \rangle$ lower than the classical bound $\beta_{\mathrm{C}}$ constitutes a quantum violation of the inequality Eq. (3). In two spatial dimensions, the many-body Hamiltonian is not exactly solvable. As a result, determining its ground-state energy is classically intractable in general. As in the 1D case, we design a VQC, now with three blocks, each containing single-qubit rotation and two-qubit controlled-NOT (CNOT) gates, as shown in Fig. 3(b).

In Fig. 3(c), we plot our experimental results for detecting multipartite Bell-operator correlations with the Bell inequality Eq. (3) and five different coupling strengths $\epsilon = 0.5, 0.6, \ldots, 0.9$. We first initialize the system to a product state $|0\rangle^{\otimes 73}$ and then let it evolve under the parametrized quantum circuit ansatz. At the beginning, the variational parameters are arbitrarily chosen to be 1.0 and the measured energy (expectation value $\langle \mathcal{H}_{\mathrm{B}} \rangle$) is larger than the classical bound $\beta_{\mathrm{C}}$; hence no quantum violation and no Bell-operator correlations are detected. We then update the variational parameters with the measured gradients to decrease the energy step by step. As shown in Fig. 3(c), the energy decreases gradually as the number of iterations increases. It goes below the classical bound after a few iterations, leading to quantum violations and the detection of Bell-operator correlations. Taking the curve for $\epsilon = 0.9$ as an example, we observe that the corresponding measured energy goes below the classical bound after 14 iterations and converges to its minimal value of $-156.29$, which is lower than the classical bound $(-131.3)$ by 48 standard deviations and unambiguously signals Bell-operator correlations. When correcting for readout errors, the energy becomes $-162.41$, which exceeds the classical bound by 57 standard deviations.

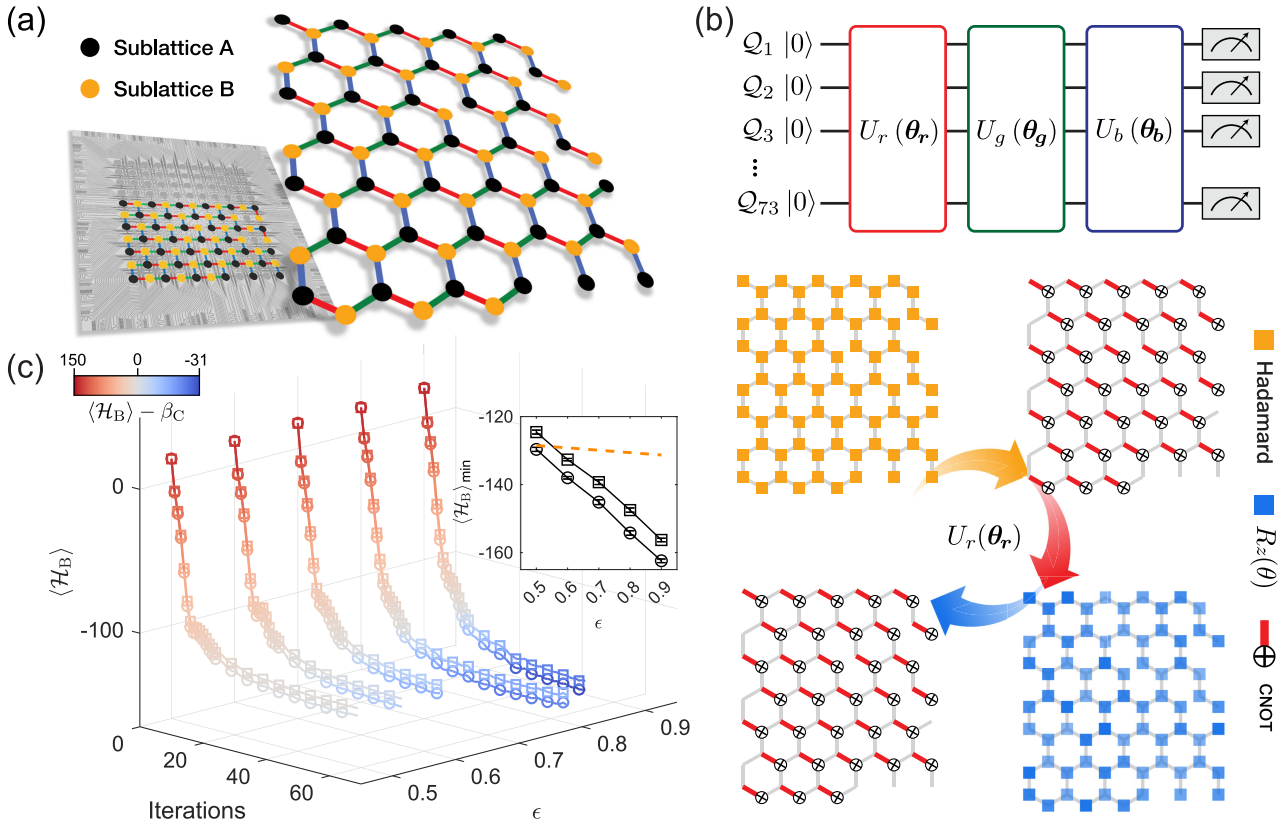


FIG. 3. Detecting Bell-operator correlation with 73 transmon qubits on a 2D honeycomb lattice. (a) Device and qubit topology. The superconducting chip contains a square lattice of $11 \times 11$ transmon qubits, from which we select 73 qubits arranged in a brick-wall (honeycomb) lattice, which consists of two triangular sublattices (denoted here by A and B, respectively). Based on the explicit form of the many-body Bell inequality Eq. (3), the system Hamiltonian is constructed by assigning different weights to the coupling terms between the nearest neighbor qubits of different orientations as colored by red, green, and blue, respectively. (b) The variational quantum circuit designed for detecting Bell-operator correlations. This variational ansatz is composed of three blocks, each of which begins with a layer of Hadamard gates, followed by a variational entangler which contains a layer of independent $R_z(\theta)$ gates and is sandwiched by two layers of controlled-NOT (CNOT) gates. For blocks with different colors, the CNOT gates are applied on qubit pairs connected by the links of the corresponding color as shown in (a) with the circuit for the red block exemplified in the lower panel. (c) Optimization trajectories for the XZ-type Hamiltonian $\mathcal{H}_B$ in Eq. (4) with different coupling configurations characterized by $\epsilon$. As the variational parameters are updated iteratively, all energy values for different $\epsilon$ decrease below the classical bound $\beta_C$, indicating a violation of the corresponding Bell inequality Eq. (3). The upper right-hand inset shows the measured minimal energies and the classical bound (yellow dashed line) versus $\epsilon$ during the variational process. Here we show both the data with (circle) and without (square) readout corrections. Error bars represent the standard deviation with ten repetitions of experiments.

## V. BELL-OPERATOR CORRELATION DEPTH

In the above discussion, Bell-operator correlations are detected once the energy is lower than the classical bound. Yet, in this scenario, only the bipartite Bell-operator correlation is certified from the experimental results. As an example, a single singlet in an otherwise classical state will register as a nonlocal state. This is not quite satisfactory in the many-body scenario. In this section, we turn to a more challenging problem of detecting genuine multipartite Bell-operator correlation. We consider a system consisting of $N$ qubits with each assigned two possible dichotomic measurements (labeled by $x_i$, $i \in \{1, 2, \ldots, N\}$). We exploit the following Svetlichny inequality [38,39]:

$$\mathcal{I}_N^S = 2^{-N/2}\left[\sum_{\mathbf{x}|s \text{ is even}} (-1)^{s/2}\langle\mathbf{A}\rangle_{\mathbf{x}} + \sum_{\mathbf{x}|s \text{ is odd}} (-1)^{(s-1)/2}\langle\mathbf{A}\rangle_{\mathbf{x}}\right] \geq -2^{(N-\lceil N/k\rceil)/2},$$

where $\langle\mathbf{A}\rangle_{\mathbf{x}} = \langle A_{1,x_1}\ldots A_{N,x_N}\rangle$ denotes the $N$-partite correlator, and $s$ is defined as $\sum_i x_i$. This inequality serves as a witness to genuine multipartite nonlocality. Any violation of this inequality would guarantee that the Bell nonlocality depth is at least $k+1$; i.e., there are at least $k+1$ parties sharing genuine nonlocal correlations. Yet, there are two challenges in the experimental detection of genuine multipartite Bell-operator correlation with this inequality.

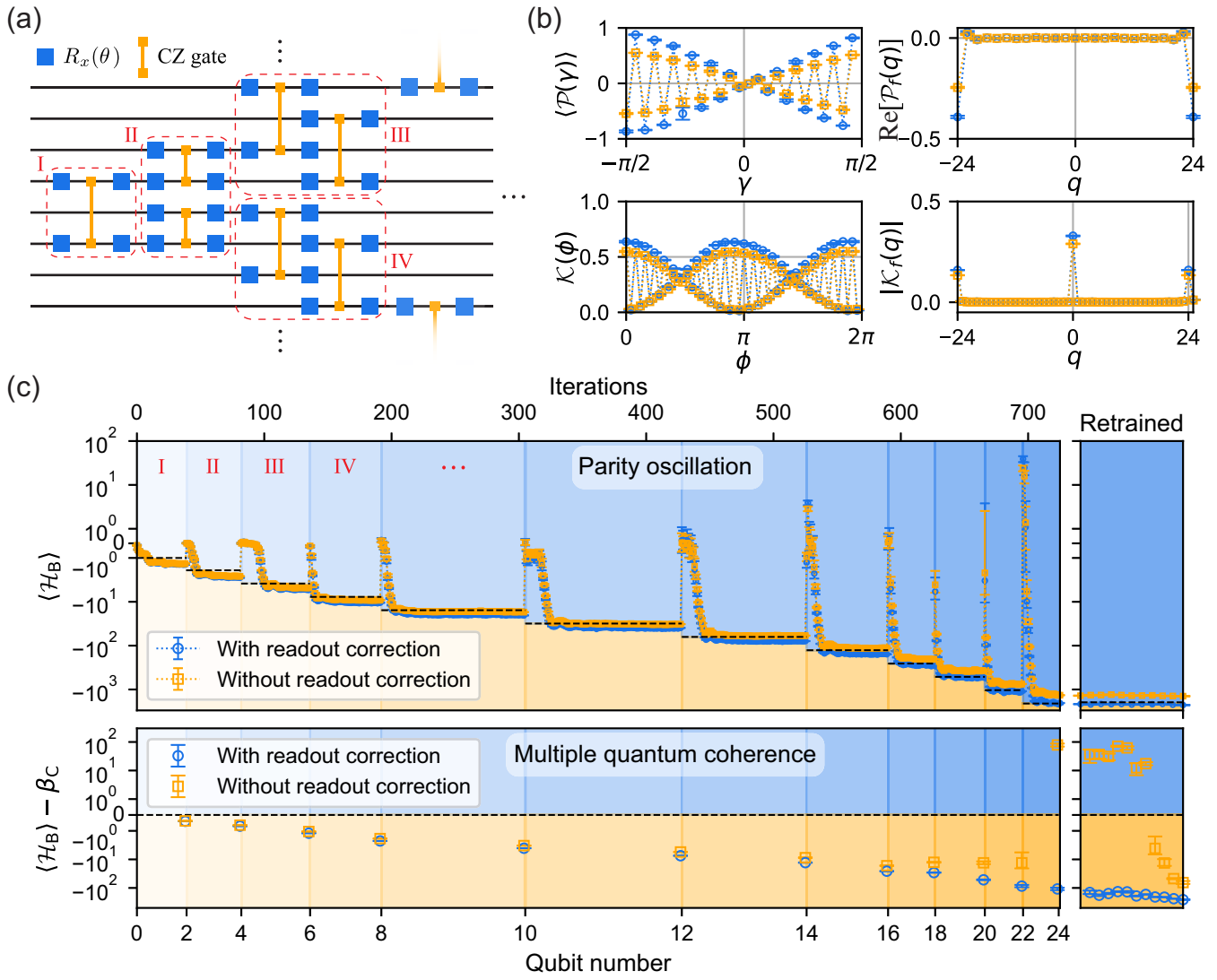


FIG. 4. Detection of Bell-operator correlation depth. (a) The variational quantum circuit with a hierarchical architecture designed for probing Bell-operator correlation depth. The circuit is divided into many subcircuits labeled I, II, III, etc., which are trained sequentially to avoid the problem of barren plateaus. Starting with two qubits, we train subcircuit I to minimize the energy of the two-qubit system, following which we expand the system by involving two more qubits and train subcircuit II to minimize the energy of the four-qubit system. The procedure goes on until all 24 qubits are involved, during which the detected Bell-operator correlation depth increases step by step. (b) Measured parity expectation $\langle\mathcal{P}(\gamma)\rangle$ and multiple quantum coherence (MQC) $\mathcal{K}(\phi)$ data for the 24-qubit Greenberger-Horne-Zeilinger (GHZ) state with (blue circles) and without (yellow squares) readout correction, alongside their corresponding Fourier spectrum [$\mathcal{P}_f(q)$ and $\mathcal{K}_f(q)$] from which the energy of the system $\langle\mathcal{H}_B\rangle$ with $\mathcal{H}_B$ shown in Eq. (5) can be obtained. Error bars denote the standard deviation from five repetitions of the experiment. (c) Optimization trajectory. During the training procedure, we assess the loss function—the energy of the system—based on parity measurement. The whole training procedure is separated into 13 phases, with the first 12 phases corresponding to the hierarchical training procedure and the last phase retraining the whole circuit again. The energy values with (blue circles) and without (orange squares) readout correction are presented for each phase, with error bars denoting the standard deviation from five repetitions of the experiment. Dashed lines indicate the $k$-nonlocal bounds (lower than which guarantees Bell-operator correlation depth $k+1$) for the corresponding phase. The lower panel shows the energy differences relative to the $k$-nonlocal bounds, with the energies measured by the MQC method at the end of each training phase, which are less affected by readout errors (see Appendix E for details).

