Testing Scalable Bell Inequalities for Quantum Graph States on IBM Quantum Devices

Bo Yang, Rudy Raymond, Hiroshi Imai, *Member, IEEE*, Hyungseok Chang, and Hidefumi Hiraishi

***Abstract*— Testing and verifying imperfect multi-qubit quantum devices are important as such noisy quantum devices are widely available today. Bell inequalities are known to be useful for testing and verifying the quality of the quantum devices from their nonlocal quantum states and local measurements. There have been many experiments demonstrating the violations of Bell inequalities, but they are limited in the number of qubits and the types of quantum states. We report violations of Bell inequalities on IBM Quantum devices based on the scalable and robust inequalities maximally violated by graph states as proposed by Baccari *et al.*. The violations are obtained from the quantum states of path graphs up to 57 and 21 qubits on a 65-qubit and two 27-qubit IBM Quantum devices, respectively, and from those of star graphs up to 11 qubits with quantum readout error mitigation (QREM). We are able to show violations of the inequalities on various graph states by constructing low-depth quantum circuits and by applying the QREM technique. We also point out that quantum circuits for star graph states of size N can be realized with circuits of depth $O(\sqrt{N})$ on subdivided honeycomb lattices which are the topology of the 65-qubit IBM Quantum device. Our experiments show encouraging results on the ability of existing quantum devices to prepare entangled quantum states and provide experimental evidence on the benefit of scalable Bell inequalities for testing them.**

***Index Terms*— Quantum computing, IBM quantum, benchmarking, graph state, bell inequality.**

I. Introduction

NONLOCALITY of quantum states–first discovered by Bell [2]–is an intriguing consequence of quantum mechanics in which correlations among quantum bits cannot be explained by classical statistics. In particular, the nonlocality implies the so-called Bell inequalities that are violated by entangled (or, nonlocal) quantum states but not by any classical (or, local) correlation. To demonstrate the violation of Bell inequalities, various concepts and experimental tools were developed [3]. One of them is the CHSH inequality [4], which can be used to test the nonlocality of two qubits. The CHSH inequality is also extended by many works, such as that by Ito *et al.* [5] whose inequality allows a wider range of quantum states to violate the classical bounds, or the CHSH-like Bell inequalities for three quantum bits or more, such as the Mermin's inequality for Greenberger–Horne–Zeilinger (GHZ) state [6] and the Bell inequality for graph states [7].

The Bell inequalities soon find their applications for witnessing entanglement [8], [9] and for self-testing [10], [11] quantum devices. The latter is useful for certifying the *quantumness* of the devices by statistical tests on the correlations resulting from the quantum states they produce without the knowledge of their internal functioning. However, most of the existing Bell inequalities require measuring correlations on quantum graph states whose number scales exponentially [7], [12] or polynomially [13], [14], [15] to the number of qubits.

Recently, Baccari *et al.* [1] proposed a family of CHSH-like Bell inequalities that are both scalable and robust. The scalability comes from the fact that the new inequalities can be tested by measuring correlations on quantum graph states whose number scales only linearly with the number of qubits. The robustness stems from the obvious gap between the maximally possible quantum correlation and the classical correlation in their inequalities. They show the fidelity of graph states can be lower bounded by the linear function of the violation magnitude of their inequalities. The scalability and robustness of the new CHSH-like Bell inequalities are therefore potentially beneficial for self-testing noisy quantum devices available today.

At the same time, we have been witnessing the proliferation of near-term quantum devices [16], [17], [18], [19], [20]. Some quantum devices are even widely accessible in the cloud. Here we focus on the IBM Quantum superconducting devices. Currently, IBM Quantum Experience makes over 25 quantum systems online among which 7 quantum devices are open for public use. Since the start of the service in 2016, IBM Quantum processors are getting larger and more sophisticated. The 54-qubit device (`ibmq_rochester`) appeared in 2019, the 65-qubit device (`ibmq_manhattan`) in 2020, and the 127-qubit device (`ibm_washington`) in 2022.

In addition, IBM Quantum Experience adopts the quantum volume (QV) [21] as the metric of capabilities and error rate of the processors. This metric, in a nutshell, is decided by the system size and the max depth of the effective quantum circuit for the randomized model circuits on the device [21], [22].

Manuscript received 13 June 2022; revised 8 August 2022; accepted 17 August 2022. Date of publication 25 August 2022; date of current version 10 October 2022. This article was recommended by Guest Editor J.-H. Jiang. *(Corresponding author: Bo Yang.)*

Bo Yang and Hiroshi Imai are with the Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan (e-mail: yh198595@gmail.com).

Rudy Raymond is with IBM Research Tokyo, Tokyo 103-8510, Japan.

Hyungseok Chang is with the Graduate School of Education, The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan.

Hidefumi Hiraishi is with the College of Science and Technology, Nihon University, Chiyoda, Tokyo 102-0074, Japan.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JETCAS.2022.3201730.

Digital Object Identifier 10.1109/JETCAS.2022.3201730

TABLE I

BASIC PROPERTIES OF IBM QUANTUM DEVICES USED IN OUR EXPERIMENTS. THESE VALUES REFLECT THE DEVICE STATUSES IN JANUARY 2022

name	processor	# qubits	QV	CLOPS
ibm_washington	Eagle r1	127	64	850
ibmq_brooklyn	Hummingbird r2	65	32	1500
ibmq_montreal	Falcon r4	27	128	2000
ibmq_mumbai	Falcon r5.1	27	128	1800
ibm_cairo	Falcon r5.11	27	64	2400
ibm_hanoi	Falcon r5.11	27	64	2300
ibmq_toronto	Falcon r4	27	32	1800
ibm_kawasaki	Falcon r5.11	27	32	-

The higher QV means the better capacity of the device. The following Table I lines up IBM Quantum devices over 20-qubit available from The University of Tokyo, where the CLOPS denotes the circuit layer operation per second, representing how many layers of a QV circuit a quantum processing unit (QPU) can execute per unit of time.

Those devices offer testbeds for investigating the quantum states they produce, i.e., to see if such topologically-limited noisy devices can entangle more qubits and in which way. For example, Wei *et al.* [23] experimentally demonstrated the ability to produce GHZ states up to 18 qubits on a 20-qubit IBM Quantum device measured by their proposed scalable entanglement metric. González *et al.* [24] and Huang *et al.* [25] used Mermin-type Bell inequalities to confirm the entanglement of GHZ states up to five qubits. Nevertheless, the previous experiments are limited, and difficult to verify different types of entangled quantum states that may depend on the restriction of underlying quantum devices and the scalability of the Bell-inequality metrics they chose.

In this work, we address the task of testing noisy quantum devices with various quantum graph states based on the family of CHSH-like Bell inequalities of Baccari *et al.* Graph states have many applications in quantum information processing, and some of them are compatible with the qubit layout topology of current IBM Quantum devices, whose qubits are sparsely connected. Since the inequalities of Baccari *et al.* are designed to be maximally violated by graph states, we used their inequalities to check how well the graph states are prepared on IBM Quantum devices. In particular, we construct path graphs, star graphs, and the qubit-connection graph corresponding to the physical qubit connectivity on the QV32, 65-qubit `ibmq_brooklyn` device, the QV32, 27-qubit `ibm_cairo`, `ibm_hanoi`, `ibm_kawasaki`, `ibmq_mumbai`, `ibmq_toronto` devices, and the QV128, 27-qubit `ibmq_montreal` device. Most of these devices appeared in the period from 2020 to 2021.

Through the experiments, we observe the violation of full 65-qubit connection graph states. We also observe the maximum violation of path graphs up to 57-qubit and the maximum violation of star graphs up to 11-qubit with the efficient quantum readout error mitigation (QREM) by Yang *et al.* [26]. These experimental results also support both the ability of IBM Quantum devices to create the well-entangled large quantum states and the applicability of the scalable CHSH-like inequalities by Baccari *et al.* to the entanglement witness and device benchmarking of the growing near-term quantum devices.

These violations are made possible by shallow-depth circuits for the corresponding graph states. Namely, path graphs are from quantum circuits with depth two as used in [27], and star graphs on five-qubit or larger are from quantum circuits avoiding SWAP gates following a similar construction shown in [23]. We also provide a generalization of quantum circuit design of N-qubit star graph states within depth $O(\sqrt{N})$ on the subdivided honeycomb lattice structure, which is the typical topology of current IBM Quantum devices.

The rest of the paper is organized as follows. Section II provides a brief introduction to graph states, the Bell inequalities of Baccari *et al.*, and the quantum readout error mitigation used in our experiments. Section III explains how to make shallow quantum circuits for graph states on the IBM Quantum devices whose qubits are sparsely connected. Section IV shows the experimental results on IBM Quantum devices showing their ability to entangle more qubits than reported before. Section V concludes with a discussion of the results and future works.

II. PRELIMINARIES

A. Graph State

First, we consider a simple undirected graph $G = (V, E)$, where $V = \{0, \ldots, N-1\}$ is a set of vertices and $E = \{\{u, v\} | u, v \in V, u \neq v\}$ is a set of edges. Let $n(v)$ denote the set of all neighbours of v, that is, the vertices that are connected to v; let also $n[v] := n(v) \cup \{v\}$ be the vertex set containing v and its neighbours.

Then the graph state $|\psi_G\rangle$ is generated from a product state $|+\rangle^{\otimes |V|}$ by applying the controlled-Z gates $CZ(i, j)$ between the qubits i and j corresponding to the edge $(i, j) \in E$.

$$|\psi_G\rangle := \prod_{(i,j)\in E} CZ(i, j)|+\rangle^{\otimes N}. \tag{1}$$

This state can also be characterized by stabilizer formalism. Let $\sigma_X^{(j)}$ be the Pauli X operator which acts on the qubit j. Likewise, let $\sigma_Z^{(j)}$ be the Pauli Z operator which acts on the qubit j. Then the set of stabilizer generators of graph state $|\psi_G\rangle$ is $\{G_v\}_{v\in V}$ where

$$G_v = \sigma_X^{(v)} \otimes \bigotimes_{i\in n(v)} \sigma_Z^{(i)}. \tag{2}$$

This means $|\psi_G\rangle$ is the simultaneous eigenstate of all of the stabilizer generators in $\{G_v\}_{v\in V}$.

For example, the pictures of the five-qubit star graph state and the five-qubit path graph state are shown in Fig. 1.

B. The Scalable Bell Inequalities of Baccari et al.

The Bell inequalities of Baccari *et al.* is designed as a family of robust and scalable Bell inequalities for self-testing quantum devices. This can be considered a generalization of CHSH inequality [4], which is suitable for demonstrating the nonlocality between two qubits. While the previous Bell inequalities for self-testing would scale exponentially in terms of the size of required resources, the inequalities by

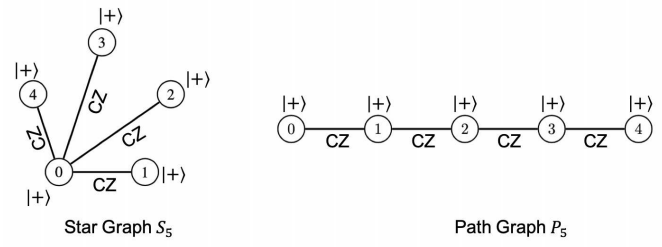


Fig. 1. Graph states associated with a star graph S_5 (left) and path graph P_5 (right) respectively. Each vertex corresponds to each qubit with $|+\rangle$ state and each edge corresponds to applying the controlled-Z (CZ) gate between the associated qubits.

Baccari *et al.* scale linearly to the system size. In this work, we used their inequalities to conduct the benchmarking of large quantum devices available through cloud access and measured the nonlocality of quantum states on them.

We recall the original notations of the scalable Bell inequalities of Bacarri *et al.* here. Selecting a vertex subset (i.e. a set of qubits) $F \subseteq V$ satisfying ${}^{\forall}u, v \in F,\ n[u] \cap n[v] = \emptyset$, the general form of their proposed physical correlation among N particles which form a graph structure G becomes

$$I_G(F) = \sum_{v\in F}\left(\deg(v)\left\langle\left(A_0^{(v)} + A_1^{(v)}\right)\prod_{i\in n(1)} A_1^{(i)}\right\rangle + \sum_{i\in n(v)}\left\langle\left(A_0^{(i)} - A_1^{(i)}\right)\prod_{j\in n(1)} A_1^{(j)}\right\rangle\right) + \sum_{i\notin\bigcup_{v\in F} n[v]}\left\langle A_0^{(i)}\prod_{j\in n(1)} A_1^{(j)}\right\rangle \quad (3)$$

where $A_b^{(i)}$ represents the observable labeled with $b \in \{0, 1\}$ on the particle (or qubit) i. Their inequalities show that this physical correlation $I_G(F)$ is upper bounded by $\beta_G^C(F)$ and $\beta_G^Q(F)$ shown below, for the classical parties and the quantum parties, respectively.

$$\beta_G^C(F) = \sum_{v\in F}\deg(v) + N - |F|$$
$$\beta_G^Q(F) = (2\sqrt{2}-1)\sum_{v\in F}\deg(v) + N - |F| \quad (4)$$

We can see $\beta_G^C(F)$ and $\beta_G^Q(F)$ will become distinct for general graph size N, if the sum of the degrees of vertices in F increases in proportion to N.

Baccari *et al.* also showed that $I_G(F)$ reaches the maximum quantum violation when each operator $A_b^{(i)}$ takes $A_0^{(i)} = \frac{1}{2}\left(\sigma_X^{(i)} + \sigma_Z^{(i)}\right), A_1^{(i)} = \frac{1}{2}\left(\sigma_X^{(i)} - \sigma_Z^{(i)}\right)$ if $i \in F$, $A_0^{(i)} = \sigma_X^{(i)}, A_1^{(i)} = \sigma_Z^{(i)}$ if $i \notin F$. In this setting, the total correlation $I_G(F)$ becomes

$$I_G(F) = \sum_{v\in F}\left(\deg(v)\langle G_v\rangle + \sum_{i\in n(v)}\langle G_i\rangle\right) + \sum_{i\notin\bigcup_{v\in F} n[v]}\langle G_i\rangle. \quad (5)$$

where every stabilizer operator G_i is used once.

Besides, choosing a feasible vertex subset F would maximize the gap between β_G^C and β_G^Q, which makes the violation on real devices clearer. The optimal choice of F for path graphs is to choose every two vertices, and the optimal choice for star graphs is to choose the central node. Or for N-node star graph S_N, $I_G(F)$ becomes quite simple, as described in the following form (6).

$$I_{S_N} = \sqrt{2}(N-1)\left\langle\sigma_X^{(1)}\sigma_Z^{(2)}\cdots\sigma_Z^{(N)}\right\rangle + \sqrt{2}\sum_{i\in V\setminus\{1\}}\left\langle\sigma_Z^{(1)}\sigma_X^{(i)}\right\rangle \quad (6)$$

The measurement operation of IBM Quantum devices is performed in the Pauli Z basis (the computational basis). To measure a qubit in the Pauli X basis, we can apply the Hadamard gate on it and measure the qubit in the Pauli Z basis.

C. Quantum Readout Error Mitigation (QREM)

Since the noise level is crucial and quantum error correction (QEC) is not yet available on the current quantum computers, we have to reduce the noise effects in classical postprocessing to achieve better performance. Quantum readout error mitigation (QREM) is one of the most practical quantum error mitigation techniques, targeted to reduce the effect of readout noise from the measurement results. The workflow of QREM is to calibrate the readout noise by running each computational basis state for the first step and then reconstructing the calibration matrix as a noise channel. For the next step, the mitigated probability distribution is obtained by applying the inverse of the calibration matrix to the noisy measurement result.

In experiments, we used the efficient QREM algorithm by Yang *et al.* [26]. This algorithm scales $O(ns^2)$ time for the measurement result with n-qubit and s shots thanks to its assumption that the unmeasured states are ignorable for sparse measurement results with few measured states. Applying QREM would significantly improve the measurement results as shown in [26], where they demonstrated the fidelity of GHZ states by mitigated results increased over 1.5 times higher than the fidelity by raw results.

III. Quantum Circuit Construction

In this experiment, we used IBM Quantum 65-qubit device (`ibmq_brooklyn`) and 27-qubit devices (`ibm_cairo`, `ibm_hanoi`, `ibm_kawasaki`, `ibmq_montreal`, `ibmq_mumbai` and `ibmq_toronto`). On these devices, we investigated the correlations of path graph P_N, star graph S_N, and the connection graphs of each device.

All the devices have the heavy-hexagonal qubit connection structure. We used the circuit designing techniques used by Wei *et al.* [23] and Mooney *et al.* [27] to prepare shallower quantum circuits.

A. Preparing the Path Graph States

Path graph state $|\psi_{P_N}\rangle$ can be prepared by shallow circuits with depth two as shown in [27]. This is completed by preparing $|+\rangle^{\otimes n}$ state first, then applying the controlled-Z gate to every other edge of the path, and finally applying the

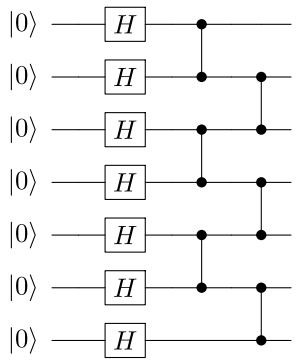


Fig. 2. Quantum circuit for the 7-qubit path graph state with the constant depth construction. After preparing the uniform superposition state by Hadamard gates, controlled-Z gates are applied in a parallel way.

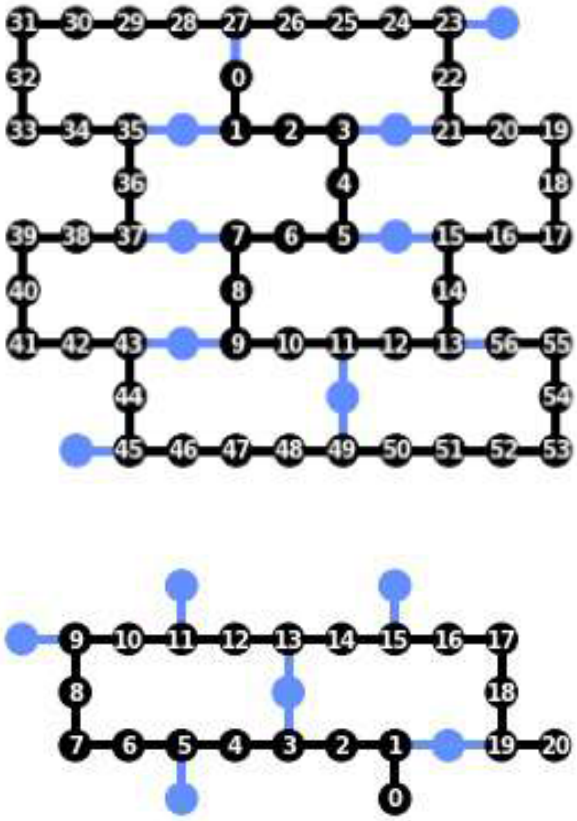


Fig. 3. Qubit layout of the path graph states on each IBM Quantum device. Each figure shows the qubit layout of 65-qubit devices and 27-qubit devices, respectively. The numbers on the figure represent the virtual positions on the physical qubits. The graph size is expanded from the qubits with smaller virtual indices to the qubits with larger virtual indices. These orders and locations of the qubits are chosen to maximize the length of path graph states on each device.

controlled-Z gate to every other remaining edge. The quantum circuit creating the graph state of the path graph P_N is shown in Fig. 2.

To make the path graph with the constant depth circuit on each device, we used the qubit layout in Fig. 3. The violation for path graph states is tested from the size two up to the maximum size that can be taken on each device.

B. Preparing the Star Graph States

Since star graphs are equivalent to GHZ states in terms of local Clifford operations [28], the quantum state corresponding to the star graph can be obtained from the GHZ state by applying the local Hadamard gate to every qubit except for the qubit representing the central node in the graph. That is, assuming that the central vertex is labeled by 0, the following equation holds.

$$|\psi_{S_N}\rangle = \left(I \otimes H^{\otimes(N-1)}\right)|\psi_{GHZ_N}\rangle \tag{7}$$

Then the remaining task is to prepare the GHZ state in a shallower manner, for which we can use the technique as shown in [23]. The main idea of this technique is that GHZ states can be prepared without qubit swapping operations on any tree-structured physical connection of qubits. The GHZ states are realized by applying the Hadamard gate to an initial qubit and then applying the X gate to other qubits controlled by the initial qubit. Since all qubits in the GHZ state are equivalently entangled, we can apply the controlled-X gates to different pairs of qubits in parallel by properly changing the control qubits. By doing so, it is possible to realize a shallower circuit with depth $O(\sqrt{N})$ for a star graph state with size N on the topology of IBM Quantum 65-qubit devices. The proof is given in the following discussion.

We first describe the construction of tree graph state $|\psi_{T_N}\rangle$ with tree depth d and see what the physical qubit topology should be taken. We then show such a graph can be embedded into the topology of the subdivided honeycomb structure. In order to create a quantum circuit, we start from vertex 0. If vertex 0 is connected with another vertex, say vertex 1, we can add it to the tree, making $|\psi_{T_2}\rangle$ with depth one.

Next, if one of the vertices $0, 1$ has degree three or larger, connected with vertex 2, and the other vertex has degree two or larger, connected with vertex 3, then we can simultaneously add vertices $2, 3$ to vertex $0, 1$. This time, the created tree $|\psi_{T_4}\rangle$ has the depth two, with three outer vertices on the qubit topology connected to different vertices of $|\psi_{T_4}\rangle$. Going one step further, if two of three neighborhoods of $|\psi_{T_4}\rangle$ have degree two or larger, and the remaining neighborhood has degree three or larger, then we can make $|\psi_{T_7}\rangle$ in one step, and assure four additional neighborhoods for $|\psi_{T_7}\rangle$.

In this way, the size of tree graph (or tree graph state) we can prepare with depth d is at least $N = \frac{1}{2}\, d(d+1) + 1$. The condition that the physical qubit topology should satisfy to achieve this construction is that they can add $d-1$ vertices with degree two and at least one vertex with degree three. Such structure can be easily found in subdivided honeycomb because every vertex with degree two in the subdivided honeycomb is adjacent to vertices with degree three, and vice versa. Note that the argument above can be applied to other two-dimensional lattice structures.

In our experiments, we prepared the star graph $|\psi_{S_N}\rangle$ of size $N = 2, \cdots, 39$ on 65-qubit devices, and of size $N = 2, \cdots, 27$ on 27-qubit devices. The detail of how we prepared star graphs on each device is shown in Fig. 4.

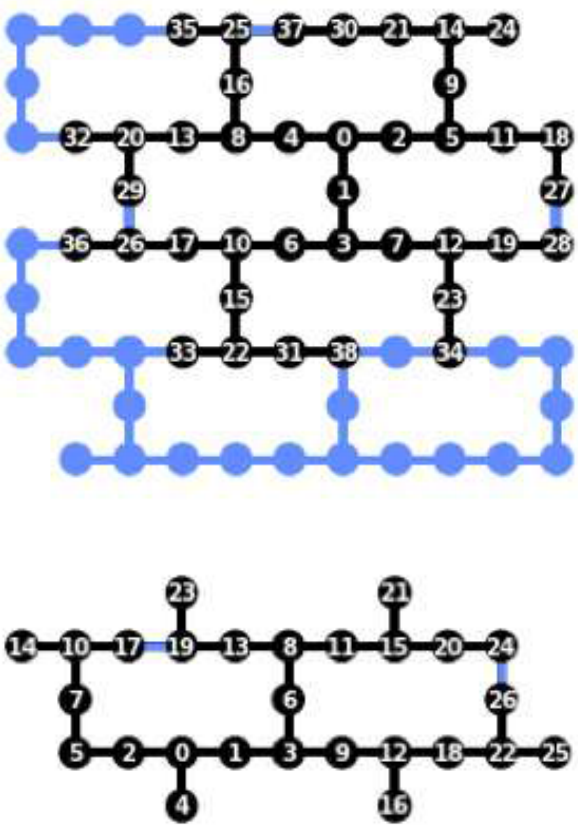


Fig. 4. Qubit layout of the star graph states on each IBM Quantum device. Each figure shows the qubit layout of 65-qubit devices and 27-qubit devices, respectively. The numbers on the figure represent the virtual positions on the physical qubits. The graph size is expanded from the qubits with smaller virtual indices to the qubits with larger virtual indices. The central vertex of the star graph is set to qubit 0.

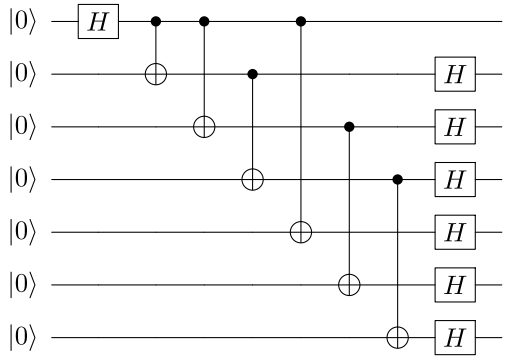


Fig. 5. Quantum circuit for the 7-qubit star graph state with $O(\sqrt{N})$ depth constructed according to the restriction of qubit connection on IBM Quantum devices. The first Hadamard gate followed by controlled-X gates makes the 7-qubit GHZ state and the final parallel Hadamard gates convert the GHZ state to the star graph state. Note that the depth of controlled-X gates is three since some of the controlled-X gates are parallelized.

In general, the qubits with small single and two-qubit operational errors are selected. The circuit example of the 7-qubit star graph on the devices is shown in Fig. 5.

C. Preparing the Qubit-Connection Graph State for Each Device

The graph structure of the qubit connection of each quantum device we used can be seen as a subdivision of the honeycomb graph. Let us define this graph of size N as TH_N. Since the maximum degree of TH_N is three, it is shown to be three-edge colorable by Vizing's theorem [29]. Therefore, we can prepare the quantum circuit for TH_N in circuit depth three. The specific construction of TH_N corresponding to the graph structure of each device is explained in Fig. 6.

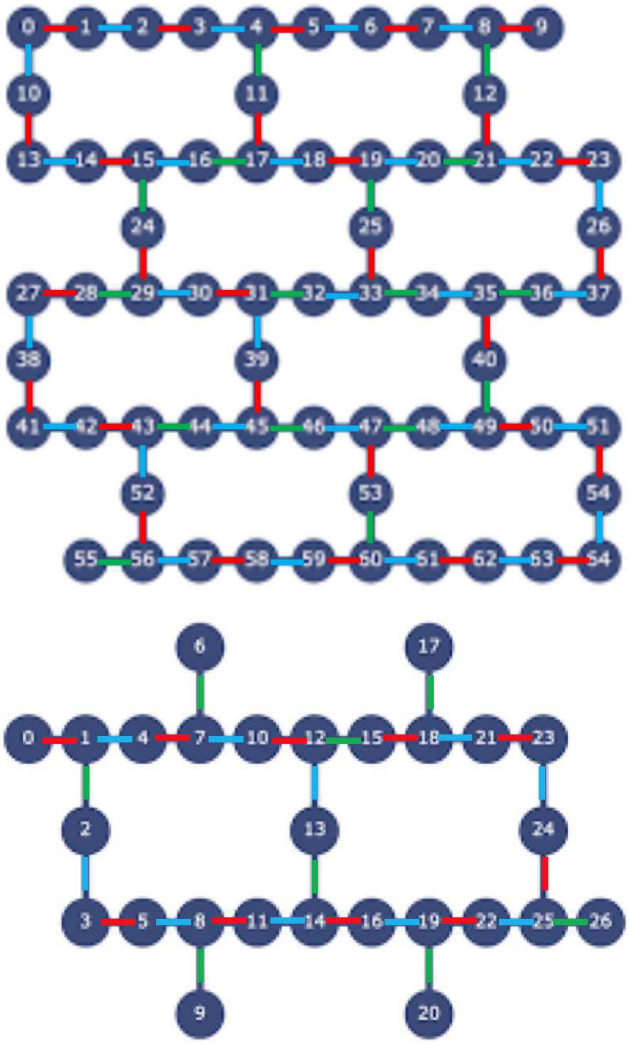


Fig. 6. Qubit layout of the connection graph states on each IBM Quantum device. Each figure shows the qubit layout of 65-qubit devices and 27-qubit devices, respectively. The numbers on the figure represent the positions of virtual qubits. In order to prepare the subdivided honeycomb graph with the whole qubits on the device, we first apply the controlled-Z gate on the red edges, then on the blue edges, and finally on the green edges. The "focused" qubits in the inequalities of Baccari *et al.* are set to [1,6,8,12,17,19,23,26] for 27-qubit devices, and [3,6,9,10,17,21,24,25,26,31,35,38,44,47,54,56,59,62] for the 65-qubit device.

D. Measurement Grouping

In the experiments, we also adopted the technique called *measurement grouping*, which has been conventionally used [30], [31], [32], [33], [34], [35], [36], [37]. With this method, we can efficiently utilize the current limited quantum computational resources by reducing the number of

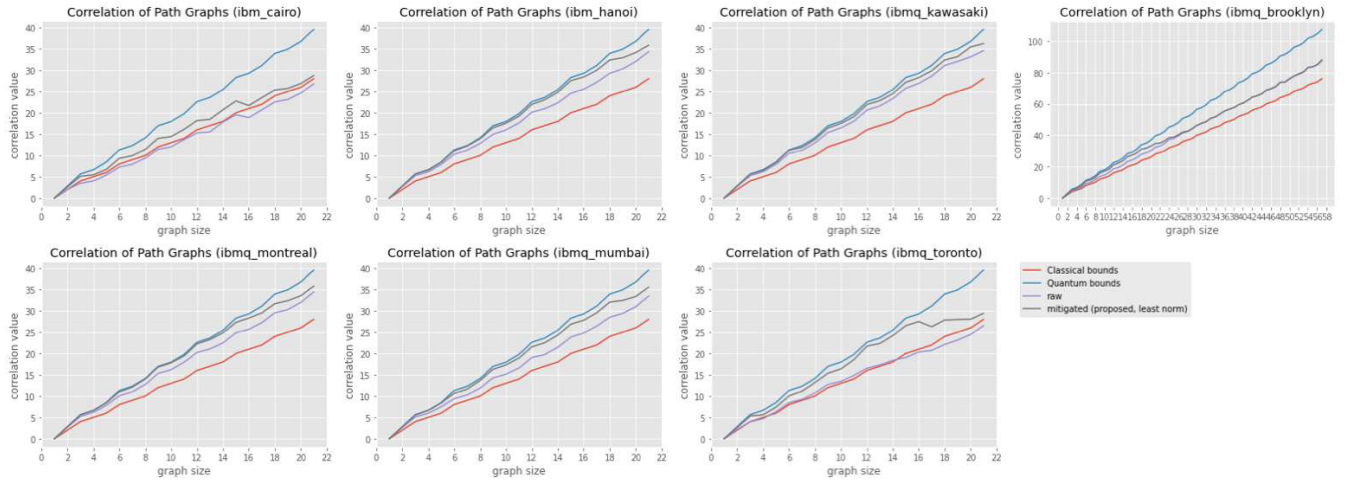


Fig. 7. Correlations of path graphs on each quantum devices. The red lines denote the upper bounds of classical correlation of the inequalities of Baccari *et al.*, and the blue lines denote the upper bounds of quantum correlation of the inequalities. The purple and black lines represent the measured correlation on real devices with and without QREM, respectively.

observables that otherwise should be measured separately. The idea of the measurement grouping is to create a grouped observable from the observables sharing the same measurement basis. For example, the observable $\sigma_X^{(0)}\sigma_Z^{(1)}$ and $\sigma_Z^{(1)}\sigma_X^{(2)}$ can be grouped into $\sigma_X^{(0)}\sigma_Z^{(1)}\sigma_X^{(2)}$. After measuring the grouped observable, we can recover the probability distribution of each observable by just taking the marginal distribution on the qubits that the observable covers.

The minimum number of grouped observable in the inequalities of Baccari *et al.* for a graph state $|\psi_G\rangle$ is characterized by the chromatic number χ_G of its associated graph G. Since the observables in the inequalities of Baccari *et al.* correspond to the stabilizers of the graph state, the vertices with the measurement basis σ_X in the grouped observable form the independent set of G. This concludes that the χ_G different measurement observables are enough instead of measuring all N different observables in the inequalities.

Since the path graph and the star graph are all 2-colorable, the stabilizers of those graph states can be grouped into two observables. For the path graph, the grouped observables become $\left\langle \sigma_X^{(0)}\sigma_Z^{(1)}\sigma_X^{(2)}\sigma_Z^{(3)}\cdots \right\rangle$ and $\left\langle \sigma_Z^{(0)}\sigma_X^{(1)}\sigma_Z^{(2)}\sigma_X^{(3)}\cdots \right\rangle$. For the star graph whose center is qubit 0, the grouped observables become $\left\langle \sigma_X^{(0)}\sigma_Z^{(1)}\sigma_Z^{(2)}\sigma_Z^{(3)}\cdots \right\rangle$ and $\left\langle \sigma_Z^{(0)}\sigma_X^{(1)}\sigma_X^{(2)}\sigma_X^{(3)}\cdots \right\rangle$. Since the connection graph of the heavy-hexagonal structure on the IBM Quantum devices is three-colorable, the number of observables in the measurement process can also be reduced from N to three.

IV. Results of Experiments

The python codes of this experiment are available at [38]. The result of each experiment is averaged over the constant 8192 shots. This can be justified by Hoeffding's inequality. The unbiased estimator $\langle O\rangle$ of $\mathrm{Tr}(O\rho)$ for the observable O on the noisy state ρ can be bounded by $\Pr[|\langle O\rangle - \mathrm{Tr}(O\rho)| \geq \delta] \leq \exp(-s\delta^2/2)$ with the shot count s [39]. This means the probability that the expectation value greatly deviates from the true expectation value will be exponentially small to the error rate and the shot count. If we set $s = 8192$, even the probability that the gap between $\langle O\rangle$ and $\mathrm{Tr}(O\rho)$ is larger than $\delta = 0.05$ is at most 3.6×10^{-5}.

The correlations for path graphs obtained from the experiments on each device is shown in Fig. 7. The correlations for star graphs on each device are shown in Fig. 8. The purple lines are the raw correlations without QREM, and the black lines are the correlations with the proposed QREM method by [26] using its least norm formulation. Fig. 9 shows the term-wise mitigated correlations in (6) for each size of star graphs. For example, the curve labeled by 0 indicates the expectation value of $\left\langle \sigma_X^{(0)}\sigma_Z^{(1)}\cdots\sigma_Z^{(N-1)} \right\rangle$ in (6) for graph size $N = 2, 3, 4, \cdots$, and the curve labeled by "3" indicates the expectation value of $\left\langle \sigma_X^{(0)}\sigma_Z^{(3)} \right\rangle$ in (6) for graph size $N = 4, 5, 6, \cdots$.

From Fig. 7, we observe the path graphs violate the inequality with a clear gap from classical bounds without QREM on most of the devices, while the plots of `ibm_cairo` and `ibmq_toronto` show the worse correlations that do not violate the classical upper bounds. However, when the QREM method is added, the correlation of path graph state for every size and on every device violated the classical bounds. With QREM, the measured correlations on the devices except for `ibm_cairo` almost reach the theoretical upper bound of quantum correlation. Besides, thanks to the constant-depth circuit preparation of path graph states, the plots of path graphs on each device seem to grow stably to the system size between the classical bound and quantum bound. Therefore, we may say the path graph state can be well prepared on the current IBM Quantum devices.

As for the star graph states, the raw correlations without QREM only violate the classical upper bounds up to four to six qubits on each device. With the proposed QREM, the maximum size of violation increased to size 11 on `ibm_hanoi` and `ibmq_toronto`, to size 10 on `ibm_cairo` and

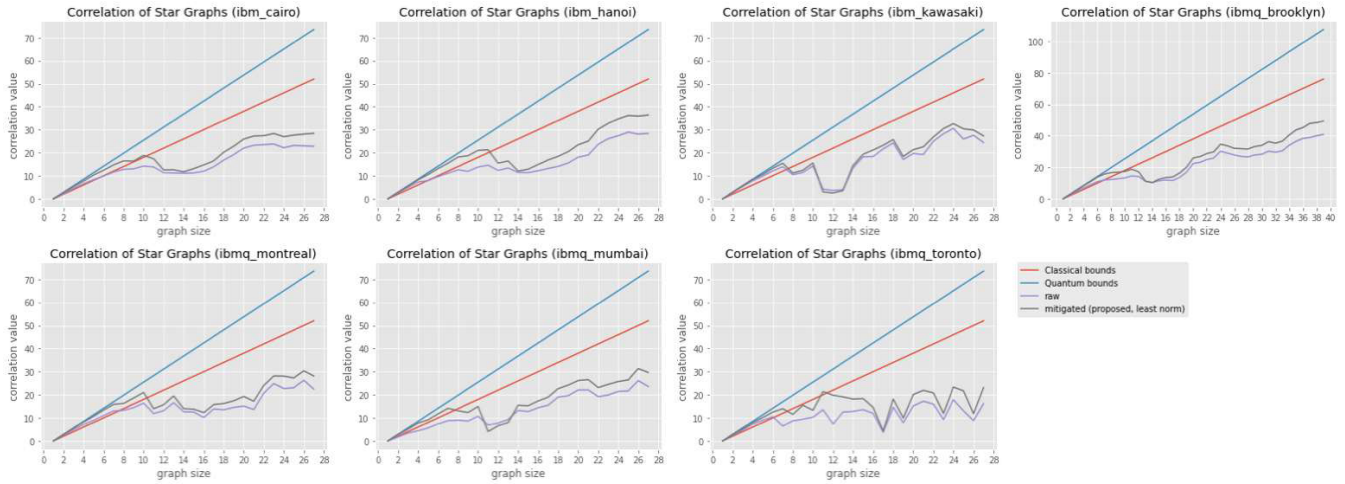


Fig. 8. Correlations of star graphs on each quantum devices. The red lines denote the upper bounds of classical correlation of the inequalities of Baccari *et al.*, and the blue lines denote the upper bounds of quantum correlation of the inequalities. The purple and black lines represent the measured correlation on real devices with and without QREM, respectively.

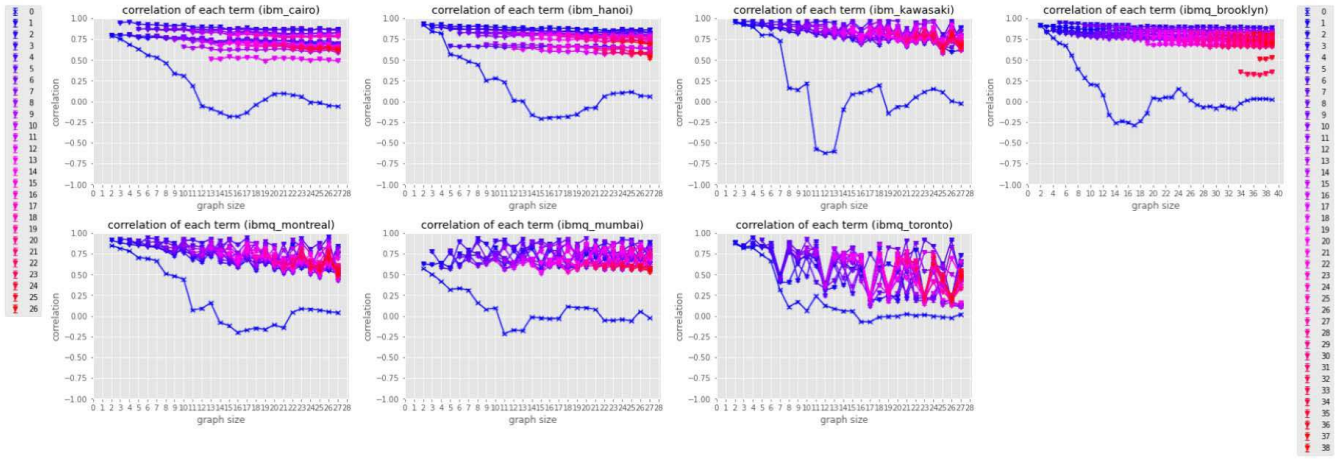


Fig. 9. Term-wise correlations of star graphs on each quantum device without QREM. The labels on the left edge represent the term-wise correlation labels for 27-qubit devices. The labels on the right edge represent the term-wise correlation labels for 65-qubit devices (ibmq_brooklyn only). The numbers of the labels are the indices of stabilizer operator associated with the virtual positions of qubits. For example, the correlation of the first term $\left\langle \sigma_X^{(0)} \sigma_Z^{(2)} \cdots \sigma_Z^{(N-1)} \right\rangle$ in (6) is represented by the label "0", and the correlation of the fifth term $\left\langle \sigma_Z^{(0)} \sigma_X^{(4)} \right\rangle$ in (6) is represented by the label "4".

`ibmq_montreal`, to size nine on `ibmq_brooklyn`, and to size seven on `ibm_kawasaki` and `ibm_mumbai`.

These small violation sizes compared to the violation sizes of path graph states can be explained by the circuit depth. Unlike the constant depth circuit for path graph states, the circuit depth of the star graph state scales $O(\sqrt{N})$ to the graph size N. A deeper circuit depth would be more vulnerable to decoherence which is one of the dominant noises in the current near-term quantum devices.

We founded out the biggest cause that prevents the star graph states from getting higher correlations is the significantly low correlation of the observable $\langle \sigma_X^{(0)} \sigma_Z^{(1)} \cdots \sigma_Z^{(N-1)} \rangle$ for the central vertex (qubit 0). The decrease of this correlation is clear between the graph sizes 10 and 11, as shown in Fig. 9, and it even scored negative values which are particularly obvious on the device of `ibm_kawasaki`.

The difference of correlation between the central vertex and the other vertices might be explained by the difference between the two grouped measurement observables $\left\langle \sigma_X^{(0)} \sigma_Z^{(1)} \sigma_Z^{(2)} \sigma_Z^{(3)} \cdots \right\rangle$ and $\left\langle \sigma_Z^{(0)} \sigma_X^{(1)} \sigma_X^{(2)} \sigma_X^{(3)} \cdots \right\rangle$. The first observable corresponds to the original observable for the central vertex. The second observable is the grouped observable of all the observables for the leaf vertices. We will discuss below how the measurement results of these two observables contain the noise effects.

The measurement operation of the second observable $\left\langle \sigma_Z^{(0)} \sigma_X^{(1)} \sigma_X^{(2)} \sigma_X^{(3)} \cdots \right\rangle$ for the star graph state is performed by first applying the Hadamard gates on the qubits except for qubit 0, and then measuring all the qubits in the computational basis. This operation corresponds to measuring the GHZ state in the computational basis. Since the GHZ state ideally has only the all-zero state and the all-one state, the noisy

TABLE II

CORRELATION VALUES OF THE WHOLE-QUBIT HONEYCOMB GRAPH ON EACH DEVICE. THE β_G^C AND β_G^Q REPRESENT THE THEORETICAL UPPER BOUNDS OF CLASSICAL CORRELATION AND QUANTUM CORRELATION RESPECTIVELY. THE "RAW $\overline{\beta}_G^Q$" AND "MITIGATED $\overline{\beta}_G^Q$" DENOTE THE MEASURED QUANTUM CORRELATIONS WITHOUT AND WITH QREM RESPECTIVELY. HERE THE PROPOSED THE QREM WITH LEAST NORM METHOD IS USED

device	β_G^C	β_G^Q	raw $\overline{\beta}_G^Q$	mitigated $\overline{\beta}_G^Q$
ibmq_brooklyn	88	121.97	94.58 ± 0.120	96.79 ± 0.12
ibm_cairo	36	50.08	41.05 ± 0.075	41.39 ± 0.07
ibm_hanoi	36	50.08	44.21 ± 0.060	44.61 ± 0.06
ibm_kawasaki	36	50.08	44.74 ± 0.060	45.39 ± 0.05
ibmq_montreal	36	50.08	38.54 ± 0.080	38.82 ± 0.08
ibmq_mumbai	36	50.08	39.94 ± 0.080	40.86 ± 0.08
ibmq_toronto	36	50.08	30.52 ± 0.100	32.21 ± 0.10

probability distribution of the GHZ state measured by the computational basis would still hold high probabilities for these two states, which would result in a higher expectation value. Accordingly, the correlations of the leaf vertices in the star graph recovered from the latter observable would reflect the performance of the GHZ state.

On the other hand, after applying the Hadamard gate to the qubit 0 in order to measure the first observable $\left\langle \sigma_X^{(0)} \sigma_Z^{(1)} \sigma_Z^{(2)} \sigma_Z^{(3)} \cdots \right\rangle$ in the computational basis the number of superposition in the star graph state would increase more. This state is equivalent to the quantum state after applying the Hadamard transform to all qubits of the GHZ state. Therefore, those results might exhibit more noise effects of T1/T2 relaxation. In addition, the correlation of the central vertex of the star graph would accumulate all the errors that happened in the N-qubit system, while the correlations of leaf vertices only use the subsystem that their original stabilizer observables are in charge of. This might also be the cause of the difference in the performance between the correlation associated with the central qubit and the correlations associated with the other qubits.

We also report the violation of the CHSH-like Bell inequalities by Baccari *et al.* in the subdivided honeycomb graph using whole qubits on both the 65-qubit device and the 27-qubit devices. The correlations of the full connection graph state associated with the qubit structure of each device are listed in Table II. This result implies that these IBM Quantum devices except for `ibmq_toronto` have the ability to prepare a large graph state unique to its qubit layout using its whole qubits, with rather good accuracy. Especially the 65-qubit connection graph states on `ibmq_brooklyn` achieves the largest graph state preparation on IBM Quantum Experience.

V. CONCLUSION

Through our experiments, we support the benefits of the CHSH-like inequalities proposed by Baccari *et al.* [1] in terms of its scalability and robustness. The inequalities of Baccari *et al.* enabled us to demonstrate the nonlocality of large quantum states up to the 65-qubits graph state. This implies the advantage of using their inequalities over using Mermin's inequality in the experiments by [24], [25], because Mermin's inequality has an exponential number of terms to the graph size.

Using Baccari *et al.*'s remarkable Bell inequalities, we also support the ability of existing IBM Quantum devices to prepare well-entangled large graph states on them. We report in this work the violation of the inequalities for several graph states with a large number of qubits. Using shallow circuits with a constant depth [27], we have seen path graphs violate the inequalities up to the maximum size on each IBM Quantum device. In particular, for the IBM Quantum 65-qubit device, path graphs showed its quantumness up to size 57. We also checked the violation of classical bounds for the graph state corresponding to the graph structure of each quantum device with its whole qubits. Our result reports the violation of star graphs up to size six which can be further improved up to 11 by applying QREM. These are the largest graph states on star graphs at the moment and our experiments can be seen as one of the milestones demonstrating the performance of IBM Quantum devices.

For future work, a possible improvement in circuit preparation can be found in the experiments by Wei *et al.* [23]. During their experiments, they added a collective π-pulse on all qubits in order to refocus low-frequency noise and reduce dephasing errors using the idea of Hahn echo [40]. As they applied the π-pulse between the entangling process of GHZ states and the disentangling process which undo the entangling process, π-pulse becomes most effective for certain time intervals decided by T1/T2 relaxation times. Since our experiments do not have the structure of symmetry in terms of entangle process and disentangle process, partial insertion of π-pulse into the entangled qubits might improve the correlations instead of the direct insertion of π-pulse in the middle of our circuits. This might significantly improve the low correlation of the observable associated with the central qubit in star graph states. Other ideas for decreasing the dephasing errors, such as dynamic decoupling methods discussed in [41], might also help us improve the total correlations of the inequality.

In conclusion, our results for the large quantum states greatly owe to the scalability of the Bell inequalities proposed by Baccari *et al.* and we experimentally support the usefulness of their inequalities as a powerful tool for the entanglement verification of large quantum states and for the benchmarking of upcoming near-term quantum devices.

ACKNOWLEDGMENT

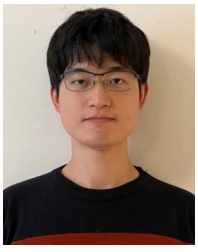
The results presented in this paper were obtained using an IBM Quantum computing system as part of the IBM Quantum Hub at The University of Tokyo.

REFERENCES

[1] F. Baccari, R. Augusiak, I. Šupić, J. Tura, and A. Acín, "Scalable Bell inequalities for qubit graph states and robust self-testing," *Phys. Rev. Lett.*, vol. 124, no. 2, Jan. 2020, Art. no. 020402, doi: 10.1103/PhysRevLett.124.020402.

[2] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Phys. Physique Fizika*, vol. 1, no. 3, pp. 195–200, 1964, doi: 10.1103/PhysicsPhysiqueFizika.1.195.

[3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," *Rev. Mod. Phys.*, vol. 86, pp. 419–478, Apr. 2014, doi: 10.1103/RevModPhys.86.419.
[4] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, no. 15, pp. 880–884, 1969, doi: 10.1103/PhysRevLett.23.880.
[5] T. Ito, H. Imai, and D. Avis, "Bell inequalities stronger than the Clauser-Horne-Shimony-Holt inequality for three-level isotropic states," *Phys. Rev. A, Gen. Phys.*, vol. 73, no. 4, Apr. 2006, Art. no. 042109, doi: 10.1103/PhysRevA.73.042109.
[6] N. D. Mermin, "Extreme quantum entanglement in a superposition of macroscopically distinct states," *Phys. Rev. Lett.*, vol. 65, no. 15, pp. 1838–1840, Oct. 1990, doi: 10.1103/PhysRevLett.65.1838.
[7] O. Gühne, G. Tóth, P. Hyllus, and H. J. Briegel, "Bell inequalities for graph states," *Phys. Rev. Lett.*, vol. 95, no. 12, Sep. 2005, doi: 10.1103/PhysRevLett.95.120405.
[8] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, "Detecting nonlocality in many-body quantum states," *Science*, vol. 344, no. 6189, pp. 1256–1258, Jun. 2014, doi: 10.1126/science.1247715.
[9] J. Tura *et al.*, "Nonlocality in many-body quantum systems detected with two-body correlators," *Ann. Phys.*, vol. 362, pp. 370–423, Nov. 2015, doi: 10.1016/j.aop.2015.07.021.
[10] D. Mayers and A. Yao, "Self testing quantum apparatus," *Quantum Inf. Comput.*, vol. 4, no. 4, pp. 273–286, Jul. 2004.
[11] A. Acín, N. Gisin, and L. Masanes, "From Bell's theorem to secure quantum key distribution," *Phys. Rev. Lett.*, vol. 97, Sep. 2006, Art. no. 120405, doi: 10.1103/PhysRevLett.97.120405.
[12] R. Schmied *et al.*, "Bell correlations in a bose-Einstein condensate," *Science*, vol. 352, no. 6284, pp. 441–444, Apr. 2016, doi: 10.1126/science.aad8665.
[13] J. Tura, A. B. Sainz, T. Vértesi, A. Acín, M. Lewenstein, and R. Augusiak, "Translationally invariant multipartite Bell inequalities involving only two-body correlators," *J. Phys. A, Math. Theor.*, vol. 47, no. 42, Oct. 2014, Art. no. 424024, doi: 10.1088/1751-8113/47/42/424024.
[14] F. Baccari, D. Cavalcanti, P. Wittek, and A. Acín, "Efficient device-independent entanglement detection for multipartite systems," *Phys. Rev. X*, vol. 7, no. 2, Jun. 2017, Art. no. 021042, doi: 10.1103/PhysRevX.7.021042.
[15] 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*, vol. 7, no. 2, Apr. 2017, Art. no. 021005, doi: 10.1103/PhysRevX.7.021005.
[16] P. Jurcevic *et al.*, "Demonstration of quantum volume 64 on a superconducting quantum computing system," *Quantum Sci. Technol.*, vol. 6, no. 2, Apr. 2021, Art. no. 025020.
[17] J. M. Pino *et al.*, "Demonstration of the QCCD trapped-ion quantum computer architecture," 2020, *arXiv:2003.01293*.
[18] F. Arute *et al.*, "Quantum approximate optimization of non-planar graph problems on a planar superconducting processor," *Nature Phys.*, vol. 17, pp. 332–336, Feb. 2020.
[19] G. A. Quantum *et al.*, "Hartree-fock on a superconducting qubit quantum computer," *Science*, vol. 369, no. 6507, pp. 1084–1089, Aug. 2020, doi: 10.1126/science.abb9811.
[20] G. Aleksandrowicz *et al.*, "Qiskit: An open-source framework for quantum computing," IBM, Armonk, NY, USA, Tech. Rep., Jan. 2019. Accessed: Jan. 7, 2022, doi: 10.5281/zenodo.2562111.
[21] N. Moll *et al.*, "Quantum optimization using variational algorithms on near-term quantum devices," *Quantum Sci. Technol.*, vol. 3, no. 3, Jun. 2018, Art. no. 030503, doi: 10.1088/2058-9565/aab822.
[22] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, "Validating quantum computers using randomized model circuits," *Phys. Rev. A, Gen. Phys.*, vol. 100, no. 3, Sep. 2019, Art. no. 032328, doi: 10.1103/PhysRevA.100.032328.
[23] K. X. Wei *et al.*, "Verifying multipartite entangled Greenberger-Horne-Zeilinger states via multiple quantum coherences," *Phys. Rev. A, Gen. Phys.*, vol. 101, no. 3, Mar. 2020, Art. no. 032343, doi: 10.1103/PhysRevA.101.032343.
[24] D. González, D. F. De La Pradilla, and G. González, "Revisiting the experimental test of Mermin's inequalities at IBMQ," *Int. J. Theor. Phys.*, vol. 59, no. 12, pp. 3756–3768, Nov. 2020, doi: 10.1007/s10773-020-04629-4.
[25] W. Huang, W. Chien, C. Cho, C. Huang, T. Huang, and C. Chang, "Mermin's inequalities of multiple qubits with orthogonal measurements on IBM Q 53-qubit system," *Quantum Eng.*, vol. 2, no. 2, p. p.e45, Jun. 2020, doi: 10.1002/que2.45.
[26] B. Yang, R. Raymond, and S. Uno, "Efficient quantum readout-error mitigation for sparse measurement outcomes of near-term quantum devices," *Phys. Rev. A, Gen. Phys.*, vol. 106, no. 1, Jul. 2022, Art. no. 012423, doi: 10.1103/PhysRevA.106.012423.
[27] G. J. Mooney, C. D. Hill, and L. C. L. Hollenberg, "Entanglement in a 20-qubit superconducting quantum computer," *Sci. Rep.*, vol. 9, no. 1, p. 13465, Dec. 2019, doi: 10.1038/s41598-019-49805-7.
[28] M. Van den Nest, J. Dehaene, and B. De Moor, "Graphical description of the action of local Clifford transformations on graph states," *Phys. Rev. A, Gen. Phys.*, vol. 69, no. 2, Feb. 2004, Art. no. 022316, doi: 10.1103/PhysRevA.69.022316.
[29] V. G. Vizing, "On an estimate of the chromatic class of a p-graph," *Discret Analiz*, vol. 3, pp. 25–30, 1964.
[30] H.-Y. Huang, R. Kueng, and J. Preskill, "Predicting many properties of a quantum system from very few measurements," *Nature Phys.* vol. 16, pp. 1050–1057, Feb. 2020, doi: 10.1038/s41567-020-0932-7.
[31] C. Hadfield, S. Bravyi, R. Raymond, and A. Mezzacapo, "Measurements of quantum Hamiltonians with locally-biased classical shadows," *Commun. Math. Phys.*, vol. 391, no. 3, pp. 951–967, May 2022, doi: 10.1007/s00220-022-04343-8.
[32] S. Hillmich, C. Hadfield, R. Raymond, A. Mezzacapo, and R. Wille, "Decision diagrams for quantum measurements with shallow circuits," 2021, *arXiv:2105.06932*.
[33] T.-C. Yen, V. Verteletskyi, and A. F. Izmaylov, "Measuring all compatible operators in one series of single-qubit measurements using unitary transformations," *J. Chem. Theory Comput.*, vol. 16, no. 4, pp. 2400–2409, Apr. 2020, doi: 10.1021/acs.jctc.0c00008.
[34] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, "Measurement optimization in the variational quantum eigensolver using a minimum clique cover," *J. Chem. Phys.*, vol. 152, no. 12, Mar. 2020, Art. no. 124114, doi: 10.1063/1.5141458.
[35] I. Hamamura and T. Imamichi, "Efficient evaluation of quantum observables using entangled measurements," *npj Quantum Inf.*, vol. 6, no. 1, pp. 1–8, Jun. 2020, doi: 10.1038/s41534-020-0284-2.
[36] B. Wu, J. Sun, Q. Huang, and X. Yuan, "Overlapped grouping measurement: A unified framework for measuring quantum states," 2021, *arXiv:2105.13091*.
[37] T.-C. Yen, A. Ganeshram, and A. F. Izmaylov, "Deterministic improvements of quantum measurements with grouping of compatible operators, non-local transformations, and covariance estimates," 2022, *arXiv:2201.01471*.
[38] (2021). *Source Code of Master Thesis Project of Bo Yang at the University of Tokyo*. [Online]. Available: https://github.com/BOBO1997/master_thesis/
[39] J. Duchi, "Cs229 supplemental lecture notes hoeffding's inequality," Stanford Univ., Stanford, CA, USA, Lect. Note CS229, 2017.
[40] E. L. Hahn, "Spin echoes," *Phys. Rev.*, vol. 80, no. 4, pp. 580–594, Nov. 1950, doi: 10.1103/PhysRev.80.580.
[41] B. Pokharel, N. Anand, B. Fortman, and D. A. Lidar, "Demonstration of fidelity improvement using dynamical decoupling with superconducting qubits," *Phys. Rev. Lett.*, vol. 121, no. 22, Nov. 2018, Art. no. 220502, doi: 10.1103/PhysRevLett.121.220502.

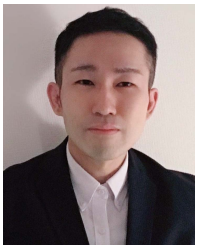


Bo Yang received the M.Sc. degree from the Department of Computer Science, Graduate School of Information Science and Technology, The University of Tokyo, Japan, in 2022. He is currently pursuing the Ph.D. degree with the Laboratoire d'Informatique de Sorbonne Université, France. His research interests include quantum error mitigation, quantum-classical hybrid algorithm, quantum verification, and quantum programming.

Rudy Raymond received the B.Eng. degree in computer science and the M.Sc. and Doctoral degrees in informatics from Kyoto University, in 2001, 2003, and 2006, respectively. He has been a Researcher at IBM Research Tokyo in Japan since 2006, a Project Researcher at the Quantum Computing Center, Keio University, since 2018, and a part-time Lecturer at the University of Tokyo since 2019. He was appointed as a Consulting Professor at the Department of Computer Science, The University of Tokyo, in August 2022.

Hiroshi Imai (Member, IEEE) received the B.Eng. degree in mathematical engineering and the D.Eng. degree in information engineering in 1981 and 1986, respectively. He became an Associate Professor at the Department of Information Engineering, Kyushu University, in 1986, and moved to the Department of Information Science, the University of Tokyo, in 1990, where he is currently a Professor with the Department of Computer Science. He led JST ERATO Quantum Computation and Information Project from 2000 to 2011. He has a leader of the field of quantum computing.

Hyungseok Chang received the M.Sc. degree from the Department of Computer Science, Graduate School of Information Science and Technology, The University of Tokyo, in 2021. His research interests include quantum algorithm of combinatorial optimization.



Hidefumi Hiraishi received the Ph.D. degree from the Department of Computer Science, University of Tokyo, in 2016. He became an Assistant professor at the Department of Computer Science, University of Tokyo, in 2016. He became an Associate Professor at the Department of Mathematics, CST, Nihon University, in 2021. His research interests include combinatorial geometry, combinatorial optimization, and quantum computation.