Experimental demonstration of a high-fidelity virtual two-qubit gate

Akhil Pratap Singh [1,*], Kosuke Mitarai,[2] Yasunari Suzuki,[3] Kentaro Heya,[4] Yutaka Tabuchi [4],
Keisuke Fujii,[2,4] and Yasunobu Nakamura [1,4]

[1]*Department of Applied Physics, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*
[2]*Graduate School of Engineering Science, Osaka University 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan*
[3]*NTT Computer and Data Science Laboratories, NTT Corporation, Musashino 180-8585, Japan*
[4]*RIKEN Center for Quantum Computing (RQC), Wako, Saitama 351-0198, Japan*

We experimentally demonstrate a virtual two-qubit gate and characterize it using quantum process tomography (QPT). The virtual two-qubit gate decomposes an actual two-qubit gate into single-qubit unitary gates and projection gates in quantum circuits for expectation-value estimation. We implement projection gates via midcircuit measurements. The deterministic sampling scheme reduces the number of experimental circuit evaluations required for decomposing a virtual two-qubit gate. We also apply quantum error mitigation to suppress the effect of measurement errors and improve the average gate fidelity of a virtual controlled-Z (CZ) gate to $f_{av} = 0.9938 \pm 0.0002$. Our results highlight a practical approach to implement virtual two-qubit gates with high fidelities, which are useful for simulating quantum circuits using fewer qubits and implementing two-qubit gates on a distant pair of qubits.

I. INTRODUCTION

Quantum computing research is progressing on a promising yet challenging path to realize a scalable fault-tolerant quantum computer, which is expected to have the capabilities to solve many problems intractable for their classical counterparts. However, current quantum devices with limited coherence times, low scalability, and non-negligible noises, termed as noisy intermediate scale quantum, or NISQ devices [1], are still far from being a full-fledged quantum computer. Nevertheless, they are still proving to be the testbeds for many promising quantum algorithms and quantum applications [2–8]. The recent experimental demonstrations of quantum computation with over 50 qubits [9–12] have motivated the practical interest in solving large problems using smaller quantum devices, despite the tradeoff of utilizing more classical resources.

To maximize the capabilities of a limited-sized NISQ quantum processor, various techniques for simulating large quantum circuits with smaller quantum devices have been proposed and demonstrated [13–20]. These techniques are useful for the expectation value estimation of a large quantum circuit, as they can reduce the hardware requirement by "cutting" the circuit, albeit with some overhead cost. Mitarai

and Fujii [21] proposed the "virtual two-qubit gate" technique, a general decomposition scheme for a two-qubit gate. The virtual two-qubit gate allows us to simulate a two-qubit gate from a quasiprobability decomposition of local single-qubit operations in the quantum circuits used for the expectation value estimation for observables. The virtual two-qubit gate scheme has been experimentally utilized on a distant pair of superconducting qubits to reduce the number of SWAP operations required and thus reducing the two-qubit errors [22]. However, the characterization of the virtual two-qubit gate was not performed, thereby limiting the ability to evaluate its quality.

In this work, we experimentally demonstrate a virtual two-qubit gate taking the example of a controlled-Z (CZ) gate and characterize it through the quantum process tomography (QPT) [23]. The virtual two-qubit gate requires the implementation of projection gates, which are nonunitary. We implement the projection gates through midcircuit measurements. However, this limits the fidelity of the virtual two-qubit gate since measurement errors are typically higher than single-qubit gate errors. We thus formulate the quantum error mitigation for midcircuit measurements and apply it to improve the average gate fidelity of the virtual two-qubit gate.

The rest of the paper is organized as follows: In Sec. II, we review the virtual two-qubit gate decomposition technique, and present our approach for its experimental implementation. We then discuss combining the quantum error mitigation with the virtual two-qubit gate. Sec. III provides details on the experimental device, gate implementation, and the characterization of the virtual CZ gate. We also demonstrate the improved average gate fidelity of the virtual CZ gate after incorporating the quantum error mitigation. Lastly, in Sec. IV

---

*apsingh@g.ecc.u-tokyo.ac.jp

we summarize our work and discuss potential directions for future work.

II. METHODS

A. Gate Decomposition

We can break down any two-qubit gate expressed in the form $e^{i\theta A \otimes B}$ into six local single-qubit operations, where $A^2 = B^2 = I$, and I is an identity matrix. We use the tilde ($\sim$) symbol to represent a superoperator $\tilde{U}$ corresponding to an operator U whose action on a density matrix ρ is defined as $\tilde{U}\rho = U\rho U^\dagger$.

The decomposition of a CZ gate is written as

$$
\begin{aligned}
\widetilde{\text{CZ}} = {}& \frac{1}{2}\left(\tilde{\mathcal{R}}_Z\left(\frac{\pi}{2}\right) \otimes \tilde{\mathcal{R}}_Z\left(\frac{\pi}{2}\right)\right) \\
& + \frac{1}{2}\left(\tilde{\mathcal{R}}_Z\left(-\frac{\pi}{2}\right) \otimes \tilde{\mathcal{R}}_Z\left(-\frac{\pi}{2}\right)\right) \\
& - \frac{1}{2}\sum_{\alpha_1,\alpha_2}\alpha_1\alpha_2\left[\tilde{\Pi}_{\alpha_1 Z} \otimes \tilde{\mathcal{R}}_Z\left((\alpha_2+1)\frac{\pi}{2}\right)\right. \\
& \left. + \tilde{\mathcal{R}}_Z\left((\alpha_1+1)\frac{\pi}{2}\right) \otimes \tilde{\Pi}_{\alpha_2 Z}\right],
\end{aligned}
\tag{1}
$$

where $\alpha_1, \alpha_2 \in \{-1, +1\}$. $\tilde{\mathcal{R}}_Z(\pm\theta)$ and $\tilde{\Pi}_{\pm Z}$ are superoperators corresponding to the operators $\mathcal{R}_Z(\pm\theta) = e^{\pm iZ\theta/2}$ and $\Pi_{\pm Z} = (I \pm Z)/2$, respectively. $\Pi_{\pm Z}$ are the (nondestructive) projection gates, which project the quantum state onto the respective eigenbasis and are nonunitary [21].