First, the number of correlators involved in the Svetlichny inequality grows exponentially with the system size. Second, each correlator is an $N$-partite correlator, which renders its measurement exceedingly difficult as $N$ increases. To overcome the first challenge, we choose a particular set of measurement settings so that the corresponding Bell operator of $\mathcal{I}_N^S$ reduces to

$$\mathcal{H}_B(N) = 2^{(N-1)/2}[(|0\rangle\langle 1|)^{\otimes N} + (|1\rangle\langle 0|)^{\otimes N}]. \quad (5)$$

$\mathcal{H}_B(N)$ contains only two terms, yet it still requires $N$-partite coincidence measurements. To overcome this challenge, we substantially upgraded the experimental setup and optimized the control and readout procedures, achieving the state-of-the-art average measurement fidelity of 99.1% (see Appendix C). In addition, we exploit both parity oscillation [52] and multiple quantum coherence (MQC) [53] techniques to measure multipartite correlations in an efficient fashion.

With the upgraded experimental setup and the strategies discussed above, we still face one more subtle challenge that prevails in VQC algorithms—the barren plateau problem (i.e., the gradients vanish exponentially with the system size [54]). Indeed, as the system size and circuit depth increase, we find that the measured gradients decrease and would be completely washed out by experimental noise. Consequently, the training of the variational circuits eventually fails and we would not be able to detect a correlation depth larger than eight in our experiment. To overcome this challenge, we build up the variational circuit step by step and design a layerwise training strategy [55], as illustrated in Fig. 4. The variational quantum circuit starts with two qubits and the algorithm is applied to prepare a low-energy state of $\mathcal{H}_B(2)$. At each step, we add two more qubits and one more layer. We fix the parameters in the previous layers and variationally update the parameters of the new layer with measured gradients corresponding to the Hamiltonian $\mathcal{H}_B(N')$. With this approach and additional retraining of the whole variational circuit, we are able to overcome the barren plateau problem, and prepare and detect genuine multipartite nonlocality up to 24 qubits.

In Fig. 4(b), we benchmark the performance of the parity oscillation and multiple quantum coherence methods exploited in our experiment. We prepare the 24-qubit Greenberger-Horne-Zeilinger (GHZ) state and measure the parity expectation $\langle \mathcal{P}(\gamma) \rangle$ and multiple quantum coherence $\mathcal{K}(\phi)$, where $\gamma$ and $\phi$ correspond to the single-qubit rotation angles in the corresponding measurement circuits (see Appendix E). We find that the experimental data agree well with the theoretical prediction. In Fig. 4(c), we plot the optimization trajectory for the whole process. From this figure, genuine $k$-correlation depths are observed with $k$ spanning from 2 to 24 in different learning phases for data both with and without readout corrections.

## VI. CONCLUSION AND OUTLOOK

In summary, we have introduced a variational-quantum-circuit approach for probing Bell-operator correlations in many-body systems, and experimentally observed genuine multipartite Bell-operator correlations up to 24 qubits with a programmable superconducting quantum processor. In contrast to previously reported detection of Bell-operator correlations in ensembles of atoms without individual addressing, our experiments are performed on a fully programmable superconducting quantum processor with state-of-the-art gate and measurement fidelities. In addition, the variational approach exploited in our experiment is generically applicable to a broad family of Bell inequalities, independent of its specific structure and the qubit connection geometry of the quantum device.

Our work is addressing a gap in the certification of multipartite quantum devices. On the one hand, we have Bell nonlocality, which yields device-independent device certification, but is too demanding for a quantum computer in practice. On the other hand, only Bell correlations, arising from Bell inequalities tailored to the physics of specific experimental platforms, have been demonstrated [34,35]. The individual addressability provided by our superconducting platform is the necessary ingredient to certify the regime in between, enabling the use of Bell inequalities much better tailored to the device, e.g., not relying on permutational invariance as a natural symmetry.

Our results not only establish a practical approach to the preparation and certification of genuine multipartite Bell-operator correlations with noisy intermediate-scale quantum (NISQ) devices, but also provide a stronger benchmark beyond entanglement for testing how quantum these devices are. The controllability and scalability demonstrated in our experiment open up several new avenues for future studies. An interesting direction is to explore VQCs with certain symmetries [56–58] or use different optimization strategies, such as dissipative cooling [59] and quantum approximate optimization algorithms [60–62], to mitigate the barren plateau problem and enhance the effectiveness of our protocol. It would also be interesting to extend our approach to alternative definitions of nonlocality depth based on local observables and shared randomness [31]. From the perspective of practical applications, an experimental demonstration of quantum learning advantages [17,18] based on Bell correlations would mark a milestone toward future utilization of quantum technologies in artificial intelligence.

## ACKNOWLEDGMENTS

We thank Antonio Acín, L.-M. Duan, Vedran Dunjko, Marc-Olivier Renou, Jens Eisert, Benjamin Schiffer, and Dong Yuan for discussions. We also thank anonymous reviewers for their helpful comments and suggestions. The device was fabricated at the Micro-Nano Fabrication Center of Zhejiang University. We acknowledge support from the Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0300200 and No. 2021ZD0302203), the National Natural Science Foundation of China (Grant No. 12174342, No. 92365301, No. 12274367, No. 12322414, No. 12274368, No. 12075128, and No. T2225008), the National Key R&D Program of China (Grant No. 2023YFB4502600), and the Zhejiang Provincial Natural Science Foundation of China (Grant No. LDQ23A040001 and No. LR24A040002). W. L., W. J., Z. L., Z.-Z. S., and D.-L. D. are supported in addition

by Tsinghua University Dushi Program, and the Shanghai Qi Zhi Institute Innovation Program No. SQZ202318. H. W. is supported by the New Cornerstone Science Foundation through the XPLORER PRIZE. C. S. is supported by the Xiaomi Young Scholars Program. P. E. and J. T. acknowledge the support received by the Dutch National Growth Fund (NGF), as part of the Quantum Delta NL programme. P. E. acknowledges the support received through the NWO-Quantum Technology programme (Grant No. NGF.1623.23.006). J. T. acknowledges the support received from the European Union's Horizon Europe research and innovation programme through the ERC StG FINE-TEA-SQUAD (Grant No. 101040729). This publication is part of the "Quantum Inspire–the Dutch Quantum Computer in the Cloud" project (with Project No. NWA.1292.19.194) of the NWA research program "Research on Routes by Consortia (ORC)," which is funded by the Netherlands Organization for Scientific Research (NWO).

The views and opinions expressed here are solely those of the authors and do not necessarily reflect those of the funding institutions.

## APPENDIX A: BELL NONLOCALITY

Bell inequalities provide a criterion to experimentally test whether the outputs from quantum mechanics can be explained by a local-hidden-variable theory. The violation of a Bell inequality signals nonlocality, which has been demonstrated in loophole-free Bell experiments across various platforms [23–26,28,63]. Furthermore, certifying nonlocality in an experimental setup serves as compelling evidence that its behavior is intrinsically quantum. In this appendix, we introduce the Bell inequalities used in this work, as well as the concepts of multipartite Bell scenarios and Bell nonlocality depth.

### 1. Clauser-Horne-Shimony-Holt inequality (CHSH)

The predictions of quantum mechanics are incompatible with the assumption of a local-hidden-variable theory. According to Bell's theorem, this distinction can be measured via Bell's inequalities. One example is the Clauser-Horne-Shimony-Holt (CHSH) inequality [7]. For simplicity, consider the scenario involving two spatially separated parties, denoted as $A$ and $B$. Each party can perform one of two measurements, represented by $x \in \{0,1\}$ for party $A$ and $y \in \{0,1\}$ for party $B$. Each measurement has two outcomes, which are labeled by $a, b \in \{-1,+1\}$ for $A$ and $B$, respectively. This process is characterized by the conditional probability distribution $P(ab|xy)$ of obtaining outcomes $(a,b)$ when performing measurements $(x,y)$.

Consider the expression

$$\mathcal{I} = \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle, \quad \text{(A1)}$$

where $\langle A_x B_y \rangle = \sum_{a,b} abP(ab|xy)$ is the expectation value of outcomes $(a,b)$ when choosing measurements $(x,y)$. If $P(ab|xy)$ agrees with a local-hidden-variable model (LHVM), meaning there exists a probability distribution $p(\lambda)$ such that $P(ab|xy) = \sum_\lambda p(\lambda) P(a|x,\lambda) P(b|y,\lambda)$, then $|\mathcal{I}| \leq 2$, which is the well-known CHSH inequality. In this work, we choose the $\mathcal{I} \geq -2$ formulation, since we are focusing on a minimization task. However, if $P(ab|xy)$ is not of local correlations, the CHSH inequality can be violated. Specifically, let these two parties share the bipartite state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

Now party $A$ performs the measurements $A_0 = \sigma_z$, $A_1 = \sigma_x$, and party $B$ performs the measurements $B_0 = (1/\sqrt{2}) \times (\sigma_z + \sigma_x)$ and $B_1 = (1/\sqrt{2})(\sigma_z - \sigma_x)$, where $\sigma_x, \sigma_z$ are Pauli operators. Then, we have $\langle A_x B_y \rangle = -(1/\sqrt{2})(-1)^{x \cdot y}$, for $x, y \in \{0,1\}$. Further, we can obtain $\mathcal{I} = -2\sqrt{2} < -2$, which violates the CHSH inequality. Thus, CHSH inequality provides a way to distinguish between the set of quantum correlations and local correlations under the assumption of local hidden variables.

