PHOTONICS Research

Scalable fast benchmarking for individual quantum gates with local twirling

Yihong Zhang, Wenjun Yu, Pei Zeng, Guoding Liu, [iD] and Xiongfeng Ma* [iD]

Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China
*Corresponding author: xma@tsinghua.edu.cn

With the development of controllable quantum systems, fast and practical characterization of multi-qubit gates has become essential for building high-fidelity quantum computing devices. The usual way to fulfill this requirement via randomized benchmarking demands complicated implementation of numerous multi-qubit twirling gates. How to efficiently and reliably estimate the fidelity of a quantum process remains an open problem. This work thus proposes a character-cycle benchmarking protocol and a character-average benchmarking protocol using only local twirling gates to estimate the process fidelity of an individual multi-qubit operation. Our protocols were able to characterize a large class of quantum gates including and beyond the Clifford group via the local gauge transformation, which forms a universal gate set for quantum computing. We demonstrated numerically our protocols for a non-Clifford gate—controlled-(TX) and a Clifford gate—five-qubit quantum error-correcting encoding circuit. The numerical results show that our protocols can efficiently and reliably characterize the gate process fidelities. Compared with the cross-entropy benchmarking, the simulation results show that the character-average benchmarking achieves three orders of magnitude improvements in terms of sampling complexity.   © 2022 Chinese Laser Press

1. INTRODUCTION

Characterizing a quantum process has great importance in both the fundamental study and practical application of quantum information science. With the recent advent of noisy intermediate-scale quantum computing [1], benchmarking of quantum operations is critical for quantum control [2,3] as it provides an indicator for assessing experimental devices. Furthermore, it is essential for the development of high-precision quantum information processing instruments. Accurate benchmarking can reliably characterize the noise levels of quantum operations, and it plays a critical role in promoting fault-tolerant universal quantum computing [4,5]. In practice, we must evaluate the performance of a quantum circuit to verify whether a quantum algorithm or an error-correcting code is properly implemented in a quantum system.

Numerous approaches have been proposed to characterize quantum processes. Conventional methods such as quantum process tomography [6] provide a full description of a channel. However, these methods are impractical for large-scale quantum systems as the required experimental resources increase exponentially with the number of qubits, even with state-of-the-art techniques such as compressed sensing [7,8]. Direct fidelity estimation [9] tackles the scaling problem and characterizes the quantum process in terms of average fidelity.

Unfortunately, the result inevitably contains extra errors from the state preparation and measurement (SPAM) and hence often over-estimates the noise levels. In reality, SPAM errors usually grow rapidly with the system size so that it is hard to characterize the quantum process accurately for large-scale quantum systems with direct fidelity estimation.

Randomized benchmarking (RB) and its variants have been proposed to avoid simultaneously the scaling problem and SPAM errors [10–16]. Standard RB estimates the average error rate of a specific gate set under the assumption of gate-independent or weakly-dependent noise. The gate set is normally selected to be the Clifford group and has been widely implemented in experiments [17–24]. Otherwise, to characterize a specific Clifford gate, a variant called interleaved RB was proposed. It utilizes random Clifford gates interleaved with the target gate [25]. The random gates here are considered to be the twirling gates for reference, whose fidelity should be measured separately to infer the fidelity of the target gate. The interleaved RB method is efficient and scalable in principle. However, it suffers from two severe problems in practice. The first is the compiling overhead for twirling operations. In reality, any operation needs to be compiled to one- and two-qubit gates native to the quantum system. Note that twirling gates are randomly picked from a gate set, like the Clifford group. In general, the

average number of native gates used for compiling a single sample grows dramatically with the system size. The second is the gate-dependent noises introduced by twirling gates. Note that different twirling operations in a gate group can vary a lot in the depths of compiled circuits. For example, a local operation such as a Pauli gate can be implemented by a single layer circuit, while a complex entangling operation requires a deep circuit with massive native gates. The strong gate-dependent noises caused by the uneven compilations may bring inaccuracy to the fidelity estimation [26,27]. Consequently, the compiling overhead and gate-dependent noises introduced by twirling gates limit the scalability of the RB method in experiments.

Recently, there have been several variants of RB attempting to address the two compiling problems. For example, character benchmarking employs the character theory so that the quality parameters can be extracted from the local twirling operations [28]. Unfortunately, for the gate groups with an exponentially increasing number of quality parameters, this method requires an exponential amount of SPAM settings. Besides, character benchmarking is still caught in the aforementioned compiling problems for the final inverse gate and can be hardly applied for a generic multi-qubit quantum operation. Another inspiring attempt called cycle benchmarking aims to estimate the fidelity of the target gate by interleaving it with the Pauli gate set. However, it is restricted to the Clifford gates [29]. In addition, cycle benchmarking requires numerous repetitions for the gates with large cyclic numbers, which is common for multi-qubit gates. Hence, this method cannot efficiently benchmark a wide class of gates. The cross-entropy benchmarking (XEB) characterizes the fidelity of a generic quantum gate reflected by linear cross-entropy using local Clifford gate twirling [30]. However, the Haar measure assumption in XEB may lead to poor fidelity estimation when the size of the target gate is large. Thus, how to efficiently and reliably estimate the fidelity of a large-scale quantum process from a universal gate set remains an open problem.

This work proposes two scalable and efficient protocols to tackle the compiling problems as well as the SPAM error issues simultaneously, which we label character-cycle benchmarking (CCB) and character-average benchmarking (CAB), respectively. The protocols utilize local twirling to reliably characterize the fidelity of an individual multi-qubit quantum operation. We employ the Pauli and the local Clifford gates for twirling and extend the applicable gate set to non-Clifford gates via the local gauge transformation. The efficiency and reliability of the protocols are shown by rigorous mathematical derivations and by numerical simulations under realistic physical assumptions.

2. CHARACTER CYCLE BENCHMARKING

Let us denote the quantum operation of a unitary matrix U acting on an n-qubit quantum state, ρ, by the calligraphic letter $\mathcal{U}$, i.e., $\mathcal{U}(\rho) = U\rho U^{-1}$, and the noisy implementation by $\tilde{\mathcal{U}}$. One can evaluate the quality of $\tilde{\mathcal{U}}$ by the process fidelity of the noise channel $\Lambda = \mathcal{U}^{-1} \circ \tilde{\mathcal{U}}$,

$$F(\Lambda) = \frac{1}{d^2} \sum_{i=0}^{4^n-1} \lambda_i, \tag{1}$$

where $\lambda_i = d^{-1}\mathrm{Tr}(P_i\Lambda(P_i))$ is the Pauli fidelity associated with the Pauli operator $P_i \in \mathsf{P}_n$, and $d = 2^n$ is the dimension of the quantum system. Here, P_n denotes the n-qubit Pauli group, containing the tensor product of the identity operation I and three Pauli matrices X, Y, Z.

In practice, it is costly to figure out all the parameters $\lambda_0, \lambda_1, \cdots, \lambda_{4^n-1}$ as their number increases exponentially with n. Instead, one can estimate the process fidelity via repeatable sampling of $\lambda_0, \lambda_1, \cdots, \lambda_{4^n-1}$. Concretely, one samples a sufficient number of Pauli operators $\{P_j\}$ and averages the corresponding $\{\lambda_j\}$,

$$F(\Lambda) \approx \frac{1}{M} \sum_{\{P_j\}} \lambda_j, \tag{2}$$

where M is the number of samples and the summation takes over the sample set.

Here, we propose a CCB protocol that employs the key techniques of the cycle [29] and character benchmarking [28]. Specifically, we extract different Pauli fidelities through applying specific initial states and measurements and utilize the character theory to fully separate the SPAM errors. The schematic circuit of the CCB protocol is shown in Fig. 1(a). Let us start with the Clifford case, where the target gate belongs to the n-qubit Clifford group C_n. The inner random gate layer consists of the target gate $\mathcal{U}$ and its inverse gate $\mathcal{U}^{-1}$, interleaved with two random Pauli gate layers. The Pauli gates are the reference gates employed to perform local Pauli twirling over the generic quantum noise channel Λ and turn it into

$$\Lambda_\mathsf{P} = \frac{1}{4^n} \sum_{\{P_j \in \mathsf{P}_n\}} \mathcal{P}_j^{-1} \circ \Lambda \circ \mathcal{P}_j, \tag{3}$$

where Λ contains the errors of Pauli gates and target gate $\mathcal{U}$; Λ_P is a Pauli channel satisfying $\Lambda_P(\rho) = \sum_j p_j P_j \rho P_j$, and p_j is the Pauli error rate related to P_j.

Note that the introduction of the inverse target gate $\mathcal{U}^{-1}$ is the major difference between CCB and cycle benchmarking. In cycle benchmarking, we need to repeat $\mathcal{U}$ for multiples of l times, where l is the cyclic number of U, i.e., $\mathcal{U}^l = I$. In general, l can be quite large for a wide class of Clifford gates, which is prohibitive for the experiments. For example, the five-qubit quantum error-correcting encoding circuit requires $l = 124$. The CCB protocol improves the efficiency and application scope via substituting a single $\mathcal{U}^{-1}$ for multiple repetitions of $\mathcal{U}$ in cycle benchmarking. In many quantum platforms, such as superconducting quantum processors, the inverse gates of the native gates are also native. Typical examples include single-qubit gates, CZ, and iSWAP. Thus, the inverse gates of native gates are normally easy to implement. More generally, if $\mathcal{U}$ is composed of several native gates, the difficulty to implement $\mathcal{U}$ and $\mathcal{U}^{-1}$ is often the same. Based on this consideration, the introduction of $\mathcal{U}^{-1}$ does not increase the implementation difficulty of CCB in most cases.

In CCB, the randomization of Pauli gates in the inner gate layers will generate a composite channel $\mathcal{U}^{-1} \circ \Lambda_\mathsf{P}^{(-)} \circ \mathcal{U} \circ \Lambda_\mathsf{P}$, where Λ_P and $\Lambda_\mathsf{P}^{(-)}$ are the Pauli-twirled channels corresponding to gates $\mathcal{U}$ and $\mathcal{U}^{-1}$, respectively. Note that this composite channel is a Pauli channel for Clifford gate $\mathcal{U}$. The fidelity we aim to estimate in the CCB protocol is defined as the CCB fidelity,
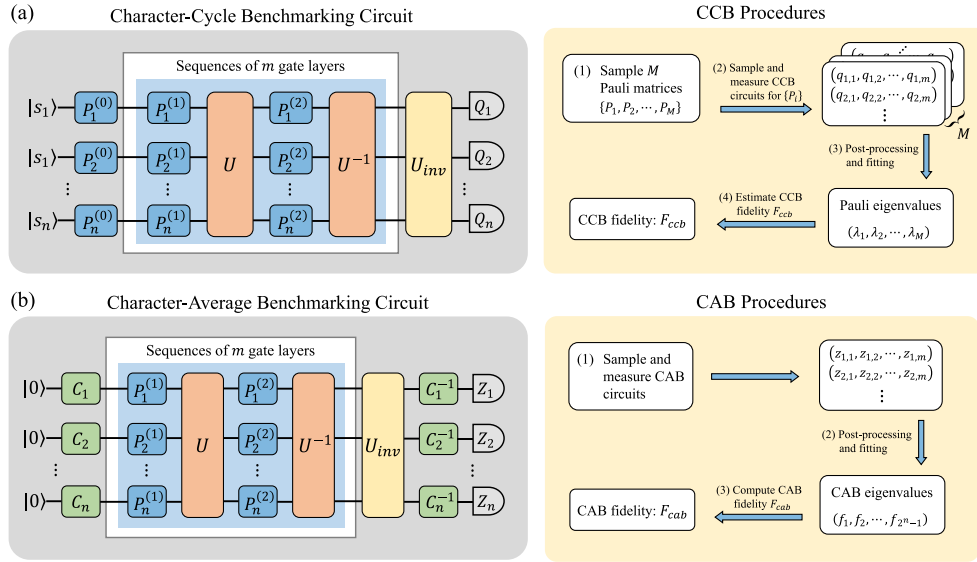
Fig. 1. Illustrations of circuit and procedures used in (a) CCB and (b) CAB protocols. The orange boxes represent the target gate $\mathcal{U}$ and its inverse gate $\mathcal{U}^{-1}$. The blue and green boxes represent the random Pauli gate and random local Clifford gate. The yellow boxes denote the inverse gate for the m inner gate layers in the light blue box. Here, $P_k^{(i)}$ is a single-qubit Pauli gate on qubit k, and $P^{(i)} = P_1^{(i)} \otimes \cdots \otimes P_n^{(i)}$ is an n-qubit Pauli gate.

$$F_{\text{ccb}} = F(\sqrt{\mathcal{U}^{-1} \circ \Lambda_{\mathsf{P}}^{(-)} \circ \mathcal{U} \circ \Lambda_{\mathsf{P}}}), \qquad (4)$$

which contains the fidelities of $\mathcal{U}$ and $\mathcal{U}^{-1}$. When the noise channel of $\mathcal{U}$ is the same as that of $\mathcal{U}^{-1}$, which is valid for most of the experimental platforms, the CCB fidelity is simplified as

$$F_{\text{ccb}}(\Lambda) = F(\sqrt{\mathcal{U}^{-1} \circ \Lambda_{\mathsf{P}} \circ \mathcal{U} \circ \Lambda_{\mathsf{P}}}). \qquad (5)$$

Equation (5) is a lower bound of the process fidelity $F(\Lambda)$ in terms of the expectation value, as proved in Appendix B.2. In the following context, we will employ Eq. (5) as our CCB fidelity metric model and our arguments apply to the general model of Eq. (4) as well. The difference between F_{ccb} and F is normally small because the physical realizations of the qubits in one experimental platform are similar and the qualities of these qubits will not differ too much. Note that if Λ_{P} is a depolarizing channel, then $F_{\text{ccb}} = F$. Thus, the CCB fidelity can be seen as a reliable metric for the noise channel Λ.