Reference [21] has quantified the decomposition cost in terms of quasiprobability simulation where the decomposition operators in Eq. (1) are sampled with probabilities proportional to their coefficients [24]. The set of these decomposition operators can also be written as

$$
\begin{aligned}
\mathcal{D}_{\text{CZ}} = {}& \left\{\mathcal{R}_Z\left(\frac{\pi}{2}\right) \otimes \mathcal{R}_Z\left(\frac{\pi}{2}\right), \mathcal{R}_Z\left(-\frac{\pi}{2}\right) \otimes \mathcal{R}_Z\left(-\frac{\pi}{2}\right),\right. \\
& \left. \mathcal{R}_Z(\pi) \otimes \Pi_{\pm Z}, I \otimes \Pi_{\pm Z}, \Pi_{\pm Z} \otimes \mathcal{R}_Z(\pi), \Pi_{\pm Z} \otimes I\right\}.
\end{aligned}
\tag{2}
$$

In this work, we employ deterministic sampling for the decomposition operators, which means that we perform uniform experiments with each operator from Eq. (2) in a deterministic manner. Unlike the quasiprobability approach, we uniformly sample the decomposition operators with a probability of unity, regardless of their coefficients. We do that because of the following two reasons. First, by implementing deterministic sampling, we can reduce the total number of circuit evaluations for each subcircuit from six to five. It is explained in detail later in this section. Second, apart from the virtual two-qubit gate implementation, we also perform its characterization. To obtain the process matrix for the virtual CZ gate, we need process matrices of all the decomposition operators $\mathcal{D}_j^i$ (see Sec. III B). Thus, we implement each decomposition operation deterministically.

Figure 1(a) illustrates the circuit decomposition for a virtual CZ gate in terms of experimental implementation. We consider a simple two-qubit quantum circuit consisting of single-qubit gates and a CZ gate. The aim here is to do the

expectation value estimation for this two-qubit circuit using only one qubit by replacing an actual CZ gate with a virtual CZ gate. For the virtual two-qubit gate decomposition, we cut the CZ gate and consider the remaining circuit consisting of two one-qubit subcircuits A and B. Each subcircuit consists of the initial state ρ^i, set of single-qubit gates implemented before and after the CZ gates denoted by U_{pre}^i and U_{post}^i, respectively, and the measurement of observable k, $\mathcal{M}_k^i$, where the superscript $i \in \{\text{A}, \text{B}\}$ refers to the subcircuit they belong to. In order to do the expectation value estimation using a virtual two-qubit gate, we do five different circuit evaluations, labeled as "Decomposition circuits" in Fig. 1(a), where we replace the actual CZ gate with local decomposition operators $\mathcal{D}_j^i$ on both control and target qubits shown on the right hand side of Fig. 1(a). In $\mathcal{D}_j^i$, $i \in \{\text{A}, \text{B}\}$ and $j \in \{1, 2, 3, 4, 5\}$ refer to the indices of the subcircuits and the decomposition circuits, respectively. Single-qubit unitary operations $\mathcal{D}_{j=1,2,4,5}^i \in \{\mathcal{R}_Z(\pi/2), \mathcal{R}_Z(-\pi/2), I, Z\}$ are shown in the yellow boxes, while projection gates $\mathcal{D}_{j=3}^i = \{\Pi_{\pm Z}\}$ are shown in the green boxes. For the expectation value estimation using the virtual two-qubit gate, we substitute the expectation values ν_j^i obtained from decomposition circuits 1–5 in Eq. (1) and get the final expectation value. This calculation is illustrated in Fig. 1(b), which is a graphical representation of Eq. (1), where the ten pairs of boxes correspond to the ten terms in Eq. (1).

Note that the utilization of the virtual two-qubit gate can allow us to decompose a quantum circuit U into two completely separate quantum circuits each acting only on qubits in subcircuits A and B, as considered in Fig. 1(a). We take input to the circuit as a separable state $\rho = \rho_{\text{A}} \otimes \rho_{\text{B}}$. As an output, we wish to obtain expectation value $P = P_{\text{A}} \otimes P_{\text{B}}$, where P_{A} and P_{B} are Pauli operators acting on groups A and B, respectively. For example, if we define two superoperators $\tilde{U}_1$ and $\tilde{U}_2$ then their tensor product is expressed as

$$
(\tilde{U}_1 \otimes \tilde{U}_1)\rho = (U_1 \otimes U_2)\rho(U_1^\dagger \otimes U_2^\dagger).
\tag{3}
$$

Using Eq. (3), U, as shown in Fig. 1(a), can be decomposed as

$$
\tilde{U} = \sum_j \left(\tilde{U}_{\text{post}}^{\text{A}} \otimes \tilde{U}_{\text{post}}^{\text{B}}\right)\left(\tilde{\mathcal{D}}_j^{\text{A}} \otimes \tilde{\mathcal{D}}_j^{\text{B}}\right)\left(\tilde{U}_{\text{pre}}^{\text{A}} \otimes \tilde{U}_{\text{pre}}^{\text{B}}\right).
\tag{4}
$$

The desired expectation value can be written as

$$
\begin{aligned}
\text{Tr}(P\tilde{U}\rho) = \sum_j \Big[& \text{Tr}\left(P_{\text{A}}\tilde{U}_{\text{post}}^{\text{A}}\tilde{\mathcal{D}}_j^{\text{A}}\tilde{U}_{\text{pre}}^{\text{A}}\rho_{\text{A}}\right) \\
& \times \text{Tr}\left(P_{\text{B}}\tilde{U}_{\text{post}}^{\text{B}}\tilde{\mathcal{D}}_j^{\text{B}}\tilde{U}_{\text{pre}}^{\text{B}}\rho_{\text{B}}\right)\Big].
\end{aligned}
\tag{5}
$$

Therefore, in this case, we can estimate all of the values of $\text{Tr}(P_{\text{A}}\tilde{U}_{\text{post}}^{\text{A}}\tilde{\mathcal{D}}_j^{\text{A}}\tilde{U}_{\text{pre}}^{\text{A}}\rho_{\text{A}})$ and $\text{Tr}(P_{\text{B}}\tilde{U}_{\text{post}}^{\text{B}}\tilde{\mathcal{D}}_j^{\text{B}}\tilde{U}_{\text{pre}}^{\text{B}}\rho_{\text{B}})$ separately for $j = 1, \ldots, 5$ [as shown in Fig. 1(a)] and then combine them according to Eqs. (1) and (5) to obtain the result for U. We note that for quasiprobability sampling of the decomposition operations, Eq. (1) needs six different circuit evaluations because there are six distinct operations [see Eq. (2)] involved in Eq. (1). However, by comparing Eqs. (1), (2), and (5), we see that if we sample the decomposition operators $\mathcal{D}_j^i$ deterministically then we can reduce the total number of circuit evaluations for each subcircuit from six to five. This can
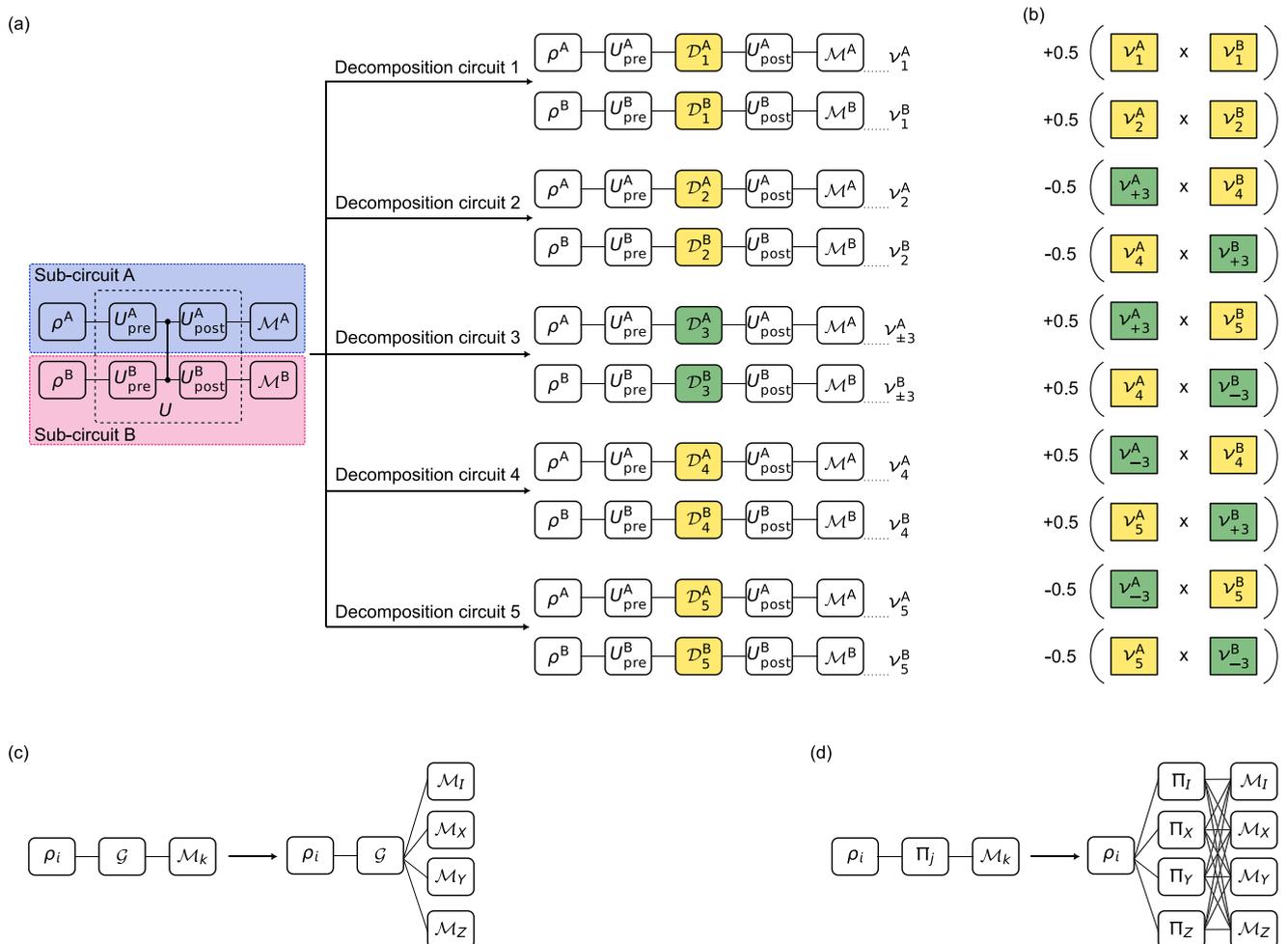
FIG. 1. Decomposition of a CZ gate and quantum error mitigation. (a) Decomposition of a simple two-qubit quantum circuit with a single CZ gate and other one-qubit gates. See text for the explanation. (b) Calculation of the final expectation value using the expectation values ν_j^i obtained in the respective decomposition circuits in (a). (c) Measurement error mitigation for a single-qubit unitary operation. (d) Quantum error mitigation for a projection gate and a measurement.

be done since, by implementing the projection gates through midcircuit measurements, we obtain both expectation values of the projection operators ($\Pi_{\pm Z}$ in case of a CZ gate) in a single experiment [$\nu_{\pm 3}^{A,B}$ in Fig. 1(a)].

B. Combining quantum error mitigation with the virtual two-qubit gate

We use midcircuit measurements to implement the projection gates. The measurement errors become the dominant contributor to the average gate infidelity of the virtual two-qubit gate since the measurement errors (3.9%) are much larger than the initialization errors ($\sim 1\%$) and single-qubit errors ($<0.05\%$) in our device. The details of the device are presented in Sec. III A. In order to mitigate the measurement errors, we perform quantum error mitigation (QEM) with the probabilistic error cancellation (PEC) method [20,25]. In this method, we define a set of modified circuits for the original circuit in which we wish to mitigate the errors. Then, the original circuit is replaced randomly by a selected set of modified circuits, where the probabilities to select the random circuits are calculated before running the quantum circuits. The mitigated result is obtained by taking the average of the

modified circuits. Here, we apply this error mitigation technique both to the midcircuit measurements for the projection gates and to the final measurements. Fig. 1(c) illustrates the measurement error mitigation circuit for single-qubit unitary operations where $\mathcal{G} = \mathcal{D}_{j=\{1,2,4,5\}}$, and instead of the measurement of the desired observable k, $\mathcal{M}_k$, we measure the set of Pauli operators denoted as $\mathcal{M}_{\mathcal{P}} = \mathcal{M}_{\{I,X,Y,Z\}}$. In the experiments, measurements on the X and Y basis are implemented by applying $\mathcal{R}_Y(\pi/2)$ and $\mathcal{R}_X(\pi/2)$ gates, respectively, just before the usual Z-basis measurement. For measurement in the I basis, we always consider the expectation value of operator I to be unity. In case of the projection gates $\mathcal{D}_{j=3}$, we apply the probabilistic quantum error mitigation sequentially [25], first for the projection gate Π_j, implemented through midcircuit measurement, and then on the final measurement $\mathcal{M}_k$ as shown in Fig. 1(d). See Appendix A for the details.

III. EXPERIMENT

A. System details

Local single-qubit operations can be classified into two categories: unitary and nonunitary/projection gates. To im-

TABLE I. Parameters of the superconducting transmon qubit used in the experiments: the qubit frequency ω_q, anharmonicity α, energy relaxation time T_1, and Ramsey dephasing time T_2^*.

	$\omega_q/2\pi$	$\alpha/2\pi$	T_1	T_2^*
Qubit	8.403 GHz	−416 MHz	20.2 μs	3.1 μs

plement a virtual CZ gate experimentally, we use a fixed-frequency superconducting transmon qubit [26]. The qubit is a part of a 16-qubit device [27]. The parameters of the qubit are summarized in Table I.

In experiments, the quantum gates and measurements are noisy, and thus we do not have perfect implementations of the above gates as mentioned in Eq. (1). Thus, hereafter we use the superscript "exp" to denote a quantum operation realized in the experiments and hence containing the errors. Let noisy quantum channels that correspond to Z, $\mathcal{R}_Z(\pm\theta)$, and $\Pi_{\pm Z}$ be Z^{exp}, $\mathcal{R}_Z^{\text{exp}}(\pm\theta)$, and $\Pi_{\pm Z}^{\text{exp}}$, respectively. To implement $\mathcal{R}_Z^{\text{exp}}(\pm\theta)$ and Z^{exp}, we use the efficient Z-gate technique [28], in which the rotation around the Z axis can be executed by adjusting the relative phase of the subsequent X and Y qubit drive pulses, which is controlled through the classical hardware and software and can be implemented almost perfectly. The $\mathcal{R}_X^{\text{exp}}(\pi/2)$, $\mathcal{R}_Y^{\text{exp}}(\pi/2)$ gates are implemented with the shaped microwave pulses [29]. For suppressing leakage, we use the DRAG pulsing technique [30,31]. To reduce the unitary control errors, we use the ORBIT technique [32]. The average gate fidelity of the single-qubit Clifford gates is 0.9992 ± 0.0004 via randomized benchmarking [33–35].

To implement $\Pi_{\pm Z}^{\text{exp}}$, we perform midcircuit measurement followed by classical postprocessing. The midcircuit measurement consists of a 90-ns-long dispersive readout pulse followed by a resonator-reset time of 500 ns. The qubit state is measured through an off-resonantly coupled resonator with the dispersive readout [36] performed via the readout resonator at 10.310 GHz. The readout signal is amplified with an impedance-matched Josephson parametric amplifier [37,38]. The averaged assignment fidelity for the qubit readout is 0.9609 ± 0.0037 (see more details in Appendix B).

B. Characterization and results

To represent quantum states, quantum channels, and measurements, we use the Pauli transfer matrix (PTM) representation. We define σ_k ($k = 0, \ldots, 3$) as the kth operator in the Pauli basis $\mathcal{P} = \{I, X, Y, Z\}$. We define a quantum channel corresponding to a decomposition operation $\mathcal{D} \in \mathcal{D}_j^i$ as $\tilde{\mathcal{D}}$. The PTM representation of a quantum channel $\tilde{\mathcal{D}}$ can be written as T, whose elements are given by

$$T(\tilde{\mathcal{D}})_{k,l} = \frac{1}{d}\text{Tr}[\sigma_k \tilde{\mathcal{D}}(\sigma_l)], \qquad (6)$$

where $d = 2^n$ is the dimension of the n-qubit system. For the virtual CZ gate characterization, we perform QPT. Fig. 1(a) also represents the circuit for virtual two-qubit gate QPT, where we perform QPT in each subcircuit for the corresponding decomposition channel $\tilde{\mathcal{D}}_{j=1,\ldots,5}^{i=\text{A,B}}$. As per the QPT procedure, we prepare $d^2 = 4$ linearly independent states from the basis $\rho_{\text{in}}^{i=\text{A,B}} = \{|0\rangle, |1\rangle, |+\rangle, |i+\rangle\}$, each of which

is subjected to the quantum channel $\tilde{\mathcal{D}}_j^i$, followed by the quantum state tomography (QST) [2], which includes measurements of the $\{X, Y, Z\}$ observables. Here, we have denoted $|\pm\rangle$, $|i\pm\rangle$, and $|0\rangle$ and $|1\rangle$ as the eigenvectors of X, Y, and Z with eigenvalues ± 1, respectively. Since we are interested in the QPT of $\tilde{\mathcal{D}}_j^i$, we take corresponding $U_{\text{pre,post}}^{i=\text{A,B}}$ to be the identity gates I^{A} and I^{B}. These set of experiments constitute a QPT circuit evaluation for the channel $\tilde{\mathcal{D}}_j^i$. After obtaining $T(\tilde{\mathcal{D}}_j^i)$ for both subcircuits A and B, we use Eq. (1) to calculate the PTM for the virtual CZ gate $T_{\text{V}-CZ}$. See Appendix C.

We present the characterization results for the virtual CZ gate and compare the cases with (PEC) and without (non-PEC) quantum error mitigation applied for the projection gates in Fig. 2. For each QPT circuit evaluation of $\tilde{\mathcal{D}}_j^i$, we take $N_s = 10\,000$ shots to calculate the average expectation value for the corresponding observables $\mathcal{M}_{k=\{X,Y,Z\}}$. Note that QPT assumes error-free state preparation and measurement (SPAM). To mitigate initialization errors, we implement initialization by measurement and postselection, in which a measurement pulse is applied at the beginning of each circuit and postselecting the shots which were prepared in the ground state. To mitigate the measurement errors, we apply error mitigation to the measurements, even in experiments without PEC for projection gates, referred to as non-PEC QPT experiments. In contrast, PEC QPT experiments involve the implementation of PEC for both projection gates and the measurements. For the measurement error mitigation in non-PEC QPT experiments, we replace each $\mathcal{M}_k$ with the measurement set of Pauli observables $\mathcal{M}_p \in \{\mathcal{M}_I, \mathcal{M}_X, \mathcal{M}_Y, \mathcal{M}_Z\}$ as shown in Fig. 1(c). The circuit for PEC QPT experiments is shown in Fig. 1(d), where PEC is applied sequentially first for the projection gates and then for the measurements. See Appendix A for the details. To get the final mitigated expectation value with its uncertainty, we repeat each QPT circuit 100 times and take its average. The average gate fidelity for the virtual CZ gate without doing the quantum error mitigation for the projection gates is $f_{\text{av}} = 0.9782 \pm 0.0001$, and after applying the quantum error mitigation to the projection gates it is improved to $f_{\text{av}}^{\text{mit}} = 0.9938 \pm 0.0002$, where "mit" denotes the fidelity after applying the PEC mitigation. The results indicate a feasible and practical approach to experimentally implement virtual two-qubit gates with enhanced fidelity by employing measurement error mitigation techniques for the projection gates performed through midcircuit measurements and implementing high-fidelity local single-qubit gates. The high-fidelity virtual two-qubit gates are crucial for simulating quantum circuits using fewer qubits and also to simulate two-qubit gates on a distant pair of qubits which otherwise would take several number of SWAP operations between them.

IV. CONCLUSION AND DISCUSSION

In this work, we experimentally demonstrated a virtual CZ gate and characterized its performance using QPT. Our approach involved formulating probabilistic error cancellation for the projection gates, implemented through midcircuit measurements. This effectively mitigated projection gate errors. Furthermore, incorporating measurement-error mitigation led
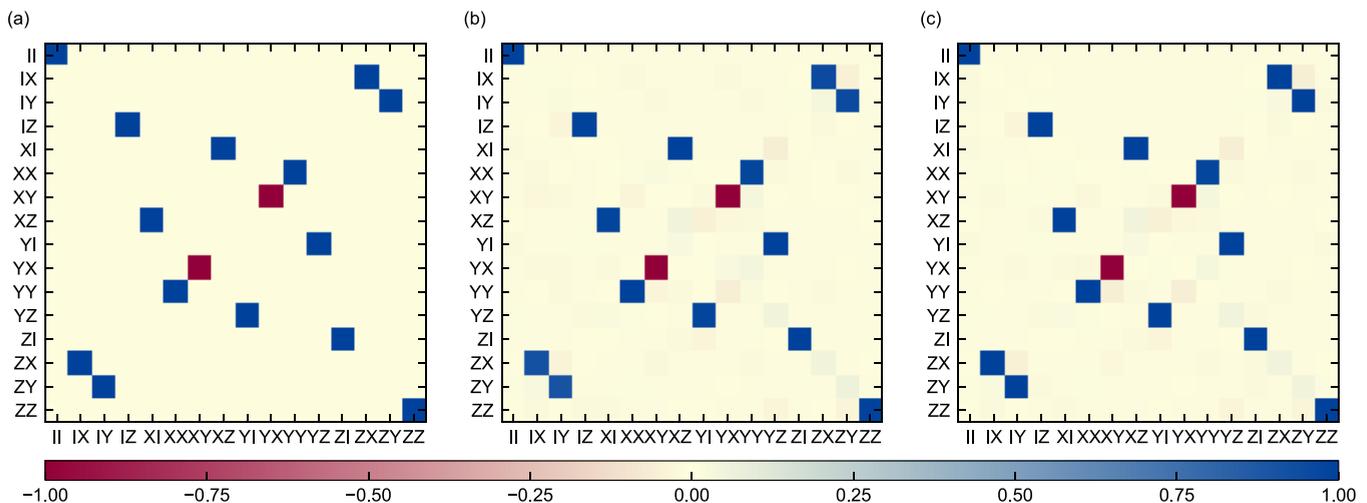
FIG. 2. Quantum process tomography results for a CZ gate. (a) Ideal CZ gate in the Pauli transfer matrix representation. (b) Experimentally-implemented virtual CZ gate without quantum error mitigation. (c) Experimentally-implemented virtual CZ gate with quantum error mitigation.

to a significant enhancement in the average gate fidelity of the virtual CZ gate, even in the presence of higher measurement errors compared to single-qubit gate errors. In one of the applications of the virtual two-qubit gates, where we broke a two-qubit circuit into two disconnected one-qubit circuits, we reduced the number of circuit evaluations for each one-qubit circuit from six to five.

As demonstrated in this work, through the implementation of projection gates using midcircuit measurements, we can eliminate the need for ancilla qubits typically used for qubit reset. This also allows us to achieve projection on all observable eigenvectors in a single experiment. However, in midcircuit measurements, the resonator-reset time after the readout pulse contributes to the overall projection gate time. To address this issue, accelerating the resonator ringdown time has the potential to further enhance the fidelities of virtual two-qubit gates [39]. This improvement is expected to be more pronounced in scenarios involving the implementation of multiple virtual two-qubit gates.

Decomposing a larger quantum circuit into smaller circuits can boost the capabilities of the limited-sized NISQ devices which can pave the way towards the goal of demonstrating the quantum advantage. Circuit-cutting schemes, such as the virtual two-qubit gate, offer potential benefits for NISQ algorithms [3] such as VQE and QAOA [40], facilitating the simulation of large circuits with fewer physical qubits. In combination with quantum error mitigation, as demonstrated in this work, the virtual two-qubit gates could potentially improve the simulation of a large quantum circuit with less physical qubits in certain scenarios. However, decomposing a quantum circuit requires large overhead in terms of additional number of circuit evaluations and classical post-processing. Using the deterministic approach considered in this work, the scaling of the required number of circuit evaluations to cut k virtual two-qubit gates is $\mathcal{O}(5^k)$. An interesting avenue for future exploration lies in developing a general approach for optimally decomposing a quantum circuit, one that can be efficiently implemented in experimental settings.

APPENDIX A: QUANTUM ERROR MITIGATION

We use the probabilistic error cancellation method for QEM [20,25]. In the PTM representation, a state ρ can be expressed as a column vector

$$|\rho\rangle\!\rangle = [\rho_0, \rho_1, \cdots, \rho_k, \cdots]^{\mathrm{T}}, \qquad (A1)$$

whose kth element is given by

$$|\rho\rangle\!\rangle_k = \mathrm{Tr}[\rho\sigma_k], \qquad (A2)$$

where σ_k is defined as the kth operator of the Pauli basis $\mathcal{P} = \{I, X, Y, Z\}$. Similarly, an observable Q can be expressed as a row vector

$$\langle\!\langle Q| = [Q_0, Q_1, \cdots, Q_k, \cdots], \qquad (A3)$$

where

$$\langle\!\langle Q|_k = \mathrm{Tr}[Q\sigma_k]. \qquad (A4)$$

A quantum channel $\tilde{\mathcal{O}}$ corresponding to the operation $\mathcal{O}$ is expressed as

$$\tilde{\mathcal{O}}\rho = \mathcal{O}\rho\mathcal{O}^\dagger, \qquad (A5)$$

whose PTM representation is given in Eq. (6).

In the QPT experiments, as also mentioned in Sec. III B, we take the initial states as $\rho_i = \{|0\rangle, |1\rangle, |+\rangle, |i+\rangle\}$ and measure the observables $Q \in \{X, Y, Z\}$. We want to compute the noise-free (mitigated) expectation values $E^{\mathrm{mit}} = \langle\!\langle Q|\tilde{O}|\rho\rangle\!\rangle$,
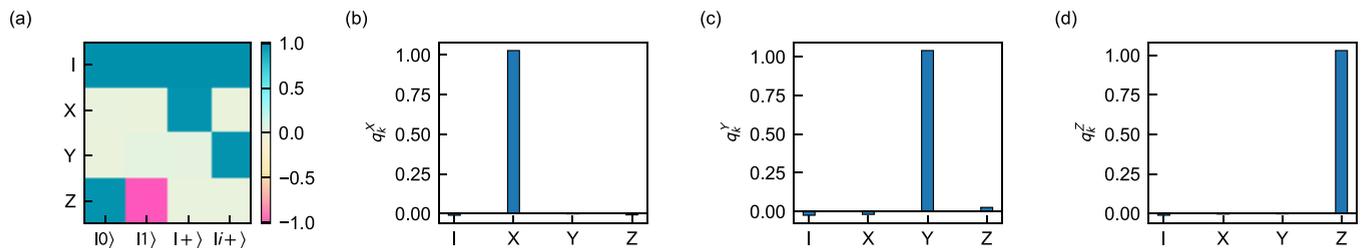
FIG. 3. Gram matrix and quasiprobability vectors. (a) Gram matrix for a single-qubit measurement. (b)–(d) Quasiprobability vectors for measurement observables X, Y, and Z, respectively.

however, in experiments we obtain $E^{\text{exp}} = \langle\langle Q^{\text{exp}}|\tilde{\mathcal{O}}^{\text{exp}}|\rho^{\text{exp}}\rangle\rangle$. Hence, we apply PEC method to mitigate the errors. The set of initial states can be expressed as the state preparation matrix A^{exp} with its elements $A_{k,i}^{\text{exp}} = \langle\langle\sigma_k|\rho_i^{\text{exp}}\rangle\rangle$ and the set of observables as the readout matrix B^{exp} with its elements $B_{k,i}^{\text{exp}} = \langle\langle\sigma_k^{\text{exp}}|\sigma_i\rangle\rangle$. Since the state preparation errors are much lower in our device, and we also perform initialization by measurement and postselection, we assume the state preparation to be almost perfect ($\rho^{\text{exp}} = \rho$) and take its estimate as error-free $A_{k,i}^{\text{exp}} \approx \langle\langle\sigma_k|\rho_i\rangle\rangle$, which is a decent guess and is expressed as

$$A^{\text{exp}} \approx \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \end{pmatrix}. \tag{A6}$$

The Gram matrix $\mathbb{G}^{\text{exp}} = B^{\text{exp}}A^{\text{exp}}$, is obtained experimentally by performing the QPT experiment for an identity gate I, as also shown in Fig. 3(a). With this, we can obtain the readout matrix $B^{\text{exp}} = \mathbb{G}^{\text{exp}}(A^{\text{exp}})^{-1}$.

In the case of unitary decomposition gates, as shown in Fig. 1(a), the single-qubit gates $\mathcal{O} \in \mathcal{D}_{j=\{1,2,4,5\}}^{i=\text{A,B}}$ are implemented with high fidelity, so we assume them to be error-free ($\tilde{\mathcal{O}}^{\text{exp}} = \tilde{\mathcal{O}}$). Due to non-negligible measurement errors, we apply QEM for the measurements, as shown in Fig. 1(c). We can express the observables in terms of experimentally obtained noisy observables [20,41] as

$$\langle\langle\sigma_k| = \sum_l q_k^l \langle\langle\sigma_l^{\text{exp}}|, \tag{A7}$$

where $k \in \{1, 2, 3\}$, $l \in \{0, 1, 2, 3\}$, q_k^l are the elements of quasiprobability vector q_k, and $\langle\langle\sigma_l^{\text{exp}}|$ is the lth row of B^{exp}. The quasiprobability vectors $q_{k\in\{X,Y,Z\}}$ are shown in Figs. 3(b)–3(d), respectively, and can be computed as

$$q_k = \langle\langle\sigma_k|(B^{\text{exp}})^{-1}. \tag{A8}$$

The probability with which σ_l^{exp} is sampled is given by

$$p_k^l = \frac{|q_k^l|}{\sum_m |q_k^m|}. \tag{A9}$$

To obtain the final mitigated expectation value E_k^{mit}, each outcome from the experiment is multiplied by a weight factor

$$w_k^l = \text{sgn}(q_k^l) \sum_m |q_k^m|. \tag{A10}$$

The weighted average is the mitigated expectation value for the observable σ_k. Using Eqs. (A7) to (A10), it can be written as

$$E_k^{\text{mit}} = \sum_l q_k^l \langle\langle\sigma_l^{\text{exp}}|\tilde{\mathcal{D}}_j^i|\rho\rangle\rangle,$$

$$= \sum_l w_k^l p_k^l \langle\langle\sigma_l^{\text{exp}}|\tilde{\mathcal{D}}_j^i|\rho\rangle\rangle, \tag{A11}$$

where $\tilde{\mathcal{D}}_j^i \in \tilde{\mathcal{D}}_{j=\{1,2,4,5\}}^{i=\{A,B\}}$. In the case of $\langle\sigma_0^{\text{exp}}\rangle = \langle I^{\text{exp}}\rangle$, we always consider the expectation value as $+1$.

Now, we discuss applying PEC method on the decomposition circuits containing projection gates $\Pi_{\pm Z}$. We perform the PEC method sequentially first for the projection gates and then for the measurements as shown in Fig. 1(d). Since we implement projection gates through midcircuit measurements, we make an approximation here, i.e., we express a projection gate operation ($\tilde{\Pi}_{\pm Z}$) as a linear combination of the basis operations $\mathcal{B}^{\text{exp}} \in \{\tilde{I}, \tilde{\Pi}_{+X}, \tilde{\Pi}_{-X}, \tilde{\Pi}_{+Y}, \tilde{\Pi}_{-Y}, \tilde{\Pi}_{+Z}, \tilde{\Pi}_{-Z}\}$. The PTMs $\mathcal{B}^{\text{exp}}$ are estimated using Eq. (6) with measurement-error mitigation. We can express the PTMs corresponding to Π_Z as

$$\tilde{\Pi}_{+Z} \simeq \sum_u \gamma_u \mathcal{B}_u^{\text{exp}}, \tag{A12}$$

$$\tilde{\Pi}_{-Z} \simeq \sum_u \delta_u \mathcal{B}_u^{\text{exp}}, \tag{A13}$$

where $u \in \{1, \ldots, 7\}$ denotes the uth element of $\mathcal{B}^{\text{exp}}$, and γ and δ are quasiprobability vectors, as shown in Fig. 4. To calculate γ and δ, we minimize the Euclidean 2-norms $||\sum_u \gamma_u \mathcal{B}_u^{\text{exp}} - \tilde{\Pi}_{+Z}||$ and $||\sum_u \delta_u \mathcal{B}_u^{\text{exp}} - \tilde{\Pi}_{-Z}||$, respectively.

The probabilities of selecting $\mathcal{B}_u^{\text{exp}}$ corresponding to $\tilde{\Pi}_{+Z}$ and $\tilde{\Pi}_{-Z}$ from Eqs. (A12) and (A13) can therefore be expressed as

$$p_u^\gamma = \frac{|\gamma_u|}{\sum_r |\gamma_u^r|}, \tag{A14a}$$

$$p_u^\delta = \frac{|\delta_u|}{\sum_s |\delta_u^s|}, \tag{A14b}$$

respectively. The respective weight factors associated with Eqs. (A14a) and (A14b) can be written as

$$w_u^\gamma = \text{sgn}(\gamma_u) \sum_r |\gamma_u^r|, \tag{A15a}$$

$$w_u^\delta = \text{sgn}(\delta_u) \sum_s |\delta_u^s|. \tag{A15b}$$

We can express the mitigated expectation value of observables σ_k' as

$$E_{k'}^{\text{mit}} = \langle\langle\sigma_{k'}|\tilde{\mathcal{D}}_3^i|\rho\rangle\rangle, \tag{A16}$$
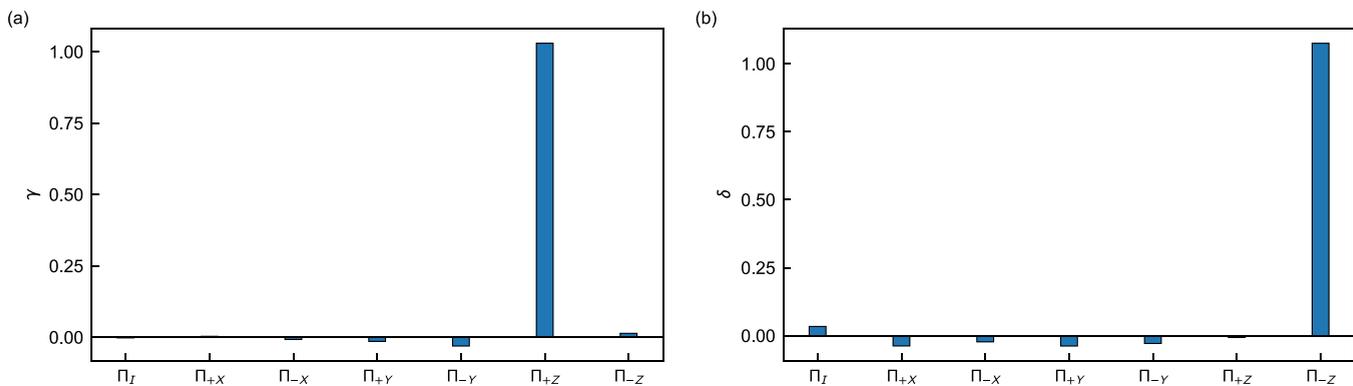
FIG. 4. Quasiprobability vectors for projection gates. (a) Quasiprobability vector γ corresponding to projection gate $\tilde{\Pi}_{+Z}$. (b) Quasiprobability vector δ corresponding to projection gate $\tilde{\Pi}_{-Z}$.

where $k' = \{1, 2, 3\}$ and $\tilde{\mathcal{D}}_3^{i=\{A,B\}} = \tilde{\Pi}_{\pm Z}$. Using Eqs. (A7) to (A10) and Eqs. (A12) to (A15), we can express Eq. (A16) for $\tilde{\Pi}_{+Z}$ and $\tilde{\Pi}_{-Z}$ as

$$E_{+Z,k'}^{\text{mit}} = \sum_l \sum_u w_{k'}^l p_k^l w_u^\gamma p_u^\gamma \langle\langle \sigma_l^{\exp} | \mathcal{B}_u^{\exp} | \rho \rangle\rangle, \quad \text{(A17a)}$$

$$E_{-Z,k'}^{\text{mit}} = \sum_l \sum_u w_{k'}^l p_k^l w_u^\delta p_u^\delta \langle\langle \sigma_l^{\exp} | \mathcal{B}_u^{\exp} | \rho \rangle\rangle, \quad \text{(A17b)}$$

respectively.

APPENDIX B: QUBIT READOUT CHARACTERIZATION

To obtain the average assignment fidelity of the single-shot readout, we apply a $\mathcal{R}_X(\pi/2)$ gate on the initial ground state $|0\rangle$ and then apply two sequential measurement pulses. The average assignment fidelity is defined as

$$\mathcal{F}_a = \tfrac{1}{2}(p(g|g) + p(e|e)), \quad \text{(B1)}$$

where $p(x|y)$ is the probability of assigning the measurement outcome x when the qubit is prepared in the state y. We obtain the assignment fidelity of 0.9609 ± 0.0037 and the ground-state initialization fidelity $\mathcal{F}_{\text{init}} = p(g|g)$ of 0.9895 ± 0.0028.

APPENDIX C: QUANTUM PROCESS TOMOGRAPHY

We use QPT to evaluate the performance of the virtual two-qubit gate. In general, for an n-qubit system with the dimension $d = 2^n$, we can express a quantum channel $\tilde{\mathcal{E}}$ acting on an arbitrary quantum state ρ as

$$\tilde{\mathcal{E}}(\rho) = \sum_{i,j=0}^{d^2-1} \chi_{ij} B_i \rho B_j^\dagger, \quad \text{(C1)}$$

where $\{B_i\}$ are the elements of the $d \times d$ matrix basis and χ is the process matrix [2]. We define the PTM representation of a quantum channel $\tilde{\mathcal{E}}$ as T, whose elements are expressed as

$$T(\tilde{\mathcal{E}})_{i,j} = \text{Tr}[\sigma_i \tilde{\mathcal{E}}(\sigma_j)]. \quad \text{(C2)}$$

In the PTM representation, we can express a state ρ as a column vector $|\rho\rangle\rangle$ with elements $|\rho\rangle\rangle_k = \text{Tr}[\sigma_k \rho]$ and an operator $\mathcal{O}$ as a row vector $\langle\langle \mathcal{O}|$ with elements $\langle\langle \mathcal{O}|_k = \text{Tr}[\sigma_k \mathcal{O}]$. In the QPT, we prepare for each qubit an initial state from the set $\{|0\rangle, |1\rangle, |+\rangle, |i+\rangle\}$. For each initial state, we measure the qubit along the Pauli basis, X, Y, and Z.

Using the PTM representation, we can define the quantum channel for the two-qubit system as the tensor product of individual PTMs. Using Eqs. (1) and (6), we can express the PTM for the virtual CZ gate T_{V-CZ} as

$$T_{V-CZ} = \frac{1}{2} \sum_{i=1,2} \left(T(\tilde{\mathcal{D}}_i^A) \otimes T(\tilde{\mathcal{D}}_i^B) \right)$$
$$- \frac{1}{2} \sum_{\alpha_1,\alpha_2,\beta} \alpha_1, \alpha_2 \left[\left(T(\tilde{\mathcal{D}}_{\alpha_13}^A) \otimes T(\tilde{\mathcal{D}}_\beta^B) \right) + \left(T(\tilde{\mathcal{D}}_\beta^A) \right. \right.$$
$$\left. \left. \otimes T(\tilde{\mathcal{D}}_{\alpha_23}^B) \right) \right], \quad \text{(C3)}$$

where $\alpha_1, \alpha_2 \in \{\pm 1\}$ and $\beta \in \{4, 5\}$. The average gate fidelity in terms of PTMs is defined as

$$f_{\text{av}}(\tilde{\mathcal{U}}, \tilde{\mathcal{E}}) = \frac{1}{d+1} \left[\frac{1}{d} \sum_{i,j} T(\tilde{\mathcal{U}})_{i,j} T(\tilde{\mathcal{E}})_{i,j} + 1 \right], \quad \text{(C4)}$$

where $d = 2^n$ is the dimension of the n-qubit system [42]. $T(\tilde{\mathcal{U}})$ is the PTM of the ideal target gate (T_{CZ} for CZ gate), and $T(\tilde{\mathcal{E}})$ is the PTM of experimentally implemented gate (T_{V-CZ} for virtual CZ gate in this case).

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

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).

[3] 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).

[4] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[5] A. Montanaro, Quantum speedup of monte carlo methods, Proc. R. Soc. A. **471**, 20150301 (2015).

[6] 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).

[7] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, Quantum chemistry in the age of quantum computing, Chem. Rev. **119**, 10856 (2019).

[8] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[9] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[10] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li *et al.*, Strong quantum computational advantage using a superconducting quantum processor, Phys. Rev. Lett. **127**, 180501 (2021).

[11] Q. Zhu, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li, S. Li *et al.*, Quantum computational advantage via 60-qubit 24-cycle random circuit sampling, Sci. Bull. **67**, 240 (2022).

[12] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, Nature (London) **618**, 500 (2023).

[13] S. Bravyi, G. Smith, and J. A. Smolin, Trading classical and quantum computational resources, Phys. Rev. X **6**, 021043 (2016).

[14] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, Simulating large quantum circuits on a small quantum computer, Phys. Rev. Lett. **125**, 150504 (2020).

[15] C. Ying, B. Cheng, Y. Zhao, H.-L. Huang, Y.-N. Zhang, M. Gong, Y. Wu, S. Wang, F. Liang, J. Lin, Y. Xu, H. Deng, H. Rong, C.-Z. Peng, M.-H. Yung, X. Zhu, and J.-W. Pan, Experimental simulation of larger quantum circuits with fewer superconducting qubits, Phys. Rev. Lett. **130**, 110601 (2023).

[16] A. Eddins, M. Motta, T. P. Gujarati, S. Bravyi, A. Mezzacapo, C. Hadfield, and S. Sheldon, Doubling the size of quantum simulators by entanglement forging, PRX Quantum **3**, 010309 (2022).

[17] C. Piveteau and D. Sutter, Circuit knitting with classical communication, IEEE Trans. Inf. Theory, 1 (2023).

[18] L. Brenner, C. Piveteau, and D. Sutter, Optimal wire cutting with classical communication, arXiv:2302.03366.

[19] K. Fujii, K. Mizuta, H. Ueda, K. Mitarai, W. Mizukami, and Y. O. Nakagawa, Deep variational quantum eigensolver: A divide-and-conquer method for solving a larger problem with smaller size quantum computers, PRX Quantum **3**, 010346 (2022).

[20] S. Endo, S. C. Benjamin, and Y. Li, Practical quantum error mitigation for near-future applications, Phys. Rev. X **8**, 031027 (2018).

[21] K. Mitarai and K. Fujii, Constructing a virtual two-qubit gate by sampling single-qubit operations, New J. Phys. **23**, 023021 (2021).

[22] T. Yamamoto and R. Ohira, Error suppression by a virtual two-qubit gate, J. Appl. Phys. **133**, 174401 (2023).

[23] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Randomized benchmarking and process tomography for gate errors in a solid-state qubit, Phys. Rev. Lett. **102**, 090502 (2009).

[24] S. Hakkaku, Y. Tashima, K. Mitarai, W. Mizukami, and K. Fujii, Quantifying fermionic nonlinearity of quantum circuits, Phys. Rev. Res. **4**, 043100 (2022).

[25] C. Song, J. Cui, H. Wang, J. Hao, H. Feng, and Y. Li, Quantum computation with universal error mitigation on a superconducting quantum processor, Sci. Adv. **5**, eaaw5686 (2019).

[26] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[27] S. Tamate, Y. Tabuchi, L. Szikszai, K. Kusuyama, K. Zuo, Z. Yan, A. Badrutdinov, Y. Hishida, W. Qiu, H. Terai, G. Fujii, K. Makise, N. Watanabe, H. Nakagawa, M. Fujino, M. Ukibe, W. Mizubayashi, K. Kikuchi, and Y. Nakamura, Scalable packaging design for large-scale superconducting quantum circuits, Bull. Am. Phys. Soc. **66**, C30.9 (2021).

[28] 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).

[29] E. Lucero, J. Kelly, R. C. Bialczak, M. Lenander, M. Mariantoni, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, T. Yamamoto, A. N. Cleland, and J. M. Martinis, Reduced phase error through optimized control of a superconducting qubit, Phys. Rev. A **82**, 042339 (2010).

[30] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple pulses for elimination of leakage in weakly nonlinear qubits, Phys. Rev. Lett. **103**, 110501 (2009).

[31] Z. Chen, J. Kelly, C. Quintana, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Lucero, E. Jeffrey, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner *et al.*, Measuring and suppressing quantum state leakage in a superconducting qubit, Phys. Rev. Lett. **116**, 020501 (2016).

[32] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White *et al.*, Optimal quantum control using randomized benchmarking, Phys. Rev. Lett. **112**, 240504 (2014).

[33] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Randomized benchmarking of quantum gates, Phys. Rev. A **77**, 012307 (2008).

[34] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, J. Opt. B: Quantum Semiclassical Opt. **7**, S347 (2005).

[35] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, Phys. Rev. Lett. **106**, 180504 (2011).

[36] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, Phys. Rev. A **69**, 062320 (2004).

[37] J. Y. Mutus, T. C. White, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, J. Kelly, A. Megrant, C. Neill, P. J. J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, K. M. Sundqvist, A. N. Cleland, and J. M. Martinis, Strong environmental coupling in a Josephson parametric amplifier, Appl. Phys. Lett. **104**, 263513 (2014).

[38] Y. Urade, K. Zuo, S. Baba, C. Chang, K.-i. Nittoh, K. Inomata, Z. Lin, T. Yamamoto, and Y. Nakamura, Flux-driven impedance-matched Josephson parametric amplifier with improved pump efficiency, Bull. Am. Phys. Soc. **66**, A28.10 (2021).

[39] D. T. McClure, H. Paik, L. S. Bishop, M. Steffen, J. M. Chow, and J. M. Gambetta, Rapid driven reset of a qubit readout resonator, Phys. Rev. Appl. **5**, 011001 (2016).

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

[41] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[42] K. Heya, Y. Suzuki, Y. Nakamura, and K. Fujii, Variational quantum gate optimization, arXiv:1810.12745.