### 2. Gisin's elegant inequality and its generalization

Gisin's elegant inequality involves two distant parties $A$ and $B$, with four measurements for party $A$ and three measurements for party $B$, respectively. All measurements produce two outcomes, $\pm 1$ [49]. Let $\langle A_x B_y \rangle = \sum_{a,b} abP(ab|xy)$ again be the expectation value of the product of the outcomes of $A$'s $x$th measurement and $B$'s $y$th measurement. Gisin's elegant inequality, denoted as $\mathcal{I} \geq \beta_C$, reads as follows,

$$\begin{aligned} \mathcal{I} = {} & \langle A_0 B_0 \rangle + \langle A_1 B_0 \rangle - \langle A_2 B_0 \rangle - \langle A_3 B_0 \rangle \\ & + \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle + \langle A_2 B_1 \rangle - \langle A_3 B_1 \rangle \\ & + \langle A_0 B_2 \rangle - \langle A_1 B_2 \rangle - \langle A_2 B_2 \rangle + \langle A_3 B_2 \rangle \\ \geq {} & \beta_C = -6, \end{aligned} \quad \text{(A2)}$$

where $\beta_C = -6$ is the classical bound. Note that the inequality $\mathcal{I} \geq -6$ is not a facet of the polytope of local correlations [49]. The maximal quantum violation is $-4\sqrt{3} \approx -6.93$, which is attained when party $A$ and $B$ share the state $|\psi\rangle = (1/\sqrt{2})(|01\rangle - |10\rangle)$ [64]. In the case of the maximal violation, the four measurements of party $A$ are

$$A_0 = \frac{\sigma_x + \sigma_y + \sigma_z}{\sqrt{3}}, \qquad A_1 = \frac{\sigma_x - \sigma_y - \sigma_z}{\sqrt{3}},$$
$$A_2 = \frac{-\sigma_x + \sigma_y - \sigma_z}{\sqrt{3}}, \qquad A_3 = \frac{-\sigma_x - \sigma_y + \sigma_z}{\sqrt{3}},$$

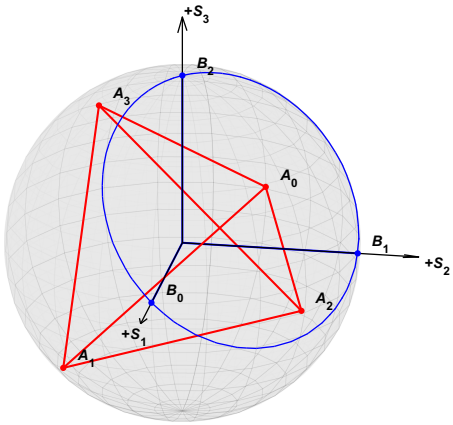


FIG. 5. Measurement settings represented on a Bloch sphere for Gisin's elegant inequality. For party $A$, the four measurements are the vertices of the tetrahedron: $A_0 = (1/\sqrt{3})(1,1,1)$, $A_1 = (1/\sqrt{3})(1,-1,-1)$, $A_2 = (1/\sqrt{3})(-1,1,-1)$, and $A_3 = (1/\sqrt{3})(-1,-1,1)$ (points in red color). For party $B$, the three measurements are mutually orthogonal: $B_0 = (1,0,0)$, $B_1 = (0,1,0)$, and $B_2 = (0,0,1)$ (points in blue color), which are parallel with the three axes.

and the three measurements of party $B$ are

$$B_0 = \sigma_x, \qquad B_1 = \sigma_y, \qquad B_2 = \sigma_z,$$

where $\sigma_x$, $\sigma_y$, $\sigma_z$ are Pauli operators. We can represent this measurement setting on a Bloch sphere as shown in Fig. 5. The basis $\{+S_1, +S_2, +S_3\}$ for this sphere is the Pauli basis $\{\sigma_x, \sigma_y, \sigma_z\}$. The four measurements of party $A$ are the vertices of a tetrahedron and the three measurements of party $B$ are mutually orthogonal. The Bell operator of $\mathcal{I}$ in Eq. (A2) is

$$\mathcal{B} = \frac{4}{\sqrt{3}}(\sigma_x^A\sigma_x^B + \sigma_y^A\sigma_y^B + \sigma_z^A\sigma_z^B), \tag{A3}$$

and one can check that $\langle\psi|\mathcal{B}|\psi\rangle = -4\sqrt{3}$ is the maximal quantum violation.

With the same measurement setting, we use a modified version of Gisin's elegant inequality,

$$\begin{aligned}\mathcal{I}_\Delta = {} & \langle A_0B_0\rangle + \langle A_1B_0\rangle - \langle A_2B_0\rangle - \langle A_3B_0\rangle \\ & + \langle A_0B_1\rangle - \langle A_1B_1\rangle + \langle A_2B_1\rangle - \langle A_3B_1\rangle \\ & + \Delta(\langle A_0B_2\rangle - \langle A_1B_2\rangle - \langle A_2B_2\rangle + \langle A_3B_2\rangle) \\ \geq {} & -(2|\Delta| + |\Delta+2| + |\Delta-2|),\end{aligned}$$

as introduced in the main text, where $\Delta$ is a real number [45]. Now its corresponding Bell operator becomes

$$\mathcal{B}_\Delta = \frac{4}{\sqrt{3}}(\sigma_x^A\sigma_x^B + \sigma_y^A\sigma_y^B + \Delta\sigma_z^A\sigma_z^B). \tag{A4}$$

The ground-state energy of this operator, i.e., the minimal eigenvalue, is $4/\sqrt{3}\Delta$ for $\Delta \leq -1$ and $-(4/\sqrt{3})(2+\Delta)$ for $\Delta > -1$.

### 3. Multipartite scenario

To explore the Bell nonlocality and entanglement properties in a larger and more complex system, it is essential to study multipartite Bell inequalities. Consider a Bell scenario where $N$ spatially separated parties $A_1, A_2, \ldots, A_N$ share some $N$-partite resource. On their share of this resource, each party $A_i$ can perform one of $m$ measurements and each measurement can yield $d$ possible outcomes. We refer to this scenario as a Bell scenario $(N, m, d)$. For party $A_i$, we label the measurement choices $x_i \in [m] = \{0, \ldots, m-1\}$ and its outcomes $a_i \in [d] = \{0, \ldots, d-1\}$, respectively. The correlations between results are described by a collection of conditional probability distributions,

$$\{P(a_1, \ldots, a_N|x_1, \ldots, x_N)\}_{a_1,\ldots,a_N;x_1,\ldots,x_N}, \tag{A5}$$

where $P(a_1, \ldots, a_N|x_1, \ldots, x_N) =: P(\mathbf{a}|\mathbf{x})$ is the probability of obtaining outcomes $\mathbf{a} := a_1, \ldots, a_N$ by performing measurements $\mathbf{x} := x_1, \ldots, x_N$. We can order these probabilities in Eq. (A5) into a vector $\vec{p}$ with $(md)^N$ components; that is,

$$\vec{p} = \{P(\mathbf{a}|\mathbf{x})\}_{\mathbf{a};\mathbf{x}} \in \mathbb{R}^{(md)^N}. \tag{A6}$$

In characterizing multipartite correlations, one typically studies whether they are compatible with a given physical model. Three notable examples are the no-signaling principle, local-hidden-variable models, and quantum theory.

Nonsignaling correlations are defined by the no-signaling principle which states that all parties are spatially separated and cannot communicate instantaneously. This ensures that the marginals observed by any subset of parties remain consistent and well defined; i.e., any subset of parties does not depend on the measurement choices of the other parties [65,66]. In terms of probability distributions in Eq. (A5), the no-signaling principle can be expressed as

$$\begin{aligned}&\sum_{a_i} P(a_1, \ldots, a_i, \ldots, a_N|x_1, \ldots, x_i, \ldots, x_N) \\ &= \sum_{a_i} P(a_1, \ldots, a_i, \ldots, a_N|x_1, \ldots, x_i', \ldots, x_N)\end{aligned} \tag{A7}$$

for all $x_i \neq x_i'$, $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N$ and $x_1, \ldots, x_{i-1}$, $x_{i+1}, \ldots, x_N$. Thus, $P(a_{i_1}, \ldots, a_{i_l}|x_{i_1}, \ldots, x_{i_l})$ is well defined on any subset $\{i_1, \ldots, i_l\} \subseteq \{1, \ldots, N\}$. Moreover, the probability distributions in Eq. (A5) need to satisfy the $m^N$ affine-linear equations enforcing the normalization of the probabilities,

$$\sum_{a_1,\ldots,a_N} P(a_1, \ldots, a_N|x_1, \ldots, x_N) = 1, \quad \forall\ x_1, \ldots, x_N, \tag{A8}$$

as well as the inequalities $P(\mathbf{a}|\mathbf{x}) \geq 0$. The feasibility region for $\vec{p}$ satisfying the no-signaling principle is a polytope, which we denote by $\mathcal{NS}$.

The second class of correlations is defined by local-hidden-variable models. To detect Bell nonlocality, we need to find Bell inequalities separating the polytope $\mathcal{L}$ of local correlations described by LHVMs from the convex set of quantum correlations $\mathcal{Q}$. The set of probability distributions in Eq. (A5) that admit the LHVM is formed by those $\vec{p}$ in which all $P(\mathbf{a}|\mathbf{x})$ can be written as

$$P(a_1, \ldots, a_N|x_1, \ldots, x_N) = \sum_{\lambda} p(\lambda) \prod_{i=1}^{N} P(a_i|x_i, \lambda), \quad \text{(A9)}$$

where $\lambda$ is some hidden variable distributed according to a probability distribution $p(\lambda)$. The compatibility region for probability distributions that admit an LHVM description forms a polytope which we denote by $\mathcal{L}$. The vertices of $\mathcal{L}$ factorize as a product of deterministic correlations as follows:

$$P(a_1, \ldots, a_N|x_1, \ldots, x_N) = \prod_{i=1}^{N} P(a_i|x_i), \quad \text{(A10)}$$

where each $P(a_i|x_i)$ is deterministic, i.e., $P(a_i|x_i) = \delta(a_i - \alpha_{i,x_i})$, $\delta$ is the Kronecker delta function and $\alpha_{i,x_i} \in [d]$ [67].

Finally, we consider the set of quantum correlations $\mathcal{Q}$, where each element $P(\mathbf{a}|\mathbf{x})$ can be represented by Born's rule:

$$P(a_1, \ldots, a_N|x_1, \ldots, x_N) = \mathrm{Tr}[\rho_N(\mathcal{M}_{1,x_1}^{a_1} \otimes \cdots \otimes \mathcal{M}_{N,x_N}^{a_N})], \quad \text{(A11)}$$

where $\rho_N$ is an $N$-partite quantum state and $\mathcal{M}_{i,x_i}^{a_i} \geq 0$ is a positive operator-valued measure corresponding to the $x_i$th measurement with outcome $a_i$ performed by party $A_i$, satisfying the normalization condition $\sum_{a_i} \mathcal{M}_{i,x_i}^{a_i} = \mathbb{I}$. Note that $\mathcal{Q}$ is a convex set but not a polytope [68]. Moreover, we know $\mathcal{L} \subsetneq \mathcal{Q} \subsetneq \mathcal{NS}$ [4].

For convenience, we can represent the correlations in Eq. (A5) using expectation values instead of conditional probabilities. We restrict the number of outcomes to $d = 2$, which is the case considered in this work. First we write the probabilities $P(a_1, \ldots, a_N|x_1, \ldots, x_N)$ using $N$-dimensional discrete Fourier transformation as follows,

$$\langle A_{1,x_1}^{(k_1)} \ldots A_{N,x_N}^{(k_N)} \rangle = \sum_{(a_1,\ldots,a_N) \in \{0,1\}^N} (-1)^{(k_1,\ldots,k_N)\cdot(a_1,\ldots,a_N)} P(a_1, \ldots, a_N|x_1, \ldots, x_N), \quad \text{(A12)}$$

where $k_1, \ldots, k_N \in [2]$. Note that by using the no-signaling principle, one can obtain that when $k_i = 0$,

$$\langle A_{1,x_1}^{(k_1)} \ldots A_{N,x_N}^{(k_N)} \rangle = \langle A_{1,x_1}^{(k_1)} \ldots A_{i-1,x_{i-1}}^{(k_{i-1})} A_{i+1,x_{i+1}}^{(k_{i+1})} \ldots A_{N,x_N}^{(k_N)} \rangle, \quad \text{(A13)}$$

and $\langle A_{1,x_1}^{(0)} \ldots A_{N,x_N}^{(0)} \rangle = 1$ for any $(x_1, \ldots, x_N)$.

Finally, for the Bell scenario $(N, m, d)$, the general form of a Bell inequality $\mathcal{I}_{N,m,d} - \beta_C \geq 0$ is

$$\mathcal{I}_{N,m,d} := \sum_{x_1,\ldots,x_N=0}^{m-1} \sum_{k_1,\ldots,k_N=0}^{1} \alpha_{x_1,\ldots,x_N}^{(k_1,\ldots,k_N)} \langle A_{1,x_1}^{(k_1)} \ldots A_{N,x_N}^{(k_N)} \rangle, \quad \text{(A14)}$$

where the correlators $\langle A_{1,x_1}^{(k_1)} \ldots A_{N,x_N}^{(k_N)} \rangle$ are defined in Eq. (A12), and $\beta_C$ is the classical bound. Such Bell inequalities constrain the local polytope $\mathcal{L}$. The violation of a Bell inequality $\mathcal{I}_{N,m,d} - \beta_C \geq 0$ with $\beta_C \geq \min_{P(\mathbf{a}|\mathbf{x}) \in \mathcal{L}} \mathcal{I}_{N,m,2}$ detects Bell nonlocality. The inequality $\mathcal{I}_{N,m,2} - \beta_Q \geq 0$, where $\beta_Q \geq \inf_{P(\mathbf{a}|\mathbf{x}) \in \mathcal{Q}} \mathcal{I}_{N,m,2}$ is called the Tsirelson bound [65].

### 4. Bell nonlocality depth

The violation of Bell inequalities $\mathcal{I}_{N,m,d} \geq \beta_C$ signals nonlocality, implying that the multipartite probability distribution $P(\mathbf{a}|\mathbf{x})$ cannot be written in the form of Eq. (A9). However, the violation of a Bell inequality does not directly provide any information about the structure of the correlations. For example, a single singlet in a product state would be enough to violate a Bell inequality [69]. More concretely, consider the tripartite probability distribution $P(a_1a_2a_3|x_1x_2x_3)$. If, for example, it can be written as a product of probability distributions such as $P(a_1a_2|x_1x_2)P(a_3|x_3)$, and $P(a_1a_2|x_1x_2)$ describes nonlocal correlations, the whole distribution is considered nonlocal. Thus, a multipartite probability distribution can manifest in various classes of nonlocality. The concept of genuine multipartite nonlocality helps us to classify them [39,41,70].

To explain nonlocality depth, we first formulate the concept of $k$-producible probability distributions. For $N$ parties sharing some multipartite correlations, we are interested in quantifying the number of parties that share genuine multipartite nonlocality. To be more precise, consider these $N$ parties are in a partition $L_k$ of the set of indices $S = \{1, \ldots, N\}$ into $L$ pairwise disjoint nonempty subset $S_i$, where $S_i \cap S_j = \varnothing$ if $i \neq j$ and the

cardinality of each subset $|S_i| \leq k$. Then we say an $N$-partite probability distribution is $k$ producible with respect to the partition $L_k$ if and only if it has the following decomposition:

$$P_{L_k}(\mathbf{a}|\mathbf{x}) = \sum_{\lambda} p(\lambda) \prod_{i=1}^{L} P(\mathbf{a}_{S_i}|\mathbf{x}_{S_i}, \lambda), \tag{A15}$$

where $\mathbf{a}_{S_i}, \mathbf{x}_{S_i}$ are the tuples of outcomes and measurement choices corresponding to the parties in $S_i$. The largest number of parties exhibiting nonlocal correlations is the largest cardinality of the subset $S_i$ which is at most $k$. Note that we restrict the distributions $P(\mathbf{a}_{S_i}|\mathbf{x}_{S_i}, \lambda)$ to satisfy the nonsignaling principle in Eq. (A7). Since there are many ways to partition the set of indices $S$ as $L_k$, we can generalize the definition of $k$-producible by using a convex combination of Eq. (A15). To list all the possible partitions $L_k$ of $S$ for fixed $k$, note that the number of subsets $L$ can be varied, we define $\mathcal{P}_k$ to be the set of all possible partitions $L_k$. Now, $P(\mathbf{a}|\mathbf{x})$ is $k$ producible if and only if

$$P(\mathbf{a}|\mathbf{x}) = \sum_{L_k \in \mathcal{P}_k} q_{L_k} P_{L_k}(\mathbf{a}|\mathbf{x}), \tag{A16}$$

where $P_{L_k}(\mathbf{a}|\mathbf{x})$ are the correlations as defined in Eq. (A15) corresponding to the partition $L_k$. The minimum $k$ of $P(\mathbf{a}|\mathbf{x})$ which can be written in the form of Eq. (A16) is the so-called nonlocality depth [39,41,70]. This is the main concept we focus on in this work.

The set of all the $k$-producible probability distributions is a polytope that we denote $\mathcal{L}_k$. To characterize this polytope $\mathcal{L}_k$, we need to consider all the possible partitions $L_k$ and their corresponding product of probability distributions $P(\mathbf{a}|\mathbf{x})$. Similarly to the local polytope $\mathcal{L}$, the vertices of the polytope $\mathcal{L}_k$ are as follows:

$$P(\mathbf{a}|\mathbf{x}) = \prod_{i=1}^{L} P(\mathbf{a}_{S_i}|\mathbf{x}_{S_i}), \tag{A17}$$

where $P(\mathbf{a}_{S_i}|\mathbf{x}_{S_i})$ is a vertex of nonsignaling polytope with respect to $S_i$. Furthermore, the probability distributions $P(\mathbf{a}_{S_i}|\mathbf{x}_{S_i})$ can also be expressed in terms of correlators as illustrated in Eq. (A12). The facets of the polytope $\mathcal{L}_k$ define tight Bell-like inequalities [38,65,71]. Moreover, the violation of such Bell-like inequalities indicates a minimum Bell nonlocality depth of $k+1$, implying that the probability distribution is not $k$ producible. Thus, we can use these Bell-like inequalities to detect and quantify the nonlocality depth. However, characterizing the polytope $\mathcal{L}_k$ is an intractable task because of the need to list all the extreme points of $\mathcal{L}_k$ and the number of extreme points grows exponentially as a function of the number of parties $N$ [72,73].

We can still focus on certain inequalities constraining $\mathcal{L}_k$, for example, Mermin and Svetlichny inequalities [38,42], to study nonlocality depth in the multipartite scenario. Although the Mermin and Svetlichny inequalities were derived in different contexts, they exhibit a form of equivalence when applied to detecting nonlocality depth [38,39,41,42,74]. In the $(N,2,2)$ scenario, each party can perform two measurements, and each measurement $x_i$ yields two outcomes $a_i$; i.e., $x_i \in \{0,1\}$ and $a_i \in \{-1,+1\}, i = 1,\ldots,N$. The Svetlichny expression in this case can be written as

$$\mathcal{I}_N^{\mathrm{S}} = 2^{-N/2}\left[\sum_{\mathbf{x}|s \text{ is even}} (-1)^{s/2}\langle \mathbf{A}\rangle_{\mathbf{x}} + \sum_{\mathbf{x}|s \text{ is odd}} (-1)^{(s-1)/2}\langle \mathbf{A}\rangle_{\mathbf{x}}\right],$$

where $\langle \mathbf{A}\rangle_{\mathbf{x}} = \langle A_{1,x_1}\ldots A_{N,x_N}\rangle$ is the $N$-partite correlator, and $s$ is defined as $\sum_i x_i$. Similarly, the Mermin Bell expression can be reformulated as

$$\mathcal{I}_N^{\mathrm{M}} = 2^{-(N-1)/2}\left[\sum_{\mathbf{x}|s \text{ is even}} (-1)^{s/2}\langle \mathbf{A}\rangle_{\mathbf{x}}\right].$$

Next, we can construct Bell operators corresponding to these two Bell expressions. The Svetlichny inequality is maximally violated when the $N$ parties share the state $|\mathrm{GHZ}\rangle_N = (1/\sqrt{N})(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})$ and the two measurements for each party $i \in \{1,\ldots,N\}$ are

$$A_{i,0} = \cos\left(-\frac{\pi}{4N}\right)\sigma_x + \sin\left(-\frac{\pi}{4N}\right)\sigma_y, \quad A_{i,1} = \cos\left(\frac{(2N-1)\pi}{4N}\right)\sigma_x + \sin\left(\frac{(2N-1)\pi}{4N}\right)\sigma_y.$$

The Mermin inequality is also maximally violated by the state $|\mathrm{GHZ}\rangle_N$, and the two measurements of each party $i \in \{1,\ldots,N\}$ are $A_{i,0} = \sigma_x$ and $A_{i,1} = \sigma_y$. These two inequalities can detect genuine multipartite nonlocality effectively for any number of parties $N$, as the $(N-1)$-nonlocal bound is always larger than the maximal quantum bound [39].

By assigning the above measurement settings, the Bell operators of the Mermin and Svetlichny Bell inequalities are both expressed as

$$\mathcal{B}_N^{\mathrm{M}} = \mathcal{B}_N^{\mathrm{S}} = \mathcal{B}_N = 2^{(N-1)/2}[(|0\rangle\langle 1|)^{\otimes N} + (|1\rangle\langle 0|)^{\otimes N}]. \tag{A18}$$

The criterion for detecting Bell-correlation depth follows from Ref. [39]:

$$\langle \mathcal{B}_N \rangle \geq -2^{(N-\lceil N/k \rceil)/2}, \tag{A19}$$

where $k$ is the Bell-correlation depth. The violation of the inequality in Eq. (A19) certifies that the Bell-correlation depth is at least $k+1$.

## APPENDIX B: VARIATIONAL QUANTUM CIRCUITS

With the rapid development in the field of quantum computation, various quantum algorithms have been developed to tackle practical problems in various areas including combinatorial optimization, cryptography, and finance [47,75]. While quantum algorithms may bring potential advantages to these areas, many of them demand quantum hardware capabilities far beyond noisy intermediate-scale quantum devices. To alleviate this requirement, the hybrid quantum-classical framework has emerged as one of the leading strategies for NISQ devices. Within this framework, variational quantum algorithms (VQAs) have drawn broad interest over recent years [47,75–78]. In this appendix, we introduce the basic framework and optimization strategies of our VQAs.

### 1. Basic framework

Our VQA aims to approximate the ground-state energy of a quantum system Hamiltonian via a VQC, which shares the idea of variational quantum eigensolvers [76]. For any given quantum state $|\psi\rangle$, the expectation value of Hamiltonian $H$ is lower bounded by the ground-state energy $E_0$; i.e.,

$$\langle \psi | \mathcal{H} | \psi \rangle \geq E_0. \tag{B1}$$

Then, the ground-state energy can be approximately obtained by minimizing the expectation value of $\mathcal{H}$ with respect to the ansatz wave function $|\psi(\boldsymbol{\theta})\rangle$ parametrized by $\boldsymbol{\theta}$:

$$E(\boldsymbol{\theta}) = \langle \psi(\boldsymbol{\theta}) | \mathcal{H} | \psi(\boldsymbol{\theta}) \rangle. \tag{B2}$$

In the VQA scheme, the $N$-qubit quantum state $|\psi\rangle$ is prepared by applying a parametrized quantum unitary $U(\boldsymbol{\theta})$ to an initial state $|\mathbf{0}\rangle = |0\rangle^{\otimes N}$, which can be expressed as

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\mathbf{0}\rangle. \tag{B3}$$

In experiments, the quantum unitary $U(\boldsymbol{\theta})$ can be realized by a quantum circuit with the variational gate parameter $\boldsymbol{\theta}$ (Fig. 6). During the VQA procedure, $\boldsymbol{\theta}$ is optimized to minimize $E(\boldsymbol{\theta})$.

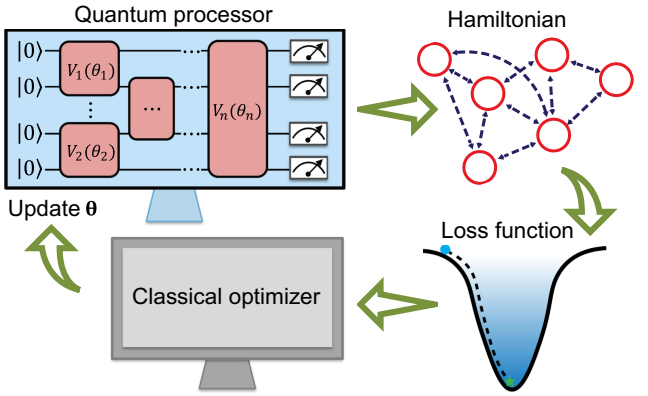


FIG. 6. Framework of the variational quantum algorithms. In VQC, quantum states are prepared on a quantum processor with parametrized gates. The expectation value of $H$ is acquired through quantum measurements, followed by classical postcalculation. The obtained result is then fed into a classical optimizer, which iteratively determines new parameter values for the quantum processor.

### 2. Optimization strategies

The aforementioned variational ansatz is optimized in experimental quantum devices. To this end, we first choose an appropriate loss function according to the practical problem and then minimize it to solve our problem. For variational quantum algorithms, our aim is to find the quantum state minimizing the expectation value of the target Hamiltonian, and the loss function is chosen as the expectation value of the energy,

$$L(\boldsymbol{\theta}) = E(\boldsymbol{\theta}) = \langle \psi(\boldsymbol{\theta}) | \mathcal{H} | \psi(\boldsymbol{\theta}) \rangle, \tag{B4}$$

where $|\psi(\boldsymbol{\theta})\rangle$ is the output state of the variational ansatz with parameters $\boldsymbol{\theta}$, and $\mathcal{H}$ is the target Hamiltonian. If we fix the initial state as $|\mathbf{0}\rangle$, and represent the variational ansatz as a parametrized unitary $U(\boldsymbol{\theta})$, the loss function becomes

$$L(\boldsymbol{\theta}) = E(\boldsymbol{\theta}) = \langle \mathbf{0} | U^\dagger(\boldsymbol{\theta}) \mathcal{H} U(\boldsymbol{\theta}) | \mathbf{0} \rangle. \tag{B5}$$

Then, the problem of finding the ground state for $\mathcal{H}$ is transformed to find the variational parameters $\boldsymbol{\theta}$ minimizing the loss function.

In practice, analytically finding the optimal variational parameters is very challenging. We usually update those parameters iteratively according to the instantaneous gradients until convergence, and the final parameters are regarded as the optimal solution. In our work, the variational ansatz consists of single-qubit rotation gates and two-qubit controlled-Z and controlled-NOT gates. The single-qubit rotation gate has form $R(\theta_k) = e^{-i\theta_k P/2}$ with $P$ chosen from the single Pauli operators $\{\sigma_x, \sigma_y, \sigma_z\}$, and $\theta_k$ being the rotation angle. All variational parameters are encoded in the rotation angles. The corresponding

derivative can be computed using the parameter shift rule [48,79],

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_k} = \frac{L(\boldsymbol{\theta}\backslash\theta_k, \theta_k + \pi/2) - L(\boldsymbol{\theta}\backslash\theta_k, \theta_k - \pi/2)}{2}, \tag{B6}$$

where $L(\boldsymbol{\theta}\backslash\theta_k, \theta_k \pm \pi/2)$ denotes the loss function where $\theta_k$ is replaced by $\theta_k \pm \pi/2$ and the other parameters are unchanged. We remark that this analytical (exact and no bias) method can be implemented in current quantum devices to provide the gradients through measurements. More optimization methods may be explored in the future [80–86].

With gradients at hand, we can update the variational parameters using the gradient descent method. As an illustrative example, the direct gradient descent works by

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \epsilon \nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}_n), \tag{B7}$$

where $\boldsymbol{\theta}_i$ is the parameters at $i$th step and $\epsilon$ is a small scalar representing the update step length. In this work, we take the Adam[0.9, 0.999] optimizer to handle the obtained gradients, where the momentum-based method further improves the training performance [87].

### 3. Nonlocal correlation detection

Here, we introduce the basic ideas of utilizing variational quantum algorithms to detect Bell-operator correlations. A typical Bell inequality has form

$$\mathcal{I} - \beta_C \geq 0, \tag{B8}$$

where $\mathcal{I}$ includes a collection of expectation values of correlators and $\beta_C$ characterizes the limits of the local-hidden-variable model. For states admitting local-hidden-variable descriptions, the inequality is always valid. However, certain quantum states and measurements may yield correlations beyond the local-hidden-variable description and a Bell inequality would be violated. For a given set of measurements, the Bell inequality can be connected to a Bell operator that can be interpreted as a Hamiltonian [45]. If there exist quantum states violating such a Bell inequality, i.e., $\langle \mathcal{H} \rangle < \beta_C$, the energy can be regarded as a witness of Bell nonlocality [45].

In this work, we extend this idea and adapt the variational quantum circuit as the variational ansatz to prepare low-energy states for the target Hamiltonian. We consider a standard Bell inequality in which the quantum correlators are expectation values for Pauli strings acting on different parties,

$$\mathcal{I} = \sum_k \langle \alpha_k M_k \rangle \geq \beta_C, \tag{B9}$$

where $M_k$ is a Pauli string and $\alpha_k$ is the corresponding coefficient. The Hamiltonian is

$$\mathcal{H} = \sum_k \alpha_k M_k. \tag{B10}$$

Within this framework, our general recipe for detecting the Bell-operator correlations through variational quantum algorithms is as follows.

(1) We begin with a fixed initial state and a fixed quantum circuit ansatz including initialized variational parameters. At the initial stage, the output state typically does not violate the given inequality.
(2) We then use the optimization strategies described before to iteratively update the variational parameters, such that the energy of the output state is minimized.

If the final minimal energy is smaller than $\beta_C$, the corresponding Bell inequality is successfully violated and nonlocality is detected, while the maximal violation of the Bell inequality (the ground state) may not always be achieved due to experimental noises and limitations of the VQC's expressive power [88–91].

## APPENDIX C: DEVICE INFORMATION

The superconducting quantum processor used in this work has $11 \times 11$ transmon qubits arranged in a square lattice and 220 tunable couplers each inserted in between neighboring two qubits [37]. Each qubit has individual microwave control and flux biases for single-qubit gates and frequency tunability, respectively. The coupling strength between any neighboring pair of qubits is dynamically tunable, allowing for high-fidelity two-qubit gates. Each qubit is dispersively coupled to a readout resonator for its state measurement. In our experiments, we use up to 73 qubits to construct the variational quantum circuit. The typical energy relaxation time $T_1$ and Hahn echo dephasing time $T_2$ of these qubits are shown in Fig. 7, with median values of $T_1 = 81.4$ μs and $T_2 = 23.5$ μs, respectively. The qubit layouts used for the three different experiments in this work are shown in Fig. 8.

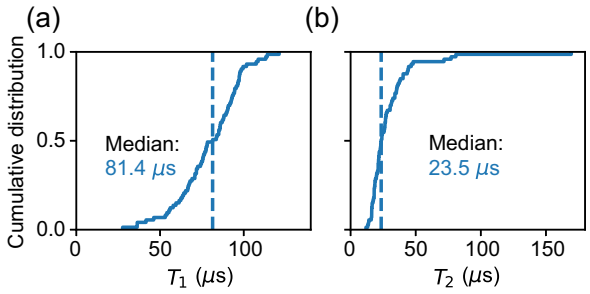


FIG. 7. Typical $T_1$ and $T_2$. (a) Cumulative distribution of $T_1$ across the 73 qubits used on the superconducting quantum processor with a dashed line indicating a median value of 81.4 μs. (b) Cumulative distribution of $T_2$ across the 73 qubits used on the superconducting quantum processor with a dashed line indicating a median value of 23.5 μs.

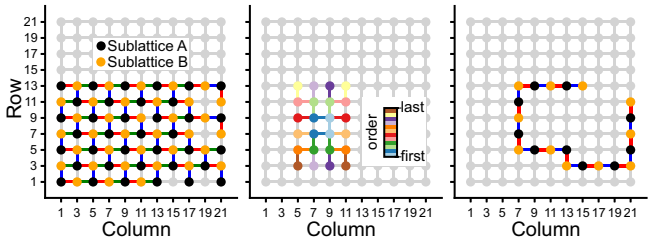


FIG. 8. Qubit layouts. Left: the 73 qubits used in the two-dimensional CHSH honeycomb model with qubit coordinates indicated as (row, column). Middle: the 24 qubits used to detect Bell-operator correlation depth. The order of training is indicated using colors, starting from two-qubit training and adding 2 qubits each time. Right: the 21 qubits used in the 1D XXZ model. The red thick bars indicate the qubit pairs with strong coupling strength, while the blue thin bars indicate weak coupling strength.

## APPENDIX D: EXPERIMENT CIRCUIT

The experimental circuit for the variational quantum algorithm contains multiple layers of single-qubit gates and two-qubit CNOT gates with different patterns. In the experiment, we realize arbitrary single-qubit gates by combining XY rotations and Z rotations, with the XY rotations implemented with 30-ns-long microwave pulses optimized according to the derivative reduction by adiabatic gate (DRAG) control theory [92] at the qubits' respective idle frequencies, and the Z rotations realized by the virtual-Z gate scheme [50]. The CNOT gate is decomposed to a generic two-qubit controlled $\pi$-phase (CZ) gate sandwiched between two Hadamard gates. We realize the CZ gates by bringing $|11\rangle$ and $|02\rangle$ (or $|20\rangle$) of the qubit pairs in near resonance and tuning the coupler to activate the effective ZZ interaction for a specific time of 30 ns via the flux controls [88]. As such, the circuit is recompiled with the native gate set $\{U(\theta,\varphi,\lambda), \mathrm{CZ}\}$, where the arbitrary single-qubit gates are parametrized as

$$U(\theta,\varphi,\lambda) = \begin{pmatrix} \cos\frac{\theta}{2} & -e^{i\lambda}\sin\frac{\theta}{2} \\ e^{i\varphi}\sin\frac{\theta}{2} & e^{i(\varphi+\lambda)}\cos\frac{\theta}{2} \end{pmatrix}. \quad \text{(D1)}$$

In our experiment, we achieve high-fidelity parallel quantum gates and projective measurements, which is essential for the demonstration of the efficacy of the VQC-based many-body Bell-operator correlation detector. For a given quantum circuit, we first reschedule the gate sequences to make them as compact as possible. Since the single-qubit gates are realized via microwave controls and the two-qubit gates involve only flux controls, we separate them into different layers to avoid unwanted crosstalk errors. Besides, for the two-qubit gate layers, we optimize the work points for every two-qubit gate as well as the idling qubits simultaneously to minimize the errors due to parasitic couplings. The performances of the single- and two-qubit gates are assessed with simultaneous cross-entropy benchmark techniques. To reduce errors of the readout procedure, we implement the excited state promotion technique [93,94], which essentially leverages the higher energy levels of the transmon qubit to reduce the state relaxation error and improve the signal-to-noise ratio of the readout procedure. In particular, we apply an $X_{21}$ pulse, which rotates the transmon in the $|2\rangle - |1\rangle$ subspace by an angle $\pi$, before applying the readout pulses. We conduct simultaneous benchmarking of single- and two-qubit gate error and readout error for each of the three different experiments. The obtained results are shown in Fig. 9. In the simultaneous two-qubit gate layer, we optimize the energy level of qubits that are not involved in the CZ gates by performing Z gates to avoid interaction with surrounding qubits [95].

Readout correction is widely used in the literature to assess the intrinsic property of the experimentally prepared states. However, to rigorously test the Bell inequality, one should use single-shot measurement to exclude the detection

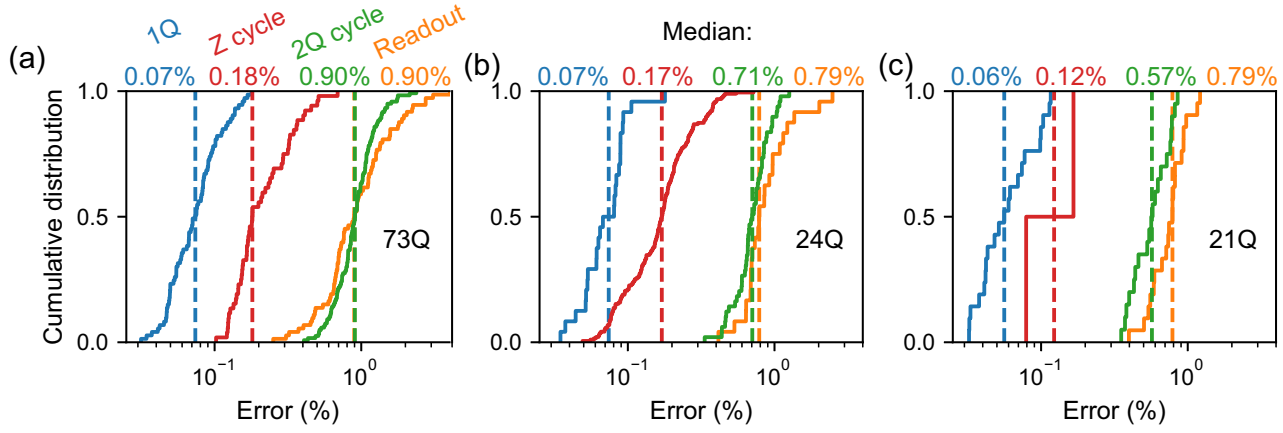


FIG. 9. Cumulative distributions of errors for simultaneous single-qubit gates (1Q), Z gates (Z cycle, including a Z gate and a single-qubit gate), two-qubit gates (2Q cycle, including a CZ gate and two single-qubit gates), and readout (the average readout error of $|0\rangle$ and $|1\rangle$). (a) Seventy-three qubits used in the two-dimensional CHSH honeycomb model. (b) Twenty-four qubits used to detect Bell-operator correlation depth. (c) Twenty-one qubits used in the 1D XXZ model.

TABLE I. Comparison of results with and without readout correction.

| Experiment | Raw | Violation | Corrected | Violation |
|---|---|---|---|---|
| Sec. III | $-163.7 \pm 0.5$ | 7 SD | $-168.9 \pm 0.5$ | 17 SD |
| Sec. IV | $-156.29 \pm 0.52$ | 48 SD | $-162.41 \pm 0.54$ | 57 SD |
| Sec. V | $-2112.7 \pm 7.7$ | 8 SD | $-2304.7 \pm 8.5$ | 30 SD |

loophole. Therefore, any data postprocessing procedures such as readout correction should be avoided. As a result, it is common to quote the results without readout correction to demonstrate the robustness of the conclusions, while also appending the results with readout correction for reference (see, e.g., Ref. [96]). In this work, we follow this convention and mainly analyze the results without readout correction. A comparison of the results with and without readout correction for the three experiments in this work are listed in Table I.

## APPENDIX E: MANY-BODY QUANTUM CORRELATION MEASUREMENT

The detection of many-body quantum correlations is typically challenging in experiments. In our study, we employ two techniques, namely parity oscillation and multiple quantum coherence, to probe the many-body quantum correlations in our system, with the corresponding quantum circuits illustrated in Fig. 10. Below we introduce the two methods and analyze the corresponding advantages and disadvantages in detail.

### 1. Parity oscillation measurement

The expectation value of the multiqubit operator $\mathcal{B}_N$, Eq. (A18), can be written as $\langle \mathcal{B}_N \rangle = 2^{(N-1)/2}(\langle \mathcal{C} \rangle + \langle \mathcal{C} \rangle^*) = 2^{(N+1)/2}\mathrm{Re}(\langle \mathcal{C} \rangle)$, where the matrix element $\mathcal{C} \equiv |0\rangle\langle 1|^{\otimes N}$, and $\mathrm{Re}(x)$ denotes the real part of $x$. Experimentally, $\langle \mathcal{C} \rangle$ can be obtained by parity measurement [52]. Consider the $N$-qubit parity operator $\mathcal{P}(\gamma) = \otimes_{j=1}^{N} [\sin(\gamma)\sigma_{y,j} + \cos(\gamma)\sigma_{x,j}]$, for a given density matrix $\rho$; its expectation value can be expressed as

$$\langle \mathcal{P}(\gamma) \rangle = \sum_{\boldsymbol{s}} \rho_{\boldsymbol{s},\bar{\boldsymbol{s}}} e^{i(2\mathbb{Z}_{\boldsymbol{s}} - N)\gamma} = \sum_{\boldsymbol{s}} |\rho_{\boldsymbol{s},\bar{\boldsymbol{s}}}| \cos\left[(2\mathbb{Z}_{\boldsymbol{s}} - N)\gamma + \phi_{\boldsymbol{s},\bar{\boldsymbol{s}}}\right], \quad \text{(E1)}$$

where $\boldsymbol{s} = s_1 s_2 \ldots s_N (s_j \in \{0,1\})$ denotes a bit string and $\bar{\boldsymbol{s}}$ represents its complement (e.g., $\boldsymbol{s} = 00\ldots0$ and $\bar{\boldsymbol{s}} = 11\ldots1$; hence $\rho_{\boldsymbol{s},\bar{\boldsymbol{s}}}$ is the antidiagonal term in the density matrix), $\phi_{\boldsymbol{s},\bar{\boldsymbol{s}}}$ is the phase of $\rho_{\boldsymbol{s},\bar{\boldsymbol{s}}}$, and $\mathbb{Z}_{\boldsymbol{s}}$ is the number of 0 in $\boldsymbol{s}$. By noticing that $\langle \mathcal{C} \rangle = \rho_{00\ldots0,11\ldots1}$, the correlation can be determined as the Fourier components $\mathcal{P}_f(N), \mathcal{P}_f(-N)$ of $\langle \mathcal{P}(\gamma) \rangle$, where $\mathcal{P}_f(q) = \mathcal{N}_s^{-1} \sum_\gamma e^{-iq\gamma} \langle \mathcal{P}(\gamma) \rangle$ is the discrete Fourier transformation and $\mathcal{N}_s$ is the sampling number of $\gamma$.

Experimentally we take a sparse sampling of $\langle \mathcal{P}(\gamma) \rangle$ with $\mathcal{N}_s = N + 1$ points, where $\gamma = -(\pi/2) + (\pi/N + 1)k$ $(k = 0, 1, \ldots, N)$. The quantum circuit is illustrated in Fig. 10(a). For each $\gamma$ we apply a single-qubit gate $U(\pi/2, \gamma - \pi, \pi - \gamma) = (\sqrt{2}/2)\begin{pmatrix} 1 & e^{-i\gamma} \\ -e^{i\gamma} & 1 \end{pmatrix}$ to each qubit before simultaneous readout. These rotations bring the axis defined by the operator $\mathcal{P}(\gamma)$ to the $z$ axis. Subsequently, we perform simultaneous measurements to all qubits and record the outcomes as binary strings over $M$ shots. From these measurements, we compute the parity expectation $\langle \mathcal{P}(\gamma) \rangle = (M_+ - M_-/M)$, where $M_+$ ($M_-$) represents the count of occurrences corresponding to eigenvalue 1 ($-1$). For each $\langle \mathcal{P}(\gamma) \rangle$, we take 900, 1500, 2400, 3600, 5000, 7200, 9600, 12 000, 15 000, 20 000, 25 000, 30 000 measurement shots for qubit number $N = 2, 4, 6, \ldots, 24$. To correct the readout error, we use the tensor product of single-qubit correction matrices.

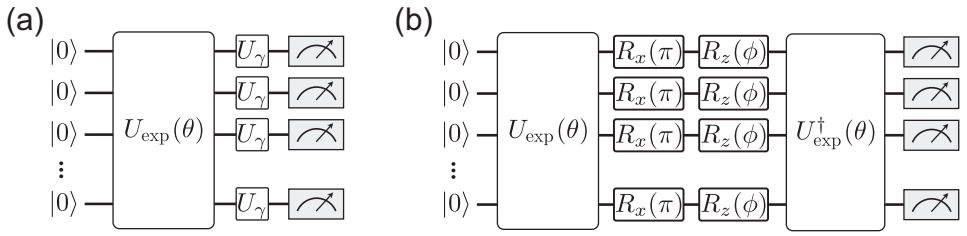


FIG. 10. Quantum circuits for measuring multiqubit correlations. (a) Circuit for measuring parity oscillation. We add a single-qubit gate $U_\gamma = (\sqrt{2}/2)\begin{pmatrix} 1 & e^{-i\gamma} \\ -e^{i\gamma} & 1 \end{pmatrix}$ to each qubit at the end of the experimental variational circuit $U_{\exp}(\theta)$ depicted in Fig. 4(a) of the main text, which brings the axis defined by the operator $\mathcal{P}(\gamma) = \sin(\gamma)\sigma_y + \cos(\gamma)\sigma_x$ to the $z$ axis. Subsequently, we perform simultaneous measurements on all qubits and calculate the parity expectations with different $\gamma$ values. (b) Circuit for measuring MQC. After applying the variational circuit $U_{\exp}(\theta)$, we add a single-qubit $R_x(\pi)$ gate and $R_z(\phi)$ gate to each qubit followed by a reversal circuit $U^\dagger_{\exp}(\theta)$. Then we measure the probability of the system being in the $|00\ldots0\rangle$ state.

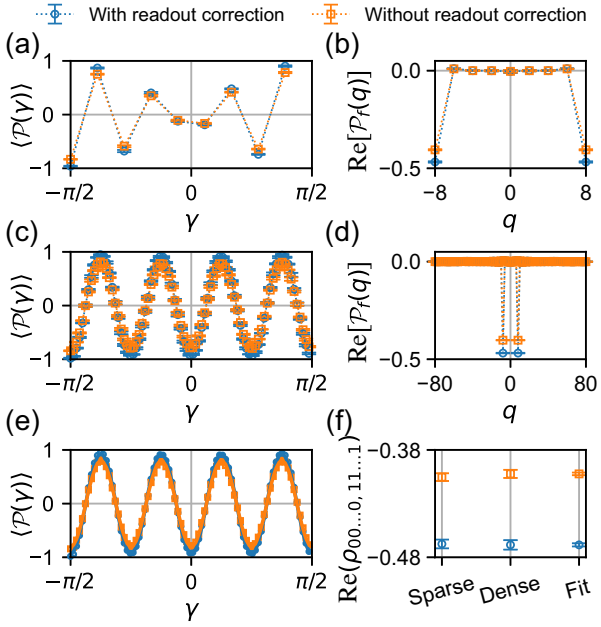


FIG. 11. Contrast three different methods for analyzing eight-qubit parity oscillation measurement results: sparse sampling, dense sampling, and sinusoid fitting. (a),(b) Sparse sampling with sampling size $\mathcal{N}_s = 9$ and the real part of its Fourier spectra. The Fourier component $\mathrm{Re}[\mathcal{P}_f(8)]$ equals to $\mathrm{Re}(\rho_{00\ldots0,11\ldots1})$. (c),(d) Dense sampling with sampling size $\mathcal{N}_s = 81$ and the real part of its Fourier spectra. (e) Sinusoid fit applied to the dense sampling data, with the frequency set to 8. (f) Comparison of the outcomes obtained from the three different methods.

In order to test the accuracy and reliability of this sparse sampling method, we compare the outcomes of sparse sampling ($\mathcal{N}_s = N + 1$) with dense sampling ($\mathcal{N}_s = 10N + 1$) case for an experimentally trained eight-qubit GHZ state. In the dense sampling case, we apply both the Fourier transformation and sinusoid fitting to the experimental data, with the results shown in Fig. 11. The extracted correlations are similar across all three cases.

### 2. Multiple quantum coherence

The parity oscillation measurement offers an efficient way to detect the many-body correlation in this work. Now we introduce another technique, i.e., the multiple quantum coherence [53], which can also measure the many-body correlation efficiently. In addition, the MQC method is less susceptible to readout errors, as we will elaborate in the next section.

Following the variationally obtained $N$-qubit state $\rho = U_{\mathrm{exp}}|00\ldots0\rangle\langle00\ldots0|U_{\mathrm{exp}}^{\dagger}$, we introduce an $X$ gate $U_X$ and a phase gate $U_\phi = e^{-i(\phi/2)\sigma_z} = \left(\begin{smallmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{smallmatrix}\right)$ to each qubit, succeeded by a reversal circuit $U_{\mathrm{exp}}^{\dagger}$ [Fig. 10(b)]. After all these operations, the probability of detecting the ground state becomes

$$\begin{aligned}\mathcal{K}(\phi) &= |\langle00\ldots0|U_{\mathrm{exp}}^{\dagger}U_{\phi}^{\otimes N}U_{X}^{\otimes N}U_{\mathrm{exp}}|00\ldots0\rangle|^2 \\ &= \sum_{\boldsymbol{s},\boldsymbol{s}'}\rho_{\boldsymbol{s},\boldsymbol{s}'}\rho_{\bar{\boldsymbol{s}}',\bar{\boldsymbol{s}}}e^{i(\mathbb{Z}_{\boldsymbol{s}}-\mathbb{Z}_{\boldsymbol{s}'})\phi}\end{aligned} \tag{E2}$$

where $\boldsymbol{s}$ ($\boldsymbol{s}'$) denotes a bit string and $\bar{\boldsymbol{s}}$ ($\bar{\boldsymbol{s}}'$) represents its complement, and $\mathbb{Z}_{\boldsymbol{s}}$ ($\mathbb{Z}_{\boldsymbol{s}'}$) is the number of 0 in $\boldsymbol{s}$ ($\boldsymbol{s}'$).

Noting that the oscillation terms with the maximum frequency $N$ ($-N$) in $\mathcal{K}(\phi)$ occur only for $\boldsymbol{s} = 00\ldots0$ ($\boldsymbol{s} = 11\ldots1$) and $\boldsymbol{s}' = 11\ldots1$ ($\boldsymbol{s}' = 00\ldots0$), and that $\rho_{00\ldots0,11\ldots1} = \rho^{*}_{11\ldots1,00\ldots0}$, we can obtain $|\rho_{00\ldots0,11\ldots1}|$ by Fourier transforming $\mathcal{K}(\phi)$ at the frequency $N$ as

$$|\rho_{00\ldots0,11\ldots1}| = \frac{1}{\sqrt{\mathcal{N}_s}}\sqrt{\left|\sum_{\phi}e^{-iN\phi}\mathcal{K}(\phi)\right|}, \tag{E3}$$

where the normalization constant $\mathcal{N}_s$ represents the sampling number of $\phi$. The phase of $\rho_{00\ldots0,11\ldots1}$ cannot be directly obtained due to potential imperfections in the reversal circuit. Therefore, we rely on the phase extracted from the parity oscillation measurement to calculate the real part of $\rho_{00\ldots0,11\ldots1}$.

Experimentally we only employ MQC as a means to detect genuine multipartite Bell-operator correlations on the variationally trained state. To accurately measure $\mathcal{K}_f(N)$ without compromising efficiency, we adopt sparse sampling on $\phi$, where $\phi = (\pi j/N + 1)$ for $j = 0, 1, 2, \ldots, 2N + 1$ (note that the number of sampling points is approximately double that of the parity oscillation measurement). For each $\mathcal{K}(\phi)$, we employ an equal number of measurement shots as utilized in the parity measurement. Moreover, we implement the same readout correction procedure as employed in the parity measurement.

### 3. Effect of readout error

As mentioned earlier, the parity oscillation measurement is more susceptible to readout errors. This is primarily due to two reasons. First, the parity is calculated from the entire set of $2^N$ probabilities, whereas MQC requires only measuring the probability of the ground state $|00\ldots0\rangle$. In addition, the MQC circuit involves a reversal circuit, which mitigates the impact of readout errors by taking the square root of the Fourier outcomes.

The projective quantum state readout procedure in superconducting circuits is error prone, meaning that we may assign a wrong state $|1\rangle$ ($|0\rangle$) to the qubit while it is actually in $|0\rangle$ ($|1\rangle$) with the probability $e_0$ ($e_1$). To understand the effect of readout error, let us first assume that $e_0 = e_1 = e$. As a reference, the experimentally calibrated median values of $e_0$ and $e_1$ are 0.8% and 0.9%, respectively. For the MQC method, if the variational circuit $U_{\mathrm{exp}}(\boldsymbol{\theta})$ perfectly generates a GHZ state, then the final state should be $\frac{1}{2}[(1 + e^{-iN\phi})|0\rangle + (1 - e^{-iN\phi})|1\rangle] \otimes |00\ldots0\rangle$, where only two bit strings, i.e., 00…0 and 10…0,

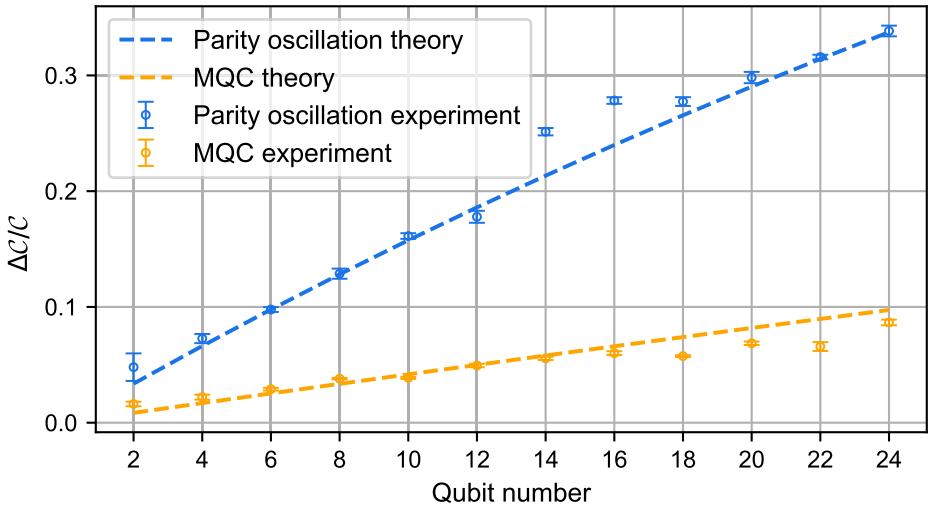


FIG. 12. The effect of readout error. We plot the effect of readout error with different qubit numbers estimated based on experimental data (dots with error bars) and theoretical calculation (dashed lines) for both the MQC (orange) and parity oscillation (blue) methods. For the experimental data, the effect of readout error is estimated as the ratio between the correlation degradation due to the readout error ($\Delta\mathcal{C}$) and the correlation after mitigating the readout error ($\mathcal{C}$). For the theoretical data, the effect of readout error can be estimated as $1-(1-e)^{N/2}$ and $1-(1-2e)^N$ for the MQC and parity oscillation methods, respectively. We use an error rate of 0.0085 for both the $|0\rangle$ and $|1\rangle$ states to estimate the theoretical values.

appear. In the presence of readout errors, the measured ground-state probability is given by

$$\begin{aligned}\mathcal{K}(\phi)_{\text{meas}} &= (1-e)^N P_{00\ldots0} + e(1-e)^{N-1}P_{10\ldots0} \\ &\approx (1-e)^N \mathcal{K}(\phi),\end{aligned} \tag{E4}$$

where $P_s$ denotes the ideal probability of measuring the bit string state $|s\rangle$. The approximation arises because $e(1-e)^{N-1}P_{10\ldots0}$ is much smaller than 1 due to $e \ll 1$, and therefore it can be ignored. In the experiment, there may be other measured bit strings besides 00…0 and 10…0 due to imperfections of the quantum gate or finite coherence time. Nonetheless, their occurrences are significantly less frequent. Combined with Eq. (E3), we see that due to readout error, the experimentally obtained correlation $|\rho_{00\ldots0,11\ldots1}|$ is reduced by a factor of $(1-e)^{N/2}$.

On the other hand, the parity expectations are calculated from the entire set of $2^N$ probabilities, which can be separated by their parity eigenvalue 1 $(-1)$ into two groups. Let us assume we take $M$ measurement shots, and the count of occurrences corresponding to eigenvalue 1 $(-1)$ is $M_+$ $(M_-)$. In the presence of readout errors, the experimentally measured parity expectation can be calculated as

$$\begin{aligned}\langle\mathcal{P}(\gamma)\rangle_{\text{meas}} &= \frac{[(1-e)^N M_+ + C_N^1 e(1-e)^{N-1}M_- + C_N^2 e^2(1-e)^{N-2}M_+ + \cdots]}{M} \\ &\quad - \frac{[(1-e)^N M_- + C_N^1 e(1-e)^{N-1}M_+ + C_N^2 e^2(1-e)^{N-2}M_- + \cdots]}{M} \\ &= \Sigma_{i=0}^N C_N^i (-e)^i (1-e)^{N-i}\frac{M_+ - M_-}{M} \\ &= (1-2e)^N \langle\mathcal{P}(\gamma)\rangle.\end{aligned} \tag{E5}$$

The first equality is based on the observation that the sign of eigenvalue will change if an odd number of qubits experience a bit-flip readout error.

Finally, the ratio of outcomes between MQC and parity measurement is approximately $(1-e)^{N/2}/(1-2e)^N$. We compare this estimation with the experimental results for different qubit numbers, as shown in Fig. 12.

### 4. Statistical analysis

For the measured many-body quantum correlation to carry statistical meaning, it needs to be supplemented with an estimate of its standard error. A smaller standard error indicates higher confidence in the measurement results and reduces the number of shots required to assert genuine multipartite Bell-operator correlations. Experimentally, we

observe a ratio of approximately 1:4 between the standard errors from MQC and parity measurements of the measured many-body quantum correlation. This observation can be attributed to the following three main factors.

First, the parity expectation is calculated from $\langle \mathcal{P}(\gamma)\rangle = (M_+ - M_-/M) = (2M_+/M) - 1$, where $M_+$ follows a binomial distribution, whereas in the MQC measurement the outcomes are obtained directly from the ground-state probability, which also follows a binomial distribution. The standard deviation of the parity expectation is given by $\sigma(\langle \mathcal{P}(\gamma)\rangle) = 2\sqrt{p(1-p)/M}$, where $p$ is the probability of getting eigenvalue 1 in a single shot. If we assume in the MQC measurement the probability of observing the ground state is also $p$ and take the same number of measurement shots, the standard deviation of the MQC outcome is only $\sqrt{p(1-p)/M}$.

Second, the number of sampling points in the discrete Fourier transformation of MQC is approximately double that of the parity oscillation measurement, which reduces the standard deviation by a ratio of $\sqrt{2}$. However, it is important to note that this reduction in standard deviation is actually due to the total increase in the number of measurement shots taken by the MQC measurement, rather than being an inherent advantage of the MQC technique itself.

Third, observation from Eq. (E4) shows that the readout error changes the range of ground-state probability from symmetry around 0.5 to below it (the standard deviation of a binomial distribution is maximal when $p = 0.5$). In contrast, in the parity oscillation measurement, the readout error causes the parity expectation to remain symmetric around 0 (the standard deviation is maximal when the parity expectation is 0). Therefore, the MQC benefits from this asymmetry induced by the readout error, leading to a further reduction in the standard deviation.


[1] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777 (1935).

[2] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum entanglement*, Rev. Mod. Phys. **81**, 865 (2009).

[3] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, *Colloquium: The Einstein-Podolsky-Rosen paradox: From concepts to applications*, Rev. Mod. Phys. **81**, 1727 (2009).

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

[5] H. M. Wiseman, S. J. Jones, and A. C. Doherty, *Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox*, Phys. Rev. Lett. **98**, 140402 (2007).

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

[7] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. **23**, 880 (1969).

[8] R. F. Werner, *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, Phys. Rev. A **40**, 4277 (1989).

[9] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, *Entanglement and nonlocality are inequivalent for any number of parties*, Phys. Rev. Lett. **115**, 030404 (2015).

[10] 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. **98**, 230501 (2007).

[11] U. Vazirani and T. Vidick, *Fully device-independent quantum key distribution*, Phys. Rev. Lett. **113**, 140501 (2014).

[12] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, *Secure quantum key distribution with realistic devices*, Rev. Mod. Phys. **92**, 025002 (2020).

[13] J. L. Bönsel, O. Gühne, and A. Cabello, *Generating multipartite nonlocality to benchmark quantum computers*, Phys. Rev. A **111**, 012207 (2025).

[14] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, *Random numbers certified by Bell's theorem*, Nature (London) **464**, 1021 (2010).

[15] S. Bravyi, D. Gosset, and R. König, *Quantum advantage with shallow circuits*, Science **362**, 308 (2018).

[16] F. Le Gall, *Average-case quantum advantage with shallow circuits*, in *Proceedings of the 34th Computational Complexity Conference (CCC), Dagstuhl, Germany, 2019* (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2020), pp. 1–20.

[17] X. Gao, E. R. Anschuetz, S.-T. Wang, J. I. Cirac, and M. D. Lukin, *Enhancing generative models via quantum correlations*, Phys. Rev. X **12**, 021037 (2022).

[18] Z. Zhang, W. Gong, W. Li, and D.-L. Deng, *Quantum-classical separations in shallow-circuit-based learning with and without noises*, Commun. Phys. **7**, 290 (2024).

[19] S. J. Freedman and J. F. Clauser, *Experimental test of local hidden-variable theories*, Phys. Rev. Lett. **28**, 938 (1972).

[20] A. Aspect, P. Grangier, and G. Roger, *Experimental tests of realistic local theories via Bell's theorem*, Phys. Rev. Lett. **47**, 460 (1981).

[21] A. Aspect, P. Grangier, and G. Roger, *Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A new violation of Bell's inequalities*, Phys. Rev. Lett. **49**, 91 (1982).

[22] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, *Violation of Bell's inequality under strict Einstein locality conditions*, Phys. Rev. Lett. **81**, 5039 (1998).

[23] M. Giustina, M. A. M. Versteegh, S. Wengerowsky *et al.*, *Significant-loophole-free test of Bell's theorem with entangled photons*, Phys. Rev. Lett. **115**, 250401 (2015).

[24] L. K. Shalm, E. Meyer-Scott, B. G. Christensen *et al.*, *Strong loophole-free test of local realism*, Phys. Rev. Lett. **115**, 250402 (2015).

[25] M.-H. Li, C. Wu, Y. Zhang *et al.*, *Test of local realism into the past without detection and locality loopholes*, Phys. Rev. Lett. **121**, 080404 (2018).

[26] B. Hensen, H. Bernien, A. E. Dréau *et al.*, *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres*, Nature (London) **526**, 682 (2015).

[27] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter, *Heralded entanglement between widely separated atoms*, Science **337**, 72 (2012).

[28] S. Storz, J. Schär, A. Kulikov *et al.*, *Loophole-free Bell inequality violation with superconducting circuits*, Nature (London) **617**, 265 (2023).

[29] H. M. Hill, *Physics Nobel honors foundational quantum entanglement experiments*, Phys. Today **75**, No. 12, 14 (2022).

[30] Y.-L. Mao, Z.-D. Li, S. Yu, and J. Fan, *Test of genuine multipartite nonlocality*, Phys. Rev. Lett. **129**, 150401 (2022).

[31] H. Cao, M.-O. Renou, C. Zhang *et al.*, *Experimental demonstration that no tripartite-nonlocal causal theory explains nature's correlations*, Phys. Rev. Lett. **129**, 150402 (2022).

[32] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, *Entanglement in many-body systems*, Rev. Mod. Phys. **80**, 517 (2008).

[33] F. Levi and F. Mintert, *Hierarchies of multipartite entanglement*, Phys. Rev. Lett. **110**, 150402 (2013).

[34] R. Schmied, J.-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, and N. Sangouard, *Bell correlations in a Bose-Einstein condensate*, Science **352**, 441 (2016).

[35] N. J. Engelsen, R. Krishnakumar, O. Hosten, and M. A. Kasevich, *Bell correlations in spin-squeezed states of 500000 atoms*, Phys. Rev. Lett. **118**, 140401 (2017).

[36] J. Preskill, *Quantum computing in the NISQ era and beyond*, Quantum **2**, 79 (2018).

[37] S. Xu, Z.-Z. Sun, K. Wang *et al.*, *Digital simulation of projective non-Abelian anyons with 68 superconducting qubits*, Chin. Phys. Lett. **40**, 060301 (2023).

[38] G. Svetlichny, *Distinguishing three-body from two-body nonseparability by a Bell-type inequality*, Phys. Rev. D **35**, 3066 (1987).

[39] F. Baccari, J. Tura, M. Fadel, A. Aloy, J.-D. Bancal, N. Sangouard, M. Lewenstein, A. Acín, and R. Augusiak, *Bell correlation depth in many-body systems*, Phys. Rev. A **100**, 022121 (2019).

[40] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, *Detecting nonlocality in many-body quantum states*, Science **344**, 1256 (2014).

[41] J.-D. Bancal, C. Branciard, N. Gisin, and S. Pironio, *Quantifying multipartite nonlocality*, Phys. Rev. Lett. **103**, 090503 (2009).

[42] N. D. Mermin, *Extreme quantum entanglement in a superposition of macroscopically distinct states*, Phys. Rev. Lett. **65**, 1838 (1990).

[43] M. Żukowski and Č. Brukner, *Bell's theorem for general N-qubit states*, Phys. Rev. Lett. **88**, 210401 (2002).

[44] I. Frérot, M. Fadel, and M. Lewenstein, *Probing quantum correlations in many-body systems: A review of scalable methods*, Rep. Prog. Phys. **86**, 114001 (2023).

[45] J. Tura, G. De las Cuevas, R. Augusiak, M. Lewenstein, A. Acín, and J. I. Cirac, *Energy as a detector of nonlocality of many-body spin systems*, Phys. Rev. X **7**, 021005 (2017).

[46] O. Gühne, G. Tóth, P. Hyllus, and H. J. Briegel, *Bell inequalities for graph states*, Phys. Rev. Lett. **95**, 120405 (2005).

[47] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, *Variational quantum algorithms*, Nat. Rev. Phys. **3**, 625 (2021).

[48] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, *Quantum circuit learning*, Phys. Rev. A **98**, 032309 (2018).

[49] N. Gisin, *Bell inequalities: Many questions, a few answers*, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle*, Vol. 73 (Springer, Dordrecht, 2009), pp. 125–138.

[50] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, *Efficient z gates for quantum computing*, Phys. Rev. A **96**, 022330 (2017).

[51] D.-L. Deng, *Machine learning detection of Bell nonlocality in quantum many-body systems*, Phys. Rev. Lett. **120**, 240402 (2018).

[52] C. Song, K. Xu, H. Li *et al.*, *Generation of multicomponent atomic Schrödinger cat states of up to 20 qubits*, Science **365**, 574 (2019).

[53] K. X. Wei, I. Lauer, S. Srinivasan, N. Sundaresan, D. T. McClure, D. Toyli, D. C. McKay, J. M. Gambetta, and S. Sheldon, *Verifying multipartite entangled Greenberger-Horne-Zeilinger states via multiple quantum coherences*, Phys. Rev. A **101**, 032343 (2020).

[54] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, *Barren plateaus in quantum neural network training landscapes*, Nat. Commun. **9**, 4812 (2018).

[55] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, *Greedy layer-wise training of deep networks*, in *Advances in Neural Information Processing Systems*, Vol. 19 (MIT Press, Cambridge, MA, 2006).

[56] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, *Group-invariant quantum machine learning*, PRX Quantum **3**, 030341 (2022).

[57] J. J. Meyer, M. Mularski, E. Gil-Fuster, A. A. Mele, F. Arzani, A. Wilms, and J. Eisert, *Exploiting symmetry in variational quantum machine learning*, PRX Quantum **4**, 010328 (2023).

[58] H. Zheng, Z. Li, J. Liu, S. Strelchuk, and R. Kondor, *Speeding up learning quantum states through group equivariant convolutional quantum Ansätze*, PRX Quantum **4**, 020327 (2023).

[59] X. Mi, A. A. Michailidis, S. Shabani *et al.*, *Stable quantum-correlated many-body states through engineered dissipation*, Science **383**, 1332 (2024).

[60] E. Farhi, J. Goldstone, and S. Gutmann, *A quantum approximate optimization algorithm*, arXiv:1411.4028.

[61] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, *Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices*, Phys. Rev. X **10**, 021067 (2020).

[62] M. Dupont, B. Evert, M. J. Hodson *et al.*, *Quantum-enhanced greedy combinatorial optimization solver*, Sci. Adv. **9**, eadi0487 (2023).

[63] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, *Event-ready Bell test using entangled atoms simultaneously closing*

*detection and locality loopholes*, Phys. Rev. Lett. **119**, 010402 (2017).
[64] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, *Optimal randomness certification from one entangled bit*, Phys. Rev. A **93**, 040102(R) (2016).
[65] B. S. Cirel'son, *Quantum generalizations of Bell's inequality*, Lett. Math. Phys. **4**, 93 (1980).
[66] S. Popescu and D. Rohrlich, *Quantum nonlocality as an axiom*, Found. Phys. **24**, 379 (1994).
[67] A. Fine, *Hidden variables, joint probability, and the Bell inequalities*, Phys. Rev. Lett. **48**, 291 (1982).
[68] K. T. Goh, J. Kaniewski, E. Wolfe, T. Vértesi, X. Wu, Y. Cai, Y.-C. Liang, and V. Scarani, *Geometry of the set of quantum correlations*, Phys. Rev. A **97**, 022104 (2018).
[69] N. Gisin, *Bell's inequality holds for all non-product states*, Phys. Lett. A **154**, 201 (1991).
[70] F. Bernards and O. Gühne, *Bell inequalities for nonlocality depth*, Phys. Rev. A **107**, 022412 (2023).
[71] J.-D. Bancal, C. Branciard, N. Brunner, N. Gisin, and Y.-C. Liang, *A framework for the study of symmetric full-correlation Bell-like inequalities*, J. Phys. A **45**, 125301 (2012).
[72] T. Fritz, *Polyhedral duality in Bell scenarios with two binary observables*, J. Math. Phys. (N.Y.) **53**, 072202 (2012).
[73] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, *Nonlocal correlations as an information-theoretic resource*, Phys. Rev. A **71**, 022101 (2005).
[74] J. T. Barreiro, J.-D. Bancal, P. Schindler, D. Nigg, M. Hennrich, T. Monz, N. Gisin, and R. Blatt, *Demonstration of genuine multipartite entanglement with device-independent witnesses*, Nat. Phys. **9**, 559 (2013).
[75] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, *Noisy intermediate-scale quantum algorithms*, Rev. Mod. Phys. **94**, 015004 (2022).
[76] A. Peruzzo *et al.*, *A variational eigenvalue solver on a photonic quantum processor*, Nat. Commun. **5**, 4213 (2014).
[77] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, *The variational quantum eigensolver: A review of methods and best practices*, Phys. Rep. **986**, 1 (2022).
[78] A. Drinko, G. I. Correr, I. Medina, P. C. Azado, A. Canabarro, and D. O. Soares-Pinto, *Benchmarking variational quantum eigensolvers for entanglement detection in many-body Hamiltonian ground states*, arXiv:2407.04453.
[79] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, *Evaluating analytic gradients on quantum hardware*, Phys. Rev. A **99**, 032331 (2019).
[80] D. Wierichs, J. Izaac, C. Wang, and C. Y.-Y. Lin, *General parameter-shift rules for quantum gradients*, Quantum **6**, 677 (2022).
[81] J. J. Meyer, J. Borregaard, and J. Eisert, *A variational toolbox for quantum multi-parameter estimation*, npj Quantum Inf. **7**, 1 (2021).
[82] O. Kyriienko and V. E. Elfving, *Generalized quantum circuit differentiation rules*, Phys. Rev. A **104**, 052417 (2021).
[83] L. Banchi and G. E. Crooks, *Measuring analytic gradients of general quantum evolution with the stochastic parameter shift rule*, Quantum **5**, 386 (2021).
[84] A. Mari, T. R. Bromley, and N. Killoran, *Estimating the gradient and higher-order derivatives on quantum hardware*, Phys. Rev. A **103**, 012405 (2021).
[85] J. Stokes, J. Izaac, N. Killoran, and G. Carleo, *Quantum natural gradient*, Quantum **4**, 269 (2020).
[86] Z. He, M. Deng, S. Zheng, L. Li, and H. Situ, *Training-free quantum architecture search*, Proc. AAAI Conf. Artif. Intell. **38**, 12430 (2024).
[87] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, arXiv:1412.6980.
[88] W. Ren *et al.*, *Experimental quantum adversarial learning with programmable superconducting qubits*, Nat. Comput. Sci. **2**, 711 (2022).
[89] X. Pan *et al.*, *Deep quantum neural networks on a superconducting processor*, Nat. Commun. **14**, 4006 (2023).
[90] W. Li, Z. Lu, and D.-L. Deng, *Quantum neural network classifiers: A tutorial*, SciPost Phys. Lect. Notes **61** (2022).
[91] L.-W. Yu, W. Li, Q. Ye, Z. Lu, Z. Han, and D.-L. Deng, *Expressibility-induced concentration of quantum neural tangent kernels*, Rep. Prog. Phys. **87**, 110501 (2024).
[92] C. Song, K. Xu, W. Liu, C.-P. Yang, S.-B. Zheng, H. Deng, Q. Xie, K. Huang, Q. Guo, L. Zhang, P. Zhang, D. Xu, D. Zheng, X. Zhu, H. Wang, Y. A. Chen, C. Y. Lu, S. Han, and J. W. Pan, *10-qubit entanglement and parallel logic operations with a superconducting circuit*, Phys. Rev. Lett. **119**, 180511 (2017).
[93] S. S. Elder, C. S. Wang, P. Reinhold *et al.*, *High-fidelity measurement of qubits encoded in multilevel superconducting circuits*, Phys. Rev. X **10**, 011001 (2020).
[94] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop *et al.*, *Demonstration of quantum volume 64 on a superconducting quantum computing system*, Quantum Sci. Technol. **6**, 025020 (2021).
[95] S. Xu, Z.-Z. Sun, K. Wang, H. Li, Z. Zhu, H. Dong, J. Deng, X. Zhang, J. Chen, Y. Wu *et al.*, *Non-Abelian braiding of Fibonacci anyons with a superconducting processor*, Nat. Phys. **20**, 1469 (2024).
[96] M. Ansmann, H. Wang, R. C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, and J. M. Martinis, *Violation of Bell's inequality in Josephson phase qubits*, Nature (London) **461**, 504 (2009).