The procedure of the CCB protocol is as follows.

1. Sample a Pauli operator P_j and initialize state $|s\rangle$ such that $P_j|s\rangle = |s\rangle$.
2. Apply a gate sequence $\mathcal{S}_{\text{ccb}}$ composed of a Pauli gate $\mathcal{P}^{(0)}$, m inner gate layers denoted by $\mathcal{S}_m$, and inverse gate $\mathcal{S}_m^{-1}$.
3. Perform measurement P_j and then calculate the P_j-weighted survival probability $f_j(m, \mathcal{S}_{\text{ccb}}) = \chi_j(P^{(0)}) \cdot \text{Tr}(P_j \mathcal{S}_{\text{ccb}}(\rho_s))$, where $\chi_j(P^{(0)}) = 1$ if P_j commutes with $P^{(0)}$ and -1 otherwise.
4. Repeat steps (2) and (3) for several times for different m and fit the P_j-weighted fidelity to $\hat{f}_j(m) = A_j \lambda_j^{2m}$.
5. Repeat steps (1)–(4) for several times and finally estimate the CCB fidelity as $F_{\text{ccb}} = \text{ave}_j \lambda_j$.

Here, the estimated fidelity F_{ccb} includes the errors from the local reference gate set P_n. To remove these extra errors, one can employ the interleaved RB technique, by performing additional CCB with a target gate of identity I to estimate the

reference fidelity F_{ccb}^I. Then, one can infer the fidelity of the target gate as $F_{\text{ccb}}/F_{\text{ccb}}^I$. In practice, the errors of local gates are often negligible, and hence, we focus on F_{ccb} in the following discussions.

Note that our inverse gate $\mathcal{S}_m^{-1}$ is a Pauli gate and thus will not introduce extra gate compiling overhead. In contrast, character benchmarking for a single multi-qubit Clifford gate [28] requires a global inverse gate and a complicated compiling process. This may cause strong gate-dependent errors and lead to inaccuracy for fidelity estimation, especially for multi-qubit quantum operations. The CCB protocol maintains the local structure of reference and inverse gates and thus avoids the compiling problems.

In the CCB protocol, one needs to average Pauli fidelities λ_j to estimate F_{ccb}. The sampling complexity for the CCB protocol is given by the following theorem.

Theorem 1 (informal version). *For an n-qubit quantum noise channel, to estimate the CCB fidelity within the confidence interval $[\hat{F}_{\text{ccb}} - \epsilon_M - \epsilon_b, \hat{F}_{\text{ccb}} + \epsilon_M + \epsilon_b]$ with probability greater than $1 - \delta$, one needs to sample M Pauli fidelities where each Pauli fidelity is estimated via K random sequences. The confidence probability of the estimation is given by*

$$\Pr(|\hat{F}_{\text{ccb}} - \bar{F}_{\text{ccb}}| \le \epsilon_M + \epsilon_b) \ge 1 - \delta, \qquad (6)$$

where $\epsilon_M \le \mathcal{O}(\frac{-\log \delta}{M})$ and $\epsilon_b \le \mathcal{O}(K^{-1}) + \mathcal{O}((\frac{-\log \delta}{K})^{3/2})$.

Here, the total number of samples, or sample complexity, depends on M and K. If the number of random sequences for each Pauli fidelity is the same, then the sample complexity is simply given by MK. Theorem 1 shows that the sample complexity only depends on fidelity precision ϵ_M, ϵ_b, and confidence level δ. The independence on system size n reflects the strong scalability of the CCB protocol. A more detailed description of the result is shown in Theorem 2.

3. LOCAL GAUGE TRANSFORMATION

Now, let us extend the applicable gates for the CCB protocol to non-Clifford gates. One can introduce local gauge transformation L to the twirling gate set, $\mathsf{P}_n \to L\mathsf{P}_n L^{-1}$, where L is an arbitrary local unitary operation $L \in \mathsf{U}(2)^{\otimes n}$. Note that the transformed twirling gate set $L\mathsf{P}_n L^{-1}$ is still local. Then, we can show that the applicable target gate set becomes $L\mathsf{C}_n L^{-1}$, where C_n is the n-qubit Clifford gate set.

To benchmark a gate LUL^{-1} from gate group $L\mathsf{C}_n L^{-1}$, we insert local gates L and L^{-1} between the twirling gates and the target gates in the original CCB circuit, as shown in Fig. 2(a). Here, $L = \otimes_{i=1}^n L_i$, where L_i can be an arbitrary single-qubit gate. In practice, the local gates are absorbed into twirling gates and target gates and do not need to be implemented individually as manifested in Fig. 2(b). The character gate $LP^{(0)}$ and the twirling gate $LP^{(1)}L^{-1}$ will be merged into a single gate in implementation as well. Details of the derivation are shown in Appendix B.3.

As shown in Fig. 2, the CCB circuit with local gauge transformation L and noise channel Λ is equivalent to the original CCB circuit with noise channel $\mathcal{L}^{-1}\Lambda\mathcal{L}$. Thus, one can obtain $F_{\text{ccb}}(\mathcal{L}^{-1}\Lambda\mathcal{L})$, which is close to the process fidelity $F(\mathcal{L}^{-1}\Lambda\mathcal{L})$. As process fidelity is gauge-invariant, that is, $F(\mathcal{L}^{-1}\Lambda\mathcal{L}) = F(\Lambda)$, one can estimate the process fidelity of Λ as the performance indicator of gate LUL^{-1}.

Now, let us check out what kinds of quantum gates belong to the set $\mathcal{S} = \{LUL^{-1}|L \in U(2)^{\otimes n}, U \in \mathsf{C}_n\}$.

First, notice that if a unitary $U \in \mathcal{S}$, then for any $L \in U(2)^{\otimes n}$, $LUL^{-1} \in \mathcal{S}$. As any unitary is generated by a Hamiltonian, that is $U = e^{iH}$ where H is Hermitian, one

can conclude that if $e^{iH} \in \mathcal{S}$, then for any $L \in U(2)^{\otimes n}$, $Le^{iH}L^{-1} = e^{iLHL^{-1}} \in \mathcal{S}$.

Take a step forward. If a controlled-$e^{iH} \in \mathcal{S}$, then through local gauge transformation $I \otimes L$, $(I \otimes L)$ controlled-$e^{iH}(I \otimes L)^{-1} = $ controlled-$e^{iLHL^{-1}} \in \mathcal{S}$. The arguments also apply to the case of multi-controlled gates.

The two observations inspire us to first represent Clifford gates in the form of e^{iH} or multiple controlled-e^{iH}, then replace H with LHL^{-1} to find other gates in $\mathcal{S}$. Take CZ as an example. $CZ = e^{i\pi|11\rangle\langle11|} = $ controlled-$e^{-i\frac{\pi}{2}Z}$. Through local gauge transformation, one can transform $|11\rangle$ to any product state $|\psi\phi\rangle$ and transform $e^{-i\frac{\pi}{2}Z}$ to any π-rotation $e^{-i\frac{\pi}{2}\vec{\sigma}\cdot\vec{\theta}}$, where $\vec{\sigma} = (X, Y, Z)$ and $\vec{\theta}$ is a unit vector. Thus, for any two-qubit product state $|\psi\phi\rangle$, we have $e^{-i\pi|\psi\phi\rangle\langle\psi\phi|} \in \mathcal{S}$. In addition, any controlled-π rotation, such as controlled-H and controlled-TX, belongs to $\mathcal{S}$.

Reversely, controlled-$S = $ controlled-$e^{-i\frac{\pi}{4}Z}$ is a controlled-$\frac{\pi}{2}$ rotation. As any controlled-$\frac{\pi}{2}$ rotation is not Clifford, one can conclude that controlled-S does not belong to $\mathcal{S}$. Similarly, Tofolli $= $ controlled-controlled-$e^{-i\frac{\pi}{2}X}$ is a controlled-controlled-π rotation. As any controlled-controlled-π rotation is not Clifford, one can conclude that Tofolli does not belong to $\mathcal{S}$ either. It is interesting to decide whether a quantum gate belongs to $\mathcal{S}$ in a more general case, and we leave it for future work.

4. CHARACTER-AVERAGE BENCHMARKING

We can take the CCB protocol one step further. Observe that in CCB, one needs to implement the fitting procedures for each sampled Pauli operator P_j to estimate Pauli fidelity λ_j.
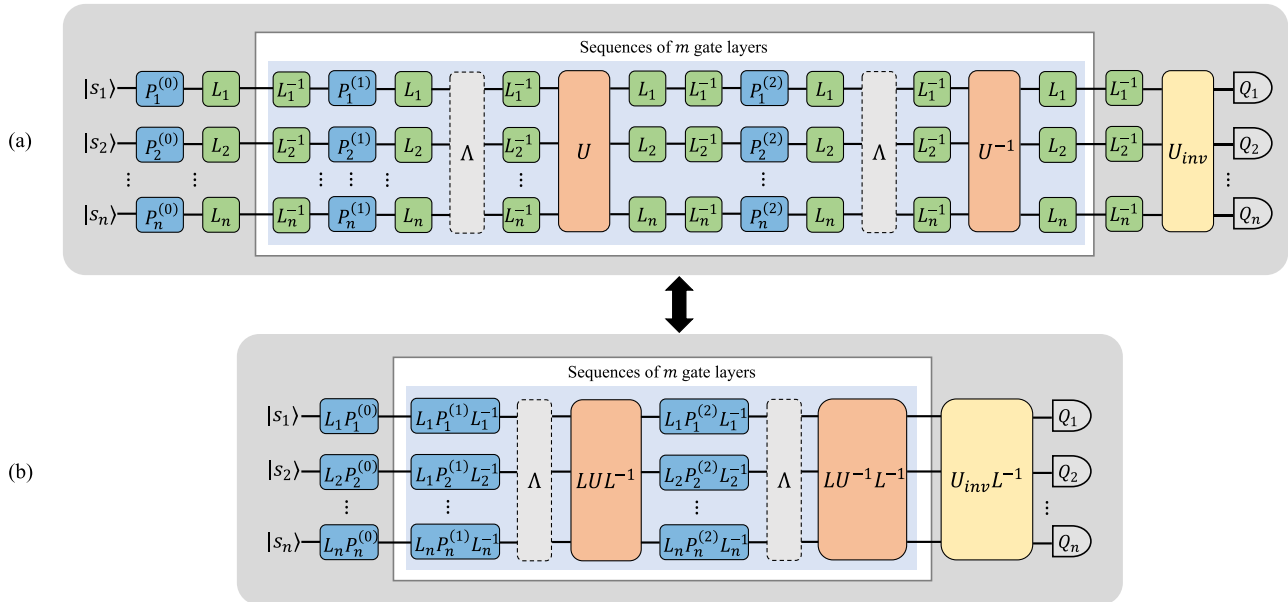
Fig. 2. Illustrations of the noisy CCB circuit with local gauge transformation. For simplicity, we show the case that the local gates are noiseless. The grey dashed boxes denote the noise channel Λ. The orange boxes represent the target gate $\mathcal{U}$ and its inverse gate $\mathcal{U}^{-1}$. The blue boxes represent the random Pauli gates. The green boxes represent the inserted local gates L and L^{-1}, where $L = \otimes_{i=1}^n L_i$. The yellow box denotes the inverse gate for the m inner gate layers in the light blue boxes. In practice, we implement gates in circuit (b) while absorbing local gates L and L^{-1} into twirling and target gates. Here, the target gate following gauge transformation becomes LUL^{-1}. Note that circuit (a) is equivalent to a CCB circuit with target gate U and noise channel $\mathcal{L}^{-1}\Lambda\mathcal{L}$. Thus, it can be implemented to estimate $F_{\text{ccb}}(\mathcal{L}^{-1}\Lambda\mathcal{L})$, which is close to $F(\Lambda)$.

Each estimation requires specific initial state, measurement, and independent randomization procedures. We can further simplify these procedures by introducing the local Clifford group $\mathsf{C}_1^{\otimes n}$. Recall that in a qubit system, the Clifford twirling depolarizes a channel via averaging the error rates in X, Y, Z bases [13]. Then for an n-qubit system, the twirling over $\mathsf{C}_1^{\otimes n}$ would partially depolarize a channel and average out 4^n Pauli fidelities λ_j into 2^n terms. These 2^n values can be obtained from the $Z^{\otimes n}$ basis measurement only with additional data post-processing.

Based on the local Clifford twirling, we propose the CAB protocol as an improvement of the CCB protocol. The schematic circuit of CAB is shown in Fig. 1(b), with the detailed procedures described in Box 1. Like the CCB protocol, we can extend the target gate set beyond the Clifford group by employing local gauge transformation. Here, to suppress statistical fluctuations, we remove the character technique. Detailed description and analysis of the CCB and CAB protocols are presented in Appendix B.

Similar to the CCB protocol, the randomization over m gate layers inside the blue box in Fig. 1(b) will generate a Pauli channel $\Lambda_{\mathsf{P}}(m) = (\mathcal{U}^{-1} \circ \Lambda_{\mathsf{P}} \circ \mathcal{U} \circ \Lambda_{\mathsf{P}})^m$. The local Clifford gates in the beginning and end of the circuit jointly perform local unitary 2-design twirling, which transforms the Pauli channel $\Lambda_{\mathsf{P}}(m)$ into a partially depolarizing channel $\Lambda_{\mathsf{C}}(m)$. Here, the quantum channel $\Lambda_{\mathsf{C}}(m)$ contains less independent parameters than the original $\Lambda_{\mathsf{P}}(m)$. It holds the unique value of fidelity for every disjoint Pauli subset in $R^n = \{\{I\}, \{X, Y, Z\}\}^{\otimes n}$. The Pauli fidelities in $\Lambda_{\mathsf{C}}(m)$ can be seen as the average values of those in $\Lambda_{\mathsf{P}}(m)$, $\mu_k^{2m} = \sum_{P_j \in \sigma_k} \lambda_j^{2m}/|\sigma_k|$, where $\{\lambda_j^{2m}\}$ are the Pauli fidelities of the channel $\Lambda_{\mathsf{P}}(m)$. The local Clifford twirling here averages multi exponential decays into one

exponential decay and captures all the information of the noise channel, as shown in Eq. (11). The comparison between F_{cab} and F_{ccb} is shown in Lemma 3 in Appendix B.5, while the CCB protocol employs a sampling method as in Eq. (2), which only contains partial information of the noise channel. Thus, one can intuitively conclude that the CAB protocol is more efficient than the CCB protocol, as demonstrated in later simulations.

5. SIMULATION

In numerical simulations, we characterize a two-qubit controlled-(TX) gate and a five-qubit quantum error correcting encoding circuit, respectively. Furthermore, we simulate the noise channel for the target gate with a realistic error model that contains a Pauli channel, an amplitude damping channel, and a correlation channel. In the simulation, the Pauli fidelities of the Pauli channel are randomly sampled from a normal distribution $\mathcal{N}(\mu, \sigma)$, which we call the $\mathcal{N}(\mu, \sigma)$-Pauli channel. Here, the error parameter μ reflects the quality of the Pauli channel and σ implies the discrepancy of the channel, i.e., the differences among Pauli fidelities. The detailed descriptions for the error models and simulations are presented in Appendix D.

For the controlled-(TX) gate, we take $(I \otimes \sqrt{T}) \cdot \mathsf{P}_n(I \otimes \sqrt{T^{-1}})$ as the twirling gate set, where T is the $\pi/8$-phase gate, $T = \exp(-i\pi Z/8)$, and $I \otimes \sqrt{T}$ is the local gauge transformation. We simulate the CAB and CCB protocols on the controlled-(TX) gate with 8 different noise channels. For each noise channel, we take 40 independent simulations for both CAB and CCB protocols. In CCB simulations, we sample $M = 10$ Pauli operators to estimate F_{ccb}.

Box 1. Procedures for character-average benchmarking

    1. Sample a gate sequence $(C, P^{(1)}, P^{(2)}, \cdots, P^{(2m)})$, where C and $P^{(i)}$ $(1 \le i \le 2m)$ are sampled uniformly at random from the local groups, $\mathsf{C}_1^{\otimes n}, \mathsf{P}_n$, respectively.

    2. Initialize the state $|\psi\rangle = |0\rangle^{\otimes n}$ and apply the gate sequence as shown in Fig. 1(b),

$$\mathcal{S}_{\mathrm{cab}} = \mathcal{C}^{-1} \circ \mathcal{U}_{\mathrm{inv}} \circ \mathcal{U}^{-1} \circ \mathcal{P}^{(2m)} \circ \ldots \circ \mathcal{U} \circ \mathcal{P}^{(1)} \circ \mathcal{C}, \tag{7}$$

where the inverse gate $\mathcal{U}_{\mathrm{inv}} = \mathcal{P}^{(1)} \circ \mathcal{U}^{-1} \circ \ldots \circ \mathcal{P}^{(2m)} \circ \mathcal{U}$ is a local gate as well.

    3. Measure in $Z^{\otimes n}$ basis and compute the survival probability for each measurement observable $Q_k \in \{I, Z\}^{\otimes n}$,

$$f_k(m, \mathcal{S}_{\mathrm{cab}}) = \mathrm{Tr}(Q_k \mathcal{S}_{\mathrm{cab}}(\rho_\psi)), \tag{8}$$

where ρ_ψ is the noisy preparation of the initial state $|\psi\rangle$.

    4. Repeat for a sufficient number of sequences and estimate the average value

$$f_k(m) = \mathop{\mathbb{E}}_{\mathcal{S}_{\mathrm{cab}}} f_k(m, \mathcal{S}_{\mathrm{cab}}). \tag{9}$$

    5. Repeat for different m and fit to the function
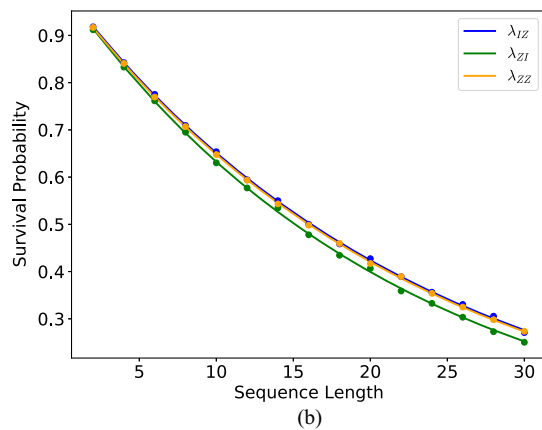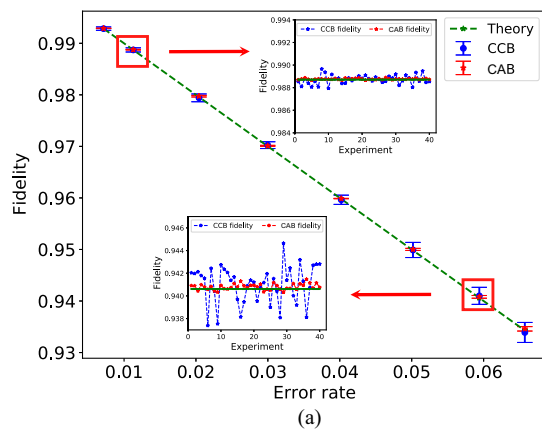
$$f_k(m) = A_k \mu_k^{2m}, \tag{10}$$

where A_k and μ_k are fitting parameters.

    6. Estimate the CAB fidelity

$$F_{\mathrm{cab}} = \frac{1}{d^2} \sum_k d_k \mu_k, \tag{11}$$

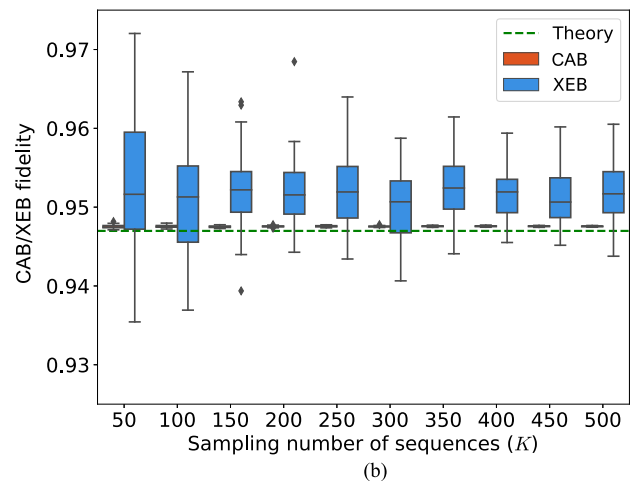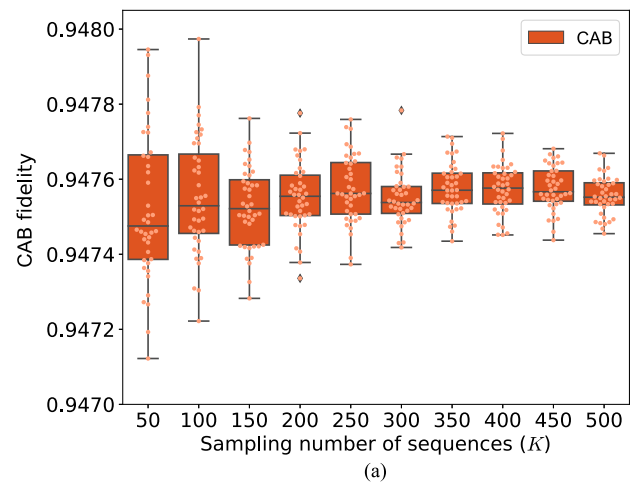where $d_k = 3^{\pi(Q_k)}$ and $\pi(Q_k)$ is the number of Z in Q_k.

Figure 3(a) shows F_{cab} and F_{ccb} versus the error rate $r = 1 - F$ for the controlled-(TX) gate with different noise channels. We observe that when the standard deviation of error parameters σ grows, the error bars of F_{cab} and F_{ccb} become larger. Intuitively, the discrepancy of the Pauli fidelities is one of the key reasons for the fluctuations of F_{cab} and F_{ccb}. The fluctuations for the estimations will reach the minimum level when the noise channel is completely depolarizing. Besides, the error bar of CAB is smaller than the error bar of CCB. This shows that under the same estimation accuracy, the sampling complexity, i.e., the amount of sampling sequences in total, of the CAB protocol is smaller than that of the CCB protocol, especially when the discrepancy of the noise

channel is large. In Fig. 3(b), we take one of the eight noise channels as an example and show the three fitting curves of Eq. (10) for the CAB protocol. The resulting CAB fidelity is $F_{cab} = 95.99\%$, which is very close to the theoretical value of process fidelity $F = 95.98\%$.

For the five-qubit error correcting encoding circuit, which is a Clifford gate, we take the Pauli group P_n as the twirling gate set. For the simplicity of simulation, we set the Pauli channel to be a depolarizing channel $\Lambda_{dep}(\rho) = p\rho + (1-p)I/d$ where $p = 0.98$. The setting of the amplitude damping and correlation channels remains the same. Furthermore, we simulate the CAB and XEB protocols to characterize the noisy five-qubit encoding circuit. For each protocol, we run 40 independent

Fig. 3. Simulation results for the controlled-(TX) gate with eight different noise channels, each containing a $\mathcal{N}(\mu, \sigma)$-Pauli channel, an amplitude damping channel, and a correlation channel. The parameters for the amplitude damping channels and correlation channels are set to be the same for the eight channels. For the Pauli channel, the error parameters are set to $\{(\mu, \sigma)\} = \{(0.995, 0.001), (0.990, 0.002), (0.980, 0.003), (0.970, 0.004), (0.960, 0.005), (0.950, 0.006), (0.940, 0.007), (0.930, 0.008)\}$. (a) The fidelity estimations with different error rates in 40 independent simulations. The green dashed line represents the theoretical fidelities. The two insert scatter plots show the fluctuations of estimated fidelities over different simulations. (b) Take the fifth simulation with a process fidelity of 95.98% as an example. The fitting curves of Eq. (10) for $Q_k = IZ, ZI, ZZ$. The decay parameters derived from the curves are $\lambda_{IZ} = 0.9580$, $\lambda_{ZI} = 0.9550$, $\lambda_{ZZ} = 0.9577$.

Fig. 4. Simulation results for the five-qubit quantum error correcting encoding circuit with a noise channel composed of a depolarizing channel $\Lambda_{dep}(\rho) = p\rho + (1-p)I/d$, where $p = 0.98$, an amplitude damping channel, and a correlation channel. The theoretical process fidelity is $F = 94.70\%$. (a) Box plot of the CAB fidelities versus sampling number K with 40 independent simulations. The red boxes represent the distributions of the CAB fidelity estimations with respect to K. The orange points represent the distributions of CAB fidelities. (b) Box plots of the CAB and XEB fidelities versus K with 40 independent simulations. The green dashed line represents the theoretical process fidelity. The plots of CAB in (a) and (b) are the same with different scaling.

simulations. In each simulation, we take the sampling number of gate sequences as $K = 50, \cdots, 500$ for each sequence length m. The box plot of F_{cab} versus K is shown in Fig. 4(a). We can see that when K grows, the fluctuations of F_{cab} become smaller. When K is not too large, like $K = 50$, the fluctuation is already small enough, which implies that the CAB protocol works well with few sampling sequences needed.

In Fig. 4(b), we show the box plots of F_{cab} and XEB fidelities F_{xeb} versus the sampling number K. It is clear to see that compared with F_{xeb}, F_{cab} is much closer to the theoretical process fidelity $F = 94.7\%$. Meanwhile, the convergence of F_{cab} is much better than that of F_{xeb}. This implies that the required K for CAB is much smaller than that of XEB under the same estimation accuracy.

To give a concrete example, we take 20 CAB simulations and 20 XEB simulations under the same noise channel. From the simulation results, we find that for CAB, when $K_{cab} = 20$, the standard deviation over the 20 simulations is $\sigma_{cab} = 3.25 \times 10^{-4}$, while for XEB, when $K_{xeb} = 20,000$, the standard deviation is $\sigma_{xeb} = 4.29 \times 10^{-4}$. This shows that to estimate the fidelities with standard deviations around 4×10^{-4}, the required K_{xeb} is over 1000 times larger than K_{cab}. Thus, we can conclude that the performance of CAB protocol is three orders of magnitude better than that of XEB protocol in terms of the sampling complexity.

The simulation results reveal the strong scalability and reliability of our protocols, especially the CAB protocol. The fluctuation of estimated CAB fidelity is small even for multi-qubit gates. We believe the CAB protocols can provide fast feedback in experimental designs and promote the development of universal fault-tolerant quantum computing.

6. CONCLUSION AND DISCUSSION

The characterization of large-scale individual quantum processes is crucial to the development of near-term quantum devices. However, to the best of our knowledge, there are no scalable and practical methods as of yet that could benchmark multi-qubit universal gate-set currently. In this work, we propose and demonstrate efficient and scalable randomized benchmarking protocols—CCB and CAB that can individually characterize a wide class of quantum gates including and beyond the Clifford set. The key technique of our protocols is using the local reference gate-set for twirling, which avoid the inaccuracy of the estimation caused by gate-compiling overhead and gate-dependent noises. The method of local gauge transformation offers a tool for characterizing non-Clifford gates. The sampling and measurement complexity are independent of the qubit number of gate, which means our benchmarking protocols can be generalized to large-scale quantum systems.

Our protocols maintain the simplicity and robustness of the conventional RB method, and estimate the quantity of most interest—process fidelity of the target gates. We believe our protocols will promote the development of universal fault-tolerant quantum computing. Furthermore, it would also be interesting to extend our randomization and estimation methods for characterizing other properties such as unitarity and coherence, which we leave for future research.

APPENDIX A: PRELIMINARIES

1. Representation Theory

The representation theory works as a general analysis of every representation for abstract groups. Informally, the representations of a group can reflect its block-diagonal structures. Let G be a finite group and $g \in G$ be a group element. The representation of G is defined as follows.

Definition 1 (Group representation). *Map ϕ is said to be a representation of group G on a linear space V if it is a group homomorphism from G to $GL(V)$:*

$$\phi : G \to GL(V),$$
$$g \mapsto \phi(g), \forall g \in G, \tag{A1}$$

where $GL(V)$ is the general linear group of V, such that $\forall g_1, g_2 \in G$:

$$\phi(g_1)\phi(g_2) = \phi(g_1 g_2). \tag{A2}$$

Given representation ϕ on V, a linear subspace $W \subseteq V$ is called *invariant* if $\forall w \in W$ and $\forall g \in G$:

$$\phi(g)w \in W. \tag{A3}$$

The restriction of ϕ to the invariant subspace W is known as a *subrepresentation* of G on W. One can further define the *irreducible representation* (or irrep for short) as follows.

Definition 2 (Irreducible representation). *Representation ϕ of group G on linear space V is irreducible if it merely has trivial subrepresentations, i.e., the invariant subspaces for V are only $\{0\}$ and V itself.*

The Maschke's theorem provides an interesting property that each representation ϕ of a finite group G can be decomposed to the irreducible representations, $\forall g \in G$:

$$\phi(g) \simeq \bigoplus_{\sigma \in R_G} \sigma(g)^{m_\sigma}, \tag{A4}$$

where $R_G = \{\sigma\}$ denotes the set of all the irreps of representation ϕ and m_σ is the *multiplicity* of the equivalent irreps of σ. In this paper, we will focus on the non-degenerate representation case, i.e., $m_\sigma = 1, \forall \sigma$.

Definition 3 (Character function). *Let σ be a representation over group G, and the character of σ is the function $\chi_\sigma : G \to \mathbb{C}$ given by $\forall g \in G$:*

$$\chi_\sigma(g) = \text{Tr}[\sigma(g)]. \tag{A5}$$

With the character function, we introduce the *generalized projection formula* used in character randomized benchmarking [28].

Lemma 1 (Generalized projection formula [31]). *Given a finite group, G, and its representation, ϕ, denote σ to be an irreducible representation contained in ϕ with its character function $\chi_\sigma : G \to \mathbb{C}$. The projector onto the support space of σ can be written as*

$$\Pi_\sigma = \frac{d_\sigma}{|G|} \sum_{g \in G} \chi_\sigma(g)\phi(g), \tag{A6}$$

where d_σ is the dimension of σ, $d_\sigma = \chi(e)$ with e being the identity element in G.

Next, we will introduce twirling over a group G.

Definition 4 (Twirling). *For representation ϕ of group G on linear space V, a random twirling for a linear map $\Lambda : V \to V$ over G is defined as*

$$\Lambda_{\mathsf{G}} = \frac{1}{|\mathsf{G}|} \sum_{g \in \mathsf{G}} \phi(g)^{\dagger} \Lambda \phi(g). \tag{A7}$$

Using Schur's Lemma [31], one can show the following proposition.

Proposition 1. *For any linear map $\Lambda : V \to V$, the twirling over group G and its representation ϕ can be written as*

$$\Lambda_{\mathsf{G}} = \sum_{\sigma \in \mathcal{R}_{\mathsf{G}}} \frac{\mathrm{Tr}(\Lambda \Pi_{\sigma})}{\mathrm{Tr}(\Pi_{\sigma})} \Pi_{\sigma}, \tag{A8}$$

where Π_{σ} denotes the projector onto the support space of σ, and $\mathcal{R}_{\mathsf{G}}$ denotes the set of all irreps of ϕ.

2. Representation for Quantum Channel

Here, we introduce the quantum channel and three frequently used channel representations that our main results rely on. Denote the Hilbert space for n qubits as $\mathcal{H}$ and the set of linear operators on $\mathcal{H}$ as $\mathcal{L}(\mathcal{H})$. Quantum channels are defined as completely positive and trace-preserving (CPTP) linear maps on $\mathcal{L}(\mathcal{H})$. Given any quantum channel $\mathcal{E} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$, we can represent it in *Kraus representation*, $\forall O \in \mathcal{L}(\mathcal{H})$:

$$\mathcal{E}(O) = \sum_{l=1}^{m} K_l O K_l^{\dagger}, \tag{A9}$$

where $\{K_l\}$ are the Kraus operators satisfying

$$\sum_{l=1}^{m} K_l^{\dagger} K_l = I. \tag{A10}$$

With the Kraus representation, the concatenation of the quantum channels or quantum gates is given by

$$\mathcal{E}_2 \circ \mathcal{E}_1(O) = \sum_{l_2=1}^{m_2} \sum_{l_1=1}^{m_1} K_{l_2} K_{l_1} O K_{l_1}^{\dagger} K_{l_2}^{\dagger}, \tag{A11}$$

where $\{K_1\}$ and $\{K_2\}$ are the Kraus operators for $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively.

To describe a long quantum circuit, the Kraus representation is not convenient. Here, we introduce another widely used representation—*Liouville representation*. The Liouville representation is defined on a set of trace-orthonormal basis on $\mathcal{L}(\mathcal{H})$. Often, we use the *normalized Pauli group*, i.e., Pauli group with a normalization factor. The n-qubit Pauli group is given by

$$\mathsf{P}_n = \bigotimes_{i=1}^{n} \{I_i, X_i, Y_i, Z_i\}, \tag{A12}$$

where X_i, Y_i, Z_i are the single-qubit Pauli matrices. Then the normalized Pauli group is given by

$$\left\{ \sigma_i = \frac{1}{\sqrt{2^n}} P_i | P_i \in \mathsf{P}_n \right\}. \tag{A13}$$

Each pair of elements (σ_i, σ_j) in this group satisfies the following constraints under the Hilbert–Schmidt inner product, $\forall \sigma_i, \sigma_j \in \mathsf{P}_n / \sqrt{2^n}$:

$$\mathrm{Tr}(\sigma_i^{\dagger} \sigma_j) = \delta_{ij}. \tag{A14}$$

Any n-qubit operator can be decomposed over the 4^n normalized Pauli operators. We can rewrite the density operator O on $\mathcal{L}(\mathcal{H})$ in a vector form:

$$|O\rangle\rangle = \sum_{i \in P^n} \mathrm{Tr}(\sigma_i^{\dagger} O) \sigma_i. \tag{A15}$$

Moreover, any quantum channel can be represented as a matrix in the Liouville representation. To be specific, we can represent an arbitrary channel Λ acting on an operator O as follows:

$$|\Lambda(O)\rangle\rangle = \Lambda |O\rangle\rangle. \tag{A16}$$

We can see the element of this matrix is given by

$$\Lambda_{ij} = \langle\langle \sigma_i | \Lambda | \sigma_j \rangle\rangle = \mathrm{Tr}(\sigma_i \Lambda(\sigma_j)). \tag{A17}$$

Consequently, in the Liouville representation, the concatenation of two channels can be depicted as the product of two matrices:

$$|\Lambda_2 \circ \Lambda_1(O)\rangle\rangle = \Lambda_2 |\Lambda_1(O)\rangle\rangle = \Lambda_2 \Lambda_1 |O\rangle\rangle. \tag{A18}$$

The measurement operator can also be vectorized with the Liouville bra-notation according to the definition of the Hilbert–Schmidt inner product. For example, the measurement probability of a state ρ on a positive operator-valued measure $\{F_i\}$ is given by

$$p_i = \langle\langle F_i | \rho \rangle\rangle = \mathrm{Tr}(F_i^{\dagger} \rho). \tag{A19}$$

We call such a Pauli–Liouville representation as the Pauli transfer matrix (PTM) representation. An n-qubit quantum channel can also be described in the χ-matrix representation:

$$\Lambda(\rho) = d \sum_{i,j} \chi_{ij} \sigma_i \rho \sigma_j^{\dagger}. \tag{A20}$$

The process matrix χ is uniquely determined by the orthonormal operator basis $\{\sigma_j\}$, where the first element is proportional to the identity matrix $\sigma_0 = I/\sqrt{d}$, and $d = 2^n$ is the dimension of the quantum system. Often, we take the normalized Pauli operators as the basis of the χ-matrix representation. If a channel is diagonal in this representation, we call it *Pauli channel*.

3. Quantum Channel Fidelity

The process fidelity of a channel Λ can be defined with its χ-matrix representation:

$$F(\Lambda) = \chi_{00}(\Lambda). \tag{A21}$$

Here, the quantity χ_{00} is independent of the choices of operator basis $\{\sigma_j\}$. In the following, we set the operator basis to be normalized Pauli group. Define the *Pauli fidelity* of a quantum channel Λ,

$$\lambda_j = d^{-1} \mathrm{Tr}(P_j \Lambda(P_j)), \tag{A22}$$

which is the diagonal term of the PTM representation of Λ. We can relate Pauli fidelities to the diagonal terms of χ-matrix representation via Walsh–Hadamard transformation:

$$\lambda_j = \sum_i (-1)^{\langle i,j \rangle} \chi_{ii}. \tag{A23}$$

Here, $\langle i, j \rangle = 0$ if P_i commutes with P_j and $\langle i, j \rangle = 1$ otherwise. Then, one can derive the process fidelity from Eq. (A23):

$$F(\Lambda) = \chi_{00}(\Lambda) = \frac{1}{d^2}\sum_j \lambda_j, \qquad \text{(A24)}$$

which can be viewed as another definition of process fidelity.

There is a relation between the commonly used average fidelity F_{ave} and the process fidelity [32]:

$$F_{ave} = (dF + 1)/(d + 1). \qquad \text{(A25)}$$

Average fidelity is defined as

$$F_{ave} = \int d\psi \text{Tr}(|\psi\rangle\langle\psi|\Lambda(|\psi\rangle\langle\psi|)), \qquad \text{(A26)}$$

where the integral is implemented over Haar measure. These two fidelity measures are both well-defined metrics to quantify the closeness of a channel to the identity.

4. Representation Theory in Randomized Benchmarking

Now, we can use the representation theory to analyze the randomized benchmarking (RB) procedures. Let us start from a quick review of the standard RB protocol. Considering an n-qubit gate set G, RB is performed via sampling random gate sequences:

$$\mathcal{S}_{rb} = \mathcal{G}_{m+1}\mathcal{G}_m \cdots \mathcal{G}_1, \qquad \text{(A27)}$$

where for $1 \le i \le m + 1$, $\mathcal{G}_i$ denotes the quantum operation in the PTM representation of a unitary matrix $g_i \in \mathsf{G}$. Here, for $1 \le i \le m$, g_i is randomly sampled from group G and $g_{m+1} = (g_m \cdots g_2 g_1)^{-1}$. Then, one applies the random gate sequence $\mathcal{S}_{rb}$ to the input state $\rho_{\psi 0}$ and performs measurement Q_0 for a sufficient number of times to estimate the average survival probability,

$$f(m) = \mathbb{E}_{\mathcal{G}_1 \cdots \mathcal{G}_m} \langle\langle Q_0|\tilde{\mathcal{G}}_{m+1}\tilde{\mathcal{G}}_m \cdots \tilde{\mathcal{G}}_2\tilde{\mathcal{G}}_1|\rho_{\psi 0}\rangle\rangle, \qquad \text{(A28)}$$

where $\tilde{\mathcal{G}}_i$ denotes the noisy implementation of $\mathcal{G}_i$, $\rho_{\psi 0}$ is the noisy preparation of the initial state $|\psi\rangle$, and measurement Q_0 also includes errors. Here, we employ the gate independent noise assumption as used in most RB protocols. That is, the noise channels attached to the gate set $\{\mathcal{G}_i\}$ are the same, $\forall i$,

$$\tilde{\mathcal{G}}_i = \Lambda_L \mathcal{G}_i \Lambda_R, \qquad \text{(A29)}$$

where Λ_L and Λ_R are left and right noise channels. The randomization over the gate sequence $\mathcal{S}_{rb}$ can be seen as performing twirling operation for the noise channels between $\mathcal{G}_i$ and $\mathcal{G}_{i+1}$. Denote $\Lambda = \Lambda_R\Lambda_L$:

$$f(m) = \langle\langle Q|(\mathbb{E}_{g\in\mathsf{G}} \mathcal{G}^\dagger\Lambda\mathcal{G})^m|\rho_\psi\rangle\rangle, \qquad \text{(A30)}$$

where the right noise channel Λ_R of the first gate is absorbed into the state preparation error, $|\rho_\psi\rangle\rangle = \Lambda_R|\rho_{\psi 0}\rangle\rangle$ and the left noise channel Λ_L of the last gate is absorbed into the measurement error, $\langle\langle Q| = \langle\langle Q_0|\Lambda_L$. According to Proposition 1, one can express f_m in a more elegant manner,

$$f(m) = \sum_{\sigma\in R_\mathsf{G}} \langle\langle Q|\Pi_\sigma^m|\rho_\psi\rangle\rangle\lambda_\sigma^m, \qquad \text{(A31)}$$

where Π_σ denotes the projector onto the irreducible subspace associated with σ, and λ_σ contains the trace information of the channel Λ on the subspace:

$$\lambda_\sigma = \frac{\text{Tr}(\Lambda\Pi_\sigma)}{\text{Tr}(\Pi_\sigma)}. \qquad \text{(A32)}$$

Because the twirling operation will not change the trace value of Λ, the process fidelity of Λ can be given by

$$F(\Lambda, I) = 4^{-n}\sum_\sigma d_\sigma\lambda_\sigma, \qquad \text{(A33)}$$

where d_σ is the dimension of the irrep σ with $d_\sigma = \text{Tr}(\Pi_\sigma)$. Here, $F(\Lambda, I)$ is also known as the entanglement fidelity, or χ_{00}. We call $\{\lambda_\sigma\}$ the quality parameters because they reflect the noisy level of a channel.

In the conventional RB protocol, the n-qubit Clifford group C_n is often picked as the target gate set, which we call Clifford RB. Any Clifford operation $C \in \mathsf{C}_n$ satisfies $\{CP_iC^{-1}$ for all $i\} = \mathsf{P}_n$, which is a transformation permuting Pauli operators. Note that in the PTM representation, C_n has only one nontrivial irrep. Thus, we only need to solve one single quality parameter. This is rather convenient, but it is hard to extend the conventional RB scheme to other groups with multiple nontrivial irreps due to the multi-variable fitting problem, as shown in Eq. (A31), which has poor confidence intervals for $\{\Lambda_\sigma\}$.

To solve the fitting problem, in the following discussions, we employ the technique of character randomized benchmarking, which utilizes the generalized projection formula of Lemma 1 in the character theory.

APPENDIX B: BENCHMARKING PROTOCOLS

1. Character Cycle Benchmarking

Here, we present further technical details of the CCB protocol. As shown in Lemma 1, we can rewrite the projection equation of Eq. (A6) for a quantum operation group $\{\mathcal{G}\}$ in the PTM representation:

$$\Pi_\sigma = \frac{d_\sigma}{|\mathsf{G}|}\sum_{g\in\mathsf{G}} \chi_\sigma(g)\mathcal{G}. \qquad \text{(B1)}$$

Then one can add an additional gate, as a character gate, to construct the projector for extracting the quality parameter λ_σ associated with irrep σ. In CCB, we estimate the process fidelity of a target gate $U \in \mathsf{C}_n$ via the twirling group $\mathsf{G} = \mathsf{P}_n$, which has 4^n irreps supported by the Pauli operator $\{P_j\}$. The schematic circuit is given in Fig. 1(a) in the main text. The detailed procedures of CCB protocol are given in Box 2.

Note that in the CCB procedures, $P^{(0)}$ is the character gate, which we will merge into the gate $P^{(1)}$ in practical implementation. In addition, $P^{(0)}$ is not included in computing the inverse gate U_{inv}. The average P_j-weighted survival probability of Eq. (B4) can be further evaluated by

$$f_i(m) = d_j\langle\langle Q_j|\left(\left(\mathbb{E}_{P_t\in\mathsf{P}_n} \mathcal{P}_t^{-1}\mathcal{U}^{-1}\Lambda\mathcal{U}\mathcal{P}_t\right)\right.$$
$$\left. \times \left(\mathbb{E}_{P_r\in\mathsf{P}_n} \mathcal{P}_r^{-1}\Lambda\mathcal{P}_r\right)\right)^m\left(\mathbb{E}_{P\in\mathsf{P}_n} \chi_j(P)\mathcal{P}\right)|\rho_{\psi j}\rangle\rangle. \qquad \text{(B7)}$$

Box 2. Procedures for character cycle benchmarking

1. Sample a Pauli operator, $P_j \in \mathsf{P}_n$, and initialize the state, $|\psi_j\rangle$, such that $|\psi_j\rangle = P_j|\psi_j\rangle$.
2. Sample a gate sequence, $(P^{(0)}, P^{(1)}, P^{(2)}, \cdots, P^{(2m)})$, where $P^{(i)}(0 \le i \le 2m)$ are sampled uniformly at random from Pauli group P^n.
3. Apply the gate sequence,

$$\mathcal{S}_{\mathrm{ccb}} = \mathcal{U}_{\mathrm{inv}} \circ \mathcal{U}^{-1} \circ \mathcal{P}^{(2m)} \circ \mathcal{U} \circ \mathcal{P}^{(2m-1)} \circ \ldots \circ \mathcal{U}^{-1} \circ \mathcal{P}^{(1)} \circ \mathcal{U} \circ \mathcal{P}^{(1)} \circ \mathcal{P}^{(0)}, \tag{B2}$$

where the inverse gate $\mathcal{U}_{\mathrm{inv}} = \mathcal{P}^{(1)} \circ \mathcal{U}^{-1} \circ \ldots \circ P^{(2m)} \circ \mathcal{U}$ is a local gate.

4. Measure in P_j basis and compute the P_j-weighted survival probability,

$$f_j(m, \mathcal{S}_{\mathrm{ccb}}) = \chi_j(P^{(0)})d_j\mathrm{Tr}(Q_j\mathcal{S}(\rho_{\psi_j})), \tag{B3}$$

where the dimension of the representation associated with P_j is $d_j = 1$, the corresponding character is $\chi_j(P^{(0)}) = 1$ if P_j commutes with $P^{(0)}$ and –1 otherwise, and Q_j and ρ_{ψ_j} are the noisy implementations of the state preparation and measurement, respectively.

5. Repeat Steps 2 to 4 for a sufficient number of sequences and estimate the average value

$$f_j(m) = \mathop{\mathbb{E}}_{\mathcal{S}_{\mathrm{ccb}}} f_j(m, \mathcal{S}_{\mathrm{ccb}}). \tag{B4}$$

6. Repeat Steps 2 to 5 for different m and fit to the function

$$f_j(m) = A_j\lambda_j^{2m}, \tag{B5}$$

where A_j and λ_j are fitting parameters.

7. Sample M Pauli operators $\{P_j\}$ in Step 1, and for each P_j, repeat Steps 2 to 6. Finally estimate the CCB fidelity:

$$F_{\mathrm{ccb}} = \frac{1}{M}\sum_{\{P_j\}}\lambda_j. \tag{B6}$$

According to Proposition 1, we have

$$\Lambda_{\mathsf{P}} \equiv \mathop{\mathbb{E}}_{P_r \in \mathsf{P}_n} \mathcal{P}_r^{-1}\Lambda\mathcal{P}_r = \sum_{P_r \in \mathsf{P}_n} \langle\langle P_r|\Lambda|P_r\rangle\rangle\Pi_{P_r},$$

$$\Lambda_{\mathsf{P}}^U \equiv \mathop{\mathbb{E}}_{P_t \in \mathsf{P}_n} \mathcal{P}_t^{-1}\mathcal{U}^{-1}\Lambda\mathcal{U}\mathcal{P}_r$$

$$= \sum_{P_t \in \mathsf{P}_n} \langle\langle P_t|\mathcal{U}^{-1}\Lambda\mathcal{U}|P_t\rangle\rangle\Pi_{P_t} = \mathcal{U}^{-1}\Lambda_{\mathsf{P}}\mathcal{U}, \tag{B8}$$

where $\Pi_{P_r} = |P_r\rangle\rangle\langle\langle P_r|$ and $\Pi_{P_t} = |P_t\rangle\rangle\langle\langle P_t|$ are the projectors onto the support spaces of P_r and P_t in the PTM representation, respectively. Then Eq. (B7) can be further simplified to

$$f_i(m) = \langle\langle Q_j|(\Lambda_{\mathsf{P}}^U\Lambda_{\mathsf{P}})^m(d_j\mathop{\mathbb{E}}_{P \in \mathsf{P}_n}\chi_j(P)\mathcal{P})|\rho_s\rangle\rangle. \tag{B9}$$

According to the generalized projection formula of Lemma 1, we have

$$\Pi_j = \frac{d_j}{4^n}\sum_{P \in \mathsf{P}_n}\chi_j(P)\mathcal{P}, \tag{B10}$$

where $\forall j$, the dimension of the subspace associated with the P_j is $d_j = 1$. Substituting Eq. (B10) and Eq. (B8) into Eq. (B9), we have

$$f_j(m) = \langle\langle Q_j|\left(\sum_{P_t \in \mathsf{P}_n}\langle\langle P_t|\mathcal{U}^{-1}\Lambda\mathcal{U}|P_t\rangle\rangle\Pi_{P_t}\right.$$

$$\left.\times\sum_{P_r \in \mathsf{P}_n}\langle\langle P_r|\Lambda|P_r\rangle\rangle\Pi_{P_r}\right)^m\Pi_j|\rho_s\rangle\rangle$$

$$= \langle\langle Q_j|\Pi_j|\rho_s(\langle\langle P_j|\mathcal{U}^{-1}\Lambda\mathcal{U}|P_j\rangle\rangle\langle\langle P_j|\Lambda|P_j\rangle\rangle)^m. \tag{B11}$$

By fitting the survival probability $f_j(m)$ to the function of Eq. (B5), one can obtain the fitting parameters,

$$\lambda_j = \sqrt{\langle\langle P_j|\mathcal{U}^{-1}\Lambda\mathcal{U}|P_j\rangle\rangle\langle\langle P_j|\Lambda|P_j\rangle\rangle},$$

$$A_j = \langle\langle Q_j|\Pi_j|\rho_s\rangle\rangle. \tag{B12}$$

Note that quality parameter λ_j is one of the Pauli fidelities of channel $\sqrt{\mathcal{U}^{-1}\Lambda_{\mathsf{P}}\mathcal{U}\Lambda_{\mathsf{P}}}$. Then we can use the CCB protocol to sample Pauli operators $\{P_j\}$ and estimate process fidelity $F(\sqrt{\mathcal{U}^{-1}\Lambda_{\mathsf{P}}\mathcal{U}\Lambda_{\mathsf{P}}}, I)$, which is close to $F(\Lambda, I)$, as we will see in Section B.2.

2. CCB Fidelity

Here, we will explain the physical meaning of the estimated process fidelity in the CCB protocol. As shown in Section B.1, the process fidelity we estimate in the CCB protocol is called the CCB fidelity:

$$F_{\mathrm{ccb}}(\Lambda) = F(\sqrt{(\mathcal{U}^{-1}\Lambda_{\mathsf{P}}\mathcal{U})\Lambda_{\mathsf{P}}}, I). \tag{B13}$$

Here, Λ_P is the noise channel Λ twirled by the Pauli gate set defined in Eq. (B8). To compute the deviation between $F_{\mathrm{ccb}}(\Lambda)$ and $F(\Lambda, I)$, we employ the rearrangement inequality.

Lemma 2 (Rearrangement inequality [33]). *Consider two sets of n real numbers $\{x_1, x_2, \cdots, x_n\}$, $\{y_1, y_2, \cdots, y_n\}$, and $x_1 \le x_2 \le \cdots \le x_n, y_1 \le y_2 \le \cdots \le y_n$, one has*

$$x_ny_1 + \cdots + x_1y_n \le x_{\tau(1)}y_1 + \cdots + x_{\tau(n)}y_n$$

$$\le x_1y_1 + \cdots + x_ny_n, \tag{B14}$$

where $\{x_{\tau(i)}\}$ is an arbitrary permutation of $\{x_i\}$.

Any Clifford unitary $U \in \mathsf{C}_n$ satisfies, $\forall i$,

$$UP_iU^{-1} = P_{u(i)}, \tag{B15}$$

where $P_i \in \mathsf{P}_n$ and the index permutation $u(i)$ is determined by the operation U. Thus, the diagonal terms of the Pauli channel $\mathcal{U}^{-1}\Lambda_\mathsf{P}\mathcal{U}$ are a permutation of those in channel Λ_P.

Denote $\{\omega_i\}$ as the Pauli fidelities of Λ_P and $\{\omega_{u(i)}\}$ as the Pauli fidelities of the channel $\mathcal{U}^{-1}\Lambda_\mathsf{P}\mathcal{U}$. Here, we assume $\forall i$, $\omega_i \geq 0$ throughout the paper. In practice, the values of ω_i are normally close to 1 for a high-fidelity gate implementation. Using the rearrangement inequality, one has

$$\omega_{u(1)}\omega_1 + \cdots + \omega_{u(4^n)}\omega_{4^n} \leq \omega_1^2 + \cdots + \omega_{4^n}^2. \tag{B16}$$

Then one can further derive that

$$\sqrt{\omega_{u(1)}\omega_1} + \cdots + \sqrt{\omega_{u(4^n)}\omega_{4^n}} \leq \omega_1 + \cdots + \omega_{4^n}. \tag{B17}$$

From the definition of the process fidelity and CCB fidelity, one can conclude that

$$F_{\mathrm{ccb}}(\Lambda) \leq F(\Lambda, I). \tag{B18}$$

3. CCB with Local Gauge Freedom

In this section, we give a detailed analysis of the CCB protocol with local gauge freedom. As shown in Fig. 5(a), we insert local gates L and L^{-1} between the twirling gates P_i and the target gates U, U^{-1}. We define the local gate as $L = \otimes_{i=1}^n L_i$, where L_i can be an arbitrary single-qubit gate. Note that the local gates can be absorbed into P_i, U, U^{-1}, as shown in Fig. 5(b), and thus they do not need to be implemented individually. Likewise, the initial character gate $LP^{(0)}$ and the twirling gate $LP^{(1)}L^{-1}$ can be treated as single gates in experiments as well.

For simplicity, we assume that the local twirling gates are noiseless and the noises of LUL^{-1} and $LU^{-1}L^{-1}$ are the same. Now, we can analyze the relationship between the sequence length of gate layers and the survival probability. Given the noise channel of the target gate Λ, the averaged P_j-weighted survival probability is given by

$$\begin{aligned}
f_j(m) = d_j \langle\langle Q_j | (\underset{P_t \in \mathsf{P}_n}{\mathbb{E}} \mathcal{P}_t^{-1} \mathcal{L}^{-1} \mathcal{L} \mathcal{U}^{-1} \mathcal{L}^{-1} \Lambda \mathcal{L} \mathcal{U} \mathcal{L}^{-1} \mathcal{L} \mathcal{P}_t)^m \\
\times \mathcal{L}^{-1} \mathcal{L} (\underset{P_r \in \mathsf{P}_n}{\mathbb{E}} \mathcal{P}_r^{-1} \mathcal{L}^{-1} \Lambda \mathcal{L} \mathcal{P}_r) \mathcal{L}^{-1} \mathcal{L} (\underset{P \in \mathsf{P}_n}{\mathbb{E}} \chi_j(P)\mathcal{P})|\rho_s\rangle\rangle \\
= d_j \langle\langle Q_j | (\underset{P_t \in \mathsf{P}_n}{\mathbb{E}} \mathcal{P}_t^{-1} \mathcal{U}^{-1} \Lambda_L \mathcal{U} \mathcal{P}_t)^m (\underset{P_r \in \mathsf{P}_n}{\mathbb{E}} \mathcal{P}_r^{-1} \Lambda_L \mathcal{P}_r) \\
\times (\underset{P \in \mathsf{P}_n}{\mathbb{E}} \chi_j(P)\mathcal{P})|\rho_s\rangle\rangle, \tag{B19}
\end{aligned}$$

where $\Lambda_L = \mathcal{L}^{-1}\Lambda\mathcal{L}$. Equation (B19) is equivalent to Eq. (B7), except for substituting Λ with Λ_L. That means running the CCB protocol with the circuit in Fig. 5 would provide an estimation of the process fidelity $F(\Lambda_L, I)$, which is equivalent to $F(\Lambda, I)$:

$$F(\Lambda_L, I) = d^{-2}\mathrm{Tr}(\Lambda_L) = d^{-2}\mathrm{Tr}(\Lambda) = F(\Lambda, I). \tag{B20}$$

This accounts for the validity of the CCB protocol with local gauge freedom.

In the Pauli–Liouville representation, the off-diagonal terms of channel Λ vanish after the Pauli twirling. When applying $L\mathsf{P}_n L^{-1}$ for twirling, the off-diagonal terms of Λ in the L-transformed Pauli–Liouville representation, defined on $L\mathsf{P}_n L^{-1}/\sqrt{d}$, would vanish. Here, the local gauge transformation merely changes the representation of Λ, while maintaining the exponential decay form of the survival probability. From another point of view, the CCB protocol with local gauge freedom is characterizing the diagonal terms of χ-matrix of the noise channel under basis $L\mathsf{P}_n L^{-1}$ instead of P_n. As the average fidelity is irrelevant to the representation basis, all the analysis in previous subsections still applies.
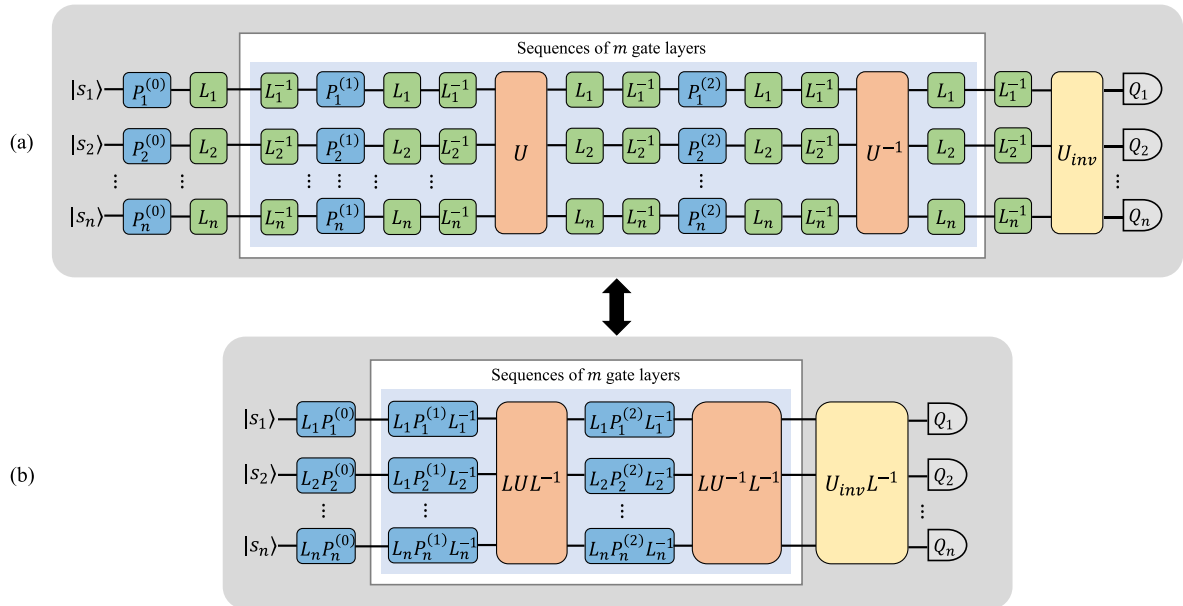
Fig. 5. Illustrations of the CCB circuit with local gauge transformation. Circuit (a) is equivalent to the original CCB circuit in the main text if all of gates are ideal. The orange boxes represent the target gate $\mathcal{U}$ and its inverse gate $\mathcal{U}^{-1}$. The blue boxes represent the random Pauli gates. The green boxes represent the inserted local gates L and L^{-1}, where $L = \otimes_{i=1}^n L_i$. The yellow box denote the inverse gate for the m inner gate layers in the light blue box. In practice, we implement gates in circuit (b) while absorbing local gates L and L^{-1} into twirling and target gates. Here, the target gate after gauge transformation becomes LUL^{-1}.

4. Character-Average Benchmarking

Here, we will provide more details of the CAB protocol. The CAB protocol can be seen as an improvement based on the CCB protocol, which adds an additional local Clifford gate $C \in \mathsf{C}_1^{\otimes n}$ to the beginning and C^{-1} to the end of the inner gate sequence, respectively, as shown in Fig. 1(b) in the main text. The detailed procedures of CAB protocol are given in Box 1 in the main text.

Let us first analyze the irreps of $\mathsf{C}_1^{\otimes n}$ in the PTM representation. According to Eq. (A5), the PTM representation of one-qubit Clifford group C_1 is the direct sum of the trivial representation, I, associated with the identity element, and a nontrivial irrep, Υ, supported by the subspace defined on the Pauli matrices $\{X, Y, Z\}$. For an n-qubit local Clifford group $\mathsf{C}_1^{\otimes n}$, there are 2^n irreps:

$$\mathcal{R}^n = \{I, \Upsilon\}^{\otimes n}. \qquad \text{(B21)}$$

According to Proposition 1, the twirling over $\mathsf{C}_1^{\otimes n}$ can be written as

$$\Lambda_{\mathsf{C}} = \sum_{\sigma_k \in \mathcal{R}^n} \frac{\mathrm{Tr}(\Lambda \Pi_{\sigma_k})}{\mathrm{Tr}(\Pi_{\sigma_k})} \Pi_{\sigma_k}, \qquad \text{(B22)}$$

where Π_{σ_k} denotes the projector onto the support space of σ_k in the PTM representation, $\mathrm{Tr}(\Pi_{\sigma_k}) = 3^{\pi(\sigma_k)}$ is the dimension of σ_k, and $\pi(\sigma_k)$ is the counting of Υ in σ_k. In the following, we will sometimes abuse the irrep notation and treat σ_k as the Pauli operator set that defines the support space of the irrep σ_k in the PTM representation.

It is obvious to see that channel Λ_{C} is a partially depolarizing channel, which has the same diagonal values in the subspace associated with the irrep σ_k for each k. We call the diagonal parameter as the *local Clifford eigenvalue*.

Definition 5 (Local Clifford eigenvalue). *The twirling channel Λ_{C} over n-qubit local Clifford group $\mathsf{C}_1^{\otimes}$ has 2^n quality parameters $\{\gamma_k\}$, which are defined as the local Clifford eigenvalues:*

$$\gamma_k = \frac{\mathrm{Tr}(\Lambda \Pi_{\sigma_k})}{\mathrm{Tr}(\Pi_{\sigma_k})} = \frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \langle\langle P_j | \Lambda | P_j \rangle\rangle, \qquad \text{(B23)}$$

where γ_k can be seen as an average value of the Pauli fidelities $\{\langle\langle P_j | \Lambda | P_j \rangle\rangle | P_j \in \sigma_k\}$.

We will give a concrete example to show how $\mathsf{C}_1^{\otimes n}$ twirls an arbitrary channel. For a two-qubit local Clifford group $\mathsf{C}_1 \otimes \mathsf{C}_1$, let us denote its irreps as

$$\sigma_1 = \{I \otimes I\} = \{II\},$$
$$\sigma_2 = \{I \otimes \Upsilon\} = \{IX, IY, IZ\},$$
$$\sigma_3 = \{\Upsilon \otimes I\} = \{XI, YI, ZI\},$$
$$\sigma_4 = \{\Upsilon \otimes \Upsilon\} = \{XX, XY, XZ, YX, YY, YZ, ZX, ZY, ZZ\},$$

$$\text{(B24)}$$

and the dimensions of these irreps are

$$\mathrm{Tr}(\Pi_{\sigma_1}) = |\sigma_1| = 3^0,$$
$$\mathrm{Tr}(\Pi_{\sigma_2}) = |\sigma_2| = 3^1,$$
$$\mathrm{Tr}(\Pi_{\sigma_3}) = |\sigma_3| = 3^1,$$
$$\mathrm{Tr}(\Pi_{\sigma_4}) = |\sigma_4| = 3^2. \qquad \text{(B25)}$$

Then the twirling of $\mathsf{C}_1^{\otimes 2}$ over a channel, Λ, according to Eq. (B22), is given by

$$\Lambda_{\mathsf{C}} = \mathrm{Tr}(\Lambda \Pi_{\sigma_1})\Pi_{\sigma_1} + \frac{\mathrm{Tr}(\Lambda \Pi_{\sigma_2})}{3}\Pi_{\sigma_2} + \frac{\mathrm{Tr}(\Lambda \Pi_{\sigma_3})}{3}\Pi_{\sigma_3}$$
$$+ \frac{\mathrm{Tr}(\Lambda \Pi_{\sigma_4})}{9}\Pi_{\sigma_4}, \qquad \text{(B26)}$$

which has four local Clifford eigenvalues $\{\mathrm{Tr}(\Lambda \Pi_{\sigma_1}),$ $\mathrm{Tr}(\Lambda \Pi_{\sigma_2})/3, \mathrm{Tr}(\Lambda \Pi_{\sigma_3})/3, \mathrm{Tr}(\Lambda \Pi_{\sigma_4})/9\}$.

Now return to the n-qubit case. The randomization over an inner gate layer in CAB is the same as CCB and generates a composite Pauli channel:

$$\Lambda_{\mathrm{in}} = \mathcal{U}^{-1} \Lambda_{\mathsf{P}} \mathcal{U} \Lambda_{\mathsf{P}}, \qquad \text{(B27)}$$

where Λ_{P} is a Pauli twirling channel defined in Eq. (B8). The initial and last random local Clifford gates $\mathcal{C}, \mathcal{C}^{-1}$ in CAB together perform a local Clifford twirling over the inner Pauli channel:

$$\Lambda_{\mathsf{C}}(m) = \underset{C \in \mathsf{C}_1^{\otimes n}}{\mathbb{E}} \mathcal{C}^{-1} \Lambda_{\mathrm{in}}^m \mathcal{C}$$
$$= \sum_{\sigma_k \in \mathcal{R}^n} \frac{\mathrm{Tr}(\Lambda_{\mathrm{in}}^m \Pi_{\sigma_k})}{\mathrm{Tr}(\Pi_{\sigma_k})} \Pi_{\sigma_k}. \qquad \text{(B28)}$$

Note that $\Lambda_{\mathsf{C}}(m)$ is a partially depolarizing channel, as we mentioned above. Thus one can extract local Clifford eigenvalues via performing corresponding measurement observables $P \in \sigma_k$. In the CAB protocol, by measuring in $Z^{\otimes n}$ basis, we can infer the measurement results of the 2^n observables $\{I, Z\}^{\otimes n}$, which span all the irrep spaces in $\mathcal{R}^n$. We assume that the measurement $Z^{\otimes n}$ is performed with negligible errors that will not influence the fidelity estimations too much. If one wants to completely remove the SPAM errors, an additional character gate from the gate set $\{I, Z\}^{\otimes n}$ can be added to the beginning of the CAB gate sequence, which is similar to the CCB protocol.

The survival probability of Eq. (9) in the main text can be derived as for Pauli operator $Q_k \in \{I, Z\}^{\otimes n}$:

$$f_k(m) = \langle\langle Q_k | \Lambda_{\mathsf{C}}(m) | \rho_s \rangle\rangle, \qquad \text{(B29)}$$

which contains fidelity information for irrep σ_k such that $Q_k \in \sigma_k$. Substituting Eqs. (B8) and (B28) to Eq. (B29), we have

$$f_k(m) = \left\langle\left\langle Q_k \left| \Pi_{Q_k} \Pi_{\sigma_k} \frac{\mathrm{Tr}(\Lambda_{\mathrm{in}}^m \Pi_{\sigma_k})}{\mathrm{Tr}(\Pi_{\sigma_k})} \right| \rho_s \right\rangle\right\rangle$$
$$= \langle\langle Q_k | \rho_s \rangle\rangle \sum_{P_j \in \sigma_k} \frac{(\langle\langle P_j | \mathcal{U}^{-1} \Lambda \mathcal{U} | P_j \rangle\rangle \langle\langle P_j | \Lambda | P_j \rangle\rangle)^m}{\mathrm{Tr}(\Pi_{\sigma_k})}$$
$$= \langle\langle Q_k | \rho_s \rangle\rangle \sum_{P_j \in \sigma_k} \frac{\lambda_j^{2m}}{\mathrm{Tr}(\Pi_{\sigma_k})}, \qquad \text{(B30)}$$

where λ_j is the Pauli fidelity of the channel $\sqrt{\mathcal{U}^{-1} \Lambda_{\mathsf{P}} \mathcal{U} \Lambda_{\mathsf{P}}}$ defined in Eq. (B12). By fitting the survival probability $f_k(m)$ to the function $f_k(m) = A_k \mu_k^{2m}$, one can solve the quality parameters $\{\mu_k\}$ and estimate the CAB fidelity as

$$F_{\mathrm{cab}} = 4^{-n} \sum_{\sigma_k \in \mathcal{R}^n} d_{\sigma_k} \mu_k, \qquad \text{(B31)}$$

where $d_{\sigma_k} = \mathrm{Tr}(\Pi_{\sigma_k})$ is the dimension of irrep σ_k.

Note that our fitting method averages multi-exponential decays $\{\lambda_j^{2m}\}$ into one exponential decay μ_k^{2m}, which leads to $\mu_k \geq \mathrm{ave}_{P_j \in \sigma_k} \lambda_j$, as proven in Section B.5. We will further prove that the CAB fidelity is the upper bound of the CCB fidelity in Section B.5.

5. Fitting Analysis for CAB

Here, we will analyze the fitting results in the CAB protocol and explain the CAB fidelity F_{cab} in detail. As shown in Eq. (B30), the survival probability in CAB is given by

$$f_k(m) = \frac{\langle\langle Q_k | \rho_s \rangle\rangle}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j^{2m}, \tag{B32}$$

where $|\sigma_k| = \mathrm{Tr}(\Pi_{\sigma_k})$. Fit $f_k(m)$ to the function,

$$f_k(m) = A_k \mu_k^{2m}, \tag{B33}$$

with $A_k > 0, \mu_k > 0$. Take the natural logarithm for both sides in Eq. (B34),

$$y = \beta_0 + \beta_1 m, \tag{B34}$$

where

$$y = \ln f_k(m),$$
$$\beta_0 = \ln A_k,$$
$$\beta_1 = 2 \ln \mu_k. \tag{B35}$$

Next, we will employ the *least-squares estimation* for the linear regression of Eq. (B34). Set $\{m_1, m_2, \cdots, m_q\}$ as inputs, and assume $m_1 < m_2 < \cdots < m_q$ without loss of generality. The regression matrix M and the observed values Y are given by

$$M = \begin{pmatrix} 1 & m_1 \\ 1 & m_2 \\ \vdots & \vdots \\ 1 & m_q \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_q \end{pmatrix}. \tag{B36}$$

Using the least-squares estimation, one can solve the optimum parameters for the model

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (M^T M)^{-1} M^T Y. \tag{B37}$$

The fitting parameter of interest in CAB is μ_k:

$$\mu_k = \exp\left(\frac{\hat{\beta}_1}{2}\right), \tag{B38}$$

where

$$\hat{\beta}_1 = \frac{q \sum_{i=1}^{q} m_i y_i - \sum_{i=1}^{q} m_i \sum_{i=1}^{q} y_i}{q \sum_{i=1}^{q} m_i^2 - \left(\sum_{i=1}^{q} m_i\right)^2}. \tag{B39}$$

Then, we have the following lemma.

Lemma 3. *Denote $\{\mu_k\}$ as the fitting parameters we solve in the CAB protocol, given in Eq. (B38). For each μ_k, we have*

$$\mu_k \geq \frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j, \tag{B40}$$

and the CAB fidelity is the upper bound of the CCB fidelity,

$$F_{cab}(\Lambda) \geq F_{ccb}(\Lambda). \tag{B41}$$

Proof. For a simple linear regression using the least-squares estimation, there are some observations above the fitting curve,

while the others are below the curve. Then one can conclude that there exist two adjacent observations whose slope in between is smaller than the fitting slope of the curve. Thus for the simple linear regression in CAB, the fitting slope holds

$$\hat{\beta}_1 \geq \min_i \left(\frac{y_{i+1} - y_i}{m_{i+1} - m_i}\right). \tag{B42}$$

According to the Chebyshev sum inequality [33], $\forall i$,

$$\frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j^{2m_{i+1}} = \frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j^{2m_i} \lambda_j^{2(m_{i+1} - m_i)}$$
$$\geq \left(\frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j^{2m_i}\right)\left(\frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j^{2(m_{i+1} - m_i)}\right). \tag{B43}$$

According to Eq. (B43), we can further derive that

$$y_{i+1} - y_i = \ln\left(\frac{\sum_{P_j \in \sigma_k} \lambda_j^{2m_{i+1}}}{\sum_{P_j \in \sigma_k} \lambda_j^{2m_i}}\right)$$
$$\geq \ln\left(\frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j^{2(m_{i+1} - m_i)}\right)$$
$$\geq 2(m_{i+1} - m_i) \ln\left(\frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j\right), \tag{B44}$$

where the last inequality comes from the convexity of function $x^{2(m_{i+1} - m_i)}$.

Substituting Eq. (B44) into Eq. (B42), we have

$$\hat{\beta}_1 \geq 2 \ln\left(\frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j\right). \tag{B45}$$

Since $\hat{\beta}_1 = 2 \ln \mu_k$, we can conclude that

$$\mu_k \geq \frac{1}{|\sigma_k|} \sum_{P_j \in \sigma_k} \lambda_j. \tag{B46}$$

Besides, the dimensions of irreps of the local Clifford gate $\mathsf{C}_1^{\otimes n}$ satisfies

$$\sum_{\sigma_k \in \mathcal{R}^n} |\sigma_k| = \sum_{\sigma_k \in \mathcal{R}^n} \mathrm{Tr}(\Pi_{\sigma_k}) = 4^n. \tag{B47}$$

Thus,

$$F_{cab} = 4^{-n} \sum_{\sigma_k \in \mathcal{R}^n} |\sigma_k| \mu_k \geq 4^{-n} \sum_{\sigma_k \in \mathcal{R}^n} \sum_{P_j \in \sigma_k} \lambda_j = F_{ccb}(\Lambda), \tag{B48}$$

which completes the proof.

APPENDIX C: STATISTICAL ANALYSIS

Here, we analyze the statistical fluctuation of the CCB protocol with finite sampling. Recall that the CCB fidelity is given by Eq. (B13),

$$F_{ccb}(\Lambda) = \frac{1}{4^n} \sum_j \lambda_j, \tag{C1}$$

where λ_j is the Pauli fidelity of channel $\sqrt{\mathcal{U}^{-1} \Lambda_\mathsf{P} \mathcal{U} \Lambda_\mathsf{P}}$ related to the Pauli operator P_j. It is impractical to solve all the Pauli fidelities using the character RB method when qubit number n becomes large. One can sample a finite number of Pauli irreps

to estimate the process fidelity of the channel. The reliability of these estimates is expressed by the confidence levels.

In the CCB protocol, the Pauli fidelity is obtained by fitting the survival probability and the gate sequence length, which is very hard for the statistical analysis. For simplicity, we take two points in the fitting diagram to analyze the fluctuation of the slope. In what follows, we use the notation where $\hat{x}$ is an estimator of a quantity, $\bar{x}$, where the bar denotes that either an expected value or a sample average has been taken over realizations of random variable x.

We begin by describing the CCB protocol in a statistical way.

1. Choose Pauli operator $P_j \in \mathsf{P}_n$.
2. Choose positive integer m_1.
3. Choose random gate sequence s_1 from gate set $\mathbb{S}_{m_1}$ and obtain an estimate of the P_j-weighted probability, $\hat{f}_j(m_1, s_1)$.
4. Repeat Step 3 K_1 times to estimate

$$\hat{f}_j(m_1) = \frac{1}{K_1}\sum_{s_1}\hat{f}_j(m_1, s_1), \tag{C2}$$

where the K_1 gate sequences form a gate sequence set, $\mathbb{S}_1 \subset \mathbb{S}_{m_1}$.
5. Choose another bigger positive integer, $m_2 > m_1$.
6. Choose random gate sequence s_2 from gate sequence set $\mathbb{S}_{m_2}$ and obtain an estimate of the P_j-weighted probability, $\hat{f}_j(m_2, s_2)$.
7. Repeat Step 6 K_2 times to estimate

$$\hat{f}_j(m_2) = \frac{1}{K_2}\sum_{s_2}\hat{f}_j(m_2, s_2), \tag{C3}$$

where the K_2 gate sequences form a gate sequence set, $\mathbb{S}_2 \subset \mathbb{S}_{m_2}$.
8. Estimate the Pauli fidelity

$$\hat{\lambda}_j = \left(\frac{\hat{f}_j(m_2)}{\hat{f}_j(m_1)}\right)^{\frac{1}{m_2-m_1}}. \tag{C4}$$

9. Sample M Pauli operators $\{P_j\}$ in Step 1, and for each P_j, repeat Steps 2–8. Finally, estimate the CCB fidelity

$$\hat{F}_{\text{ccb}} = \frac{1}{M}\sum_{\{P_j\}}\hat{\lambda}_j. \tag{C5}$$

The main statistical errors are divided into two parts in the above protocol. The first comes from the sampling randomness of the Pauli fidelity estimation $\hat{\lambda}_j$ for each P_j in Eq. (C4). The second comes from the sampling randomness of the Pauli operators $\{P_j\}$ in Eq. (C5). We shall calculate the confidence levels for these two sampling randomness, respectively.

We first calculate the bias of the Pauli fidelity estimation:

$$\Delta\lambda_j := |\underset{\mathbb{S}_1, \mathbb{S}_2}{\mathbb{E}}[\hat{\lambda}_j] - \bar{\lambda}_j|, \tag{C6}$$

where $\bar{\lambda}_j$ denotes the theoretical value of the Pauli fidelity,

$$\bar{\lambda}_j = \left(\frac{\bar{f}_j(m_2)}{\bar{f}_j(m_1)}\right)^{\frac{1}{m_2-m_1}}. \tag{C7}$$

Here, the expectations of the probability estimators are taken over the gate sequences,

$$\bar{f}_j(m_1) = \underset{s_1 \in \mathbb{S}_{m_1}}{\mathbb{E}}[\hat{f}_j(m_1, s_1)],$$

$$\bar{f}_j(m_2) = \underset{s_2 \in \mathbb{S}_{m_2}}{\mathbb{E}}[\hat{f}_j(m_2, s_2)]. \tag{C8}$$

To calculate the expectation value of the ratio estimator in Eq. (C6), we take the expectations over the gate sequence sets $\mathbb{S}_1, \mathbb{S}_2$ on both sides of Eq. (C4),

$$\underset{\mathbb{S}_1, \mathbb{S}_2}{\mathbb{E}}[\hat{\lambda}_j] = \underset{\mathbb{S}_2}{\mathbb{E}}[\hat{f}_j(m_2)^{\frac{1}{m_2-m_1}}]\underset{\mathbb{S}_1}{\mathbb{E}}[\hat{f}_j(m_1)^{-\frac{1}{m_2-m_1}}]. \tag{C9}$$

The expectations for $\hat{f}_j(m_1)$ and $\hat{f}_j(m_2)$ can be separated since the random variables s_1 and s_2 are independent. By denoting

$$\hat{b} := \hat{f}_j(m_2), \quad \bar{b} := \bar{f}_j(m_2),$$

$$\hat{a} := \hat{f}_j(m_1), \quad \bar{a} := \bar{f}_j(m_1),$$

$$t := \frac{1}{m_2 - m_1}, \tag{C10}$$

we have

$$\underset{\mathbb{S}_1, \mathbb{S}_2}{\mathbb{E}}[\hat{\lambda}_j] = \underset{\mathbb{S}_1}{\mathbb{E}}[\hat{a}^{-t}]\underset{\mathbb{S}_2}{\mathbb{E}}[\hat{b}^t]. \tag{C11}$$

Supposing that

$$\delta_a := \frac{\hat{a} - \bar{a}}{\bar{a}} \ll 1,$$

$$\delta_b := \frac{\hat{b} - \bar{b}}{\bar{b}} \ll 1, \tag{C12}$$

and using the second-order approximation of the Taylor expansion at $\delta_a = 0, \delta_b = 0$ for $\hat{a}^{-t}, \hat{b}^t$, respectively, we have

$$\hat{a}^{-t} = \bar{a}^{-t}(1 + \delta_a)^{-t}$$
$$= \bar{a}^{-t}(1 - t\delta_a + t(t+1)\delta_a^2 + \mathcal{O}(\delta_a^3)),$$
$$\hat{b}^t = \bar{b}^t(1 + \delta_b)^t$$
$$= \bar{b}^t(1 + t\delta_b + t(t-1)\delta_b^2 + \mathcal{O}(\delta_b^3)). \tag{C13}$$

Take expectations over s_1, s_2 for Eq. (C13),

$$\underset{\mathbb{S}_1}{\mathbb{E}}[\hat{a}^{-t}] = \bar{a}^{-t}\left(1 + t(t+1)\frac{\text{Var}[\hat{a}]}{\bar{a}^2} + \mathcal{O}(\delta_a^3)\right),$$

$$\underset{\mathbb{S}_2}{\mathbb{E}}[\hat{b}^t] = \bar{b}^t\left(1 + t(t-1)\frac{\text{Var}[\hat{b}]}{\bar{b}^2} + \mathcal{O}(\delta_b^3)\right). \tag{C14}$$

Then the expectation value of the ratio estimator is given by

$$\underset{\mathbb{S}_1, \mathbb{S}_2}{\mathbb{E}}[\hat{\lambda}_j] = \left(\frac{\bar{b}}{\bar{a}}\right)^t\left(1 + t(t+1)\frac{\text{Var}[\hat{a}]}{\bar{a}^2}\right)\left(1 + t(t-1)\frac{\text{Var}[\hat{b}]}{\bar{b}^2}\right)$$
$$+ \mathcal{O}(\delta_a^3, \delta_b^3)$$
$$= \bar{\lambda}_j\left(1 + t(t+1)\frac{\text{Var}[\hat{a}]}{\bar{a}^2} + t(t-1)\frac{\text{Var}[\hat{b}]}{\bar{b}^2}\right)$$
$$+ \mathcal{O}(\delta_a^3, \delta_b^3). \tag{C15}$$

Substituting Eq. (C15) into Eq. (C6), we can derive the bias of the Pauli fidelity estimation,

$$\Delta\lambda_j = \bar{\lambda}_j\left(t(t+1)\frac{\mathrm{Var}[\hat{a}]}{\bar{a}^2} + t(t-1)\frac{\mathrm{Var}[\hat{b}]}{\bar{b}^2}\right) + \mathcal{O}(\delta_a^3, \delta_b^3). \tag{C16}$$

Recall that the assumptions of Eq. (D17) are established with specific failure probabilities. To calculate the confidence intervals for the aforementioned assumptions, we apply Bernstein's inequality.

Lemma 4 (Bernstein's inequality [34]). *Consider a set of n independent random variables $\{X_1, \cdots, X_n\}$ with $\forall i, X_i \leq b$. Let $X = \frac{1}{n}\sum_i X_i$, denote $\mathbb{V}^2 = n^{-1}\sum_{i=1}^n \mathrm{Var}(X_i)$, then $\forall \epsilon > 0$,*

$$\mathrm{Pr}(|X - \mathbb{E}[X]| > \epsilon) \leq 2\exp\left(-\frac{n\epsilon^2/2}{\mathbb{V}^2 + b\epsilon/3}\right). \tag{C17}$$

Assuming that $\forall m, \hat{f}_j(m) \leq 1$, we derive the following proposition.

Proposition 2. *Let $\hat{f}_j(m)$ be the estimator of P_j-weighted probability $\bar{f}_j(m)$ for fixed K, m, and j. Then $\forall \epsilon > 0$,*

$$\mathrm{Pr}(|\hat{f}_j(m) - \bar{f}_j(m)| > \epsilon) \leq 2\exp\left(-\frac{K\epsilon^2/2}{\mathrm{Var}[\hat{f}_j(m,s)] + \epsilon/3}\right). \tag{C18}$$

Then for the estimators $\hat{a}$ and $\hat{b}$, we can derive the corresponding confidence intervals

$$\mathrm{Pr}(|\hat{a} - \bar{a}| > \epsilon_1) \leq 2\exp\left(-\frac{K_1\epsilon_1^2/2}{\mathrm{Var}[\hat{f}_j(m_1, s_1)] + \epsilon_1/3}\right)$$
$$:= \Upsilon_a(K_1, \epsilon_1),$$

$$\mathrm{Pr}(|\hat{b} - \bar{b}| > \epsilon_2) \leq 2\exp\left(-\frac{K_2\epsilon_2^2}{\mathrm{Var}[\hat{f}_j(m_2, s_2)] + \epsilon_2/3}\right)$$
$$:= \Upsilon_b(K_2, \epsilon_2). \tag{C19}$$

Substituting Eq. (C19) into Eq. (C16), we obtain the following lemma.

Lemma 5. *For any given Pauli operator $P_j \in P_n$ and some fixed m_1, m_2, K_1, K_2, the bias of the Pauli fidelity estimation is upper bounded by*

$$\Delta\lambda_j \leq \frac{1}{(m_2 - m_1)^2}\left|\frac{1 + m_2 - m_1}{K_1\bar{f}_j^2(m_1)}\mathrm{Var}[\hat{f}_j(m_1, s_1)]\right.$$
$$\left. + \frac{1 - m_2 + m_1}{K_2\bar{f}_j^2(m_2)}\mathrm{Var}[\hat{f}_j(m_2, s_2)]\right| + \mathcal{O}(\epsilon_1^3, \epsilon_2^3)$$
$$:= \epsilon_b(m_1, m_2; K_1, K_2; \epsilon_1, \epsilon_2), \tag{C20}$$

with a failure probability bounded by $\Upsilon_j(K_1, K_2; \epsilon_1, \epsilon_2) = \Upsilon_a(K_1, \epsilon_1) + \Upsilon_b(K_2, \epsilon_2)$.

The notation $\epsilon_b(m_1, m_2; K_1, K_2; \epsilon_1, \epsilon_2)$ is abbreviated as ϵ_b in the following analysis. Next, we compute the confidence interval for the CCB fidelity $\hat{F}_{\mathrm{ccb}}$ defined in Eq. (C5). We denote

$$\bar{F}'_{\mathrm{ccb}} = \frac{1}{M}\sum_{\{P_j\}}\bar{\lambda}_j, \quad \bar{F}_{\mathrm{ccb}} = \frac{1}{4^n}\sum_j\bar{\lambda}_j. \tag{C21}$$

Assume the Pauli fidelity $0 \leq \bar{\lambda}_j \leq 1$ for all P_j. One can apply the Hoeffding's inequality directly.

Lemma 6 (Hoeffding's inequality [35]). *Consider a set of n independent random variables $\{X_1, \cdots, X_n\}$ and $\forall i, a_i \leq X_i \leq b_i$. Let $X = \frac{1}{n}\sum_i X_i$, then $\forall \epsilon > 0$,*

$$\mathrm{Pr}(|X - \mathbb{E}[X]| > \epsilon) \leq 2\exp\left(-\frac{2n^2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \tag{C22}$$

Then we can derive that $\forall \epsilon_M > 0$,

$$\mathrm{Pr}(|\bar{F}_{\mathrm{ccb}} - \bar{F}'_{\mathrm{ccb}}| > \epsilon_M) \leq 2\exp(-2M\epsilon_M^2). \tag{C23}$$

According to Lemma 5, we have

$$\mathrm{Pr}(|\hat{F}_{\mathrm{ccb}} - \bar{F}'_{\mathrm{ccb}}| > \epsilon_b) \leq M\Upsilon_j(K_1, K_2; \epsilon_1, \epsilon_2), \tag{C24}$$

where $\hat{F}_{\mathrm{ccb}}$ is the estimator of $\bar{F}_{\mathrm{ccb}}$, as defined in Eq. (C5). By combining Eqs. (C23) and (C24) and applying the union bound, we can compute the confidence interval for the process fidelity estimator

$$\mathrm{Pr}(|\hat{F}_{\mathrm{ccb}} - \bar{F}_{\mathrm{ccb}}| > \epsilon_M + \epsilon_b) \leq 2\exp(-M\epsilon_M^2/2)$$
$$+ M\Upsilon_j(K_1, K_2; \epsilon_1, \epsilon_2). \tag{C25}$$

Let us assume that

$$0 \leq \mathrm{Var}[\hat{f}_j(m_1, s_1)] \leq 1, \quad 0 \leq \mathrm{Var}[\hat{f}_j(m_2, s_2)] \leq 1, \tag{C26}$$

and then Eq. (C25) can be simplified to

$$\mathrm{Pr}(|\hat{F}_{\mathrm{ccb}} - \bar{F}_{\mathrm{ccb}}| > \epsilon_M + \epsilon_b) \leq 2\exp(-M\epsilon_M^2/2)$$
$$+ 2M\exp\left(-\frac{K_1\epsilon_1^2/2}{1 + \epsilon_1/3}\right) + 2M\exp\left(-\frac{K_2\epsilon_2^2/2}{1 + \epsilon_2/3}\right). \tag{C27}$$

We further assume that

$$\frac{1}{2} < \bar{f}_j(m_1) < 1, \tag{C28}$$

and then we can derive the following theorem.

Theorem 2. *Consider a CCB implementation with sampling numbers K_1, K_2 at sequence lengths m_1, m_2 with estimation errors ϵ_1, ϵ_2, respectively, and the expected survival probability $\bar{f}_j(m_1)$ that satisfies $\forall j, 1/2 < \bar{f}_j(m_1) < 1$. The estimated CCB fidelity, $\hat{F}_{\mathrm{ccb}}$, is given by the average over M Pauli fidelities $\{\lambda_j\}$. The confidence probability for the CCB fidelity falling into the estimated interval $[\hat{F}_{\mathrm{ccb}} - \epsilon_M - \epsilon_b, \hat{F}_{\mathrm{ccb}} + \epsilon_M + \epsilon_b]$ is greater than $1 - \delta$:*

$$\mathrm{Pr}(|\hat{F}_{\mathrm{ccb}} - \bar{F}_{\mathrm{ccb}}| \leq \epsilon_M + \epsilon_b) \geq 1 - \delta, \tag{C29}$$

with

$$\epsilon_b \leq \frac{4}{K_1}\frac{1}{m_2 - m_1}\left(\frac{1}{m_2 - m_1} + 1\right) + \mathcal{O}(\epsilon_1^3, \epsilon_2^3), \tag{C30}$$

and

$$2\exp(-2M\epsilon_M^2) + 2M\exp\left(-\frac{K_1\epsilon_1^2/2}{1 + \epsilon_1/3}\right)$$
$$+ 2M\exp\left(-\frac{K_2\epsilon_2^2/2}{1 + \epsilon_2/3}\right) = \delta. \tag{C31}$$

A simplified informal version of the above theorem is presented in the main text as Theorem 1.

APPENDIX D: SIMULATION

Here, we will present the noise model for the simulation in the main text and give more details on the two-qubit controlled-(TX) gate and the five-qubit error correcting circuit. In addition, we compare a CAB process and an interleaved character randomized benchmarking (ICRB) process [28,36] for

benchmarking the two-qubit CZ gate. We provide its simulation details and the results at the end of this part.

1. Error Model
The noise channel, Λ_t, we consider here for the target gate is composed of a Pauli channel Λ_0, an amplitude damping channel Λ_1, and a qubit–qubit correlation channel Λ_2, $\Lambda_t = \Lambda_0 \circ \Lambda_1 \circ \Lambda_2$. We consider the noise channel Λ_{ref} for the twirling gate set P_n is a gate-independent Pauli channel, which is negligible compared with Λ_t.

1. Stochastic Pauli channel Λ_0.
Pauli channel Λ_0 can be written as

$$\Lambda_0(\rho) = \sum_i p_i P_i \rho P_i^{-1}, \tag{D1}$$

where p_i is the Pauli error rate related to the Pauli operator P_i. As for Λ_0, it is equivalent to say that operator $P_i \in \mathsf{P}_n$ applies on the density matrix ρ with probability p_i. We can further rewrite Λ_0 in the PTM representation,

$$\Lambda_0 = \sum_i \lambda_i |P_i\rangle\rangle\langle\langle P_i|, \tag{D2}$$

where λ_i is the Pauli fidelity. Pauli channel Λ_0 in the simulation contains dephasing errors and cross-talk errors. In reality, the fidelity of error channel Λ_t is mainly determined by Λ_0.

2. Amplitude-damping channel Λ_1.
Each qubit in the simulation is subject to an amplitude damping channel,

$$\Lambda_1 = \bigotimes_{i=1}^n \Lambda_i^d, \tag{D3}$$

where Λ_i^d is the single-qubit damping channel for qubit i,

$$\Lambda_i^d = K_i^{(0)}\rho K_i^{(0)\dagger} + K_i^{(1)}\rho K_i^{(1)\dagger},$$

$$K_i^{(0)} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\alpha_i} \end{pmatrix},$$

$$K_i^{(1)} = \begin{pmatrix} 0 & \sqrt{\alpha_i} \\ 0 & 0 \end{pmatrix}, \tag{D4}$$

with damping parameter α_i.

3. Qubit–qubit correlation channel Λ_2.
The qubit–qubit correlation channel Λ_2 is a coherent error channel in the simulation,

$$\Lambda_2 = \bigotimes_{i<j} e^{i\beta_{ij}\text{SWAP}_{ij}}, \tag{D5}$$

where β_{ij} is the correlation parameter describing the interacting strength between qubits i and j.

2. Simulations for the Controlled-(TX) Gate
The controlled-(TX) gate can be decomposed as

$$\text{CTX} = (I \otimes \sqrt{T})\text{CNOT}(I \otimes \sqrt{T^{-1}}). \tag{D6}$$

Then, we can take $I \otimes \sqrt{T}$ as the local gauge transformation, and the twirling gate set turns to

$$P_{\text{ctx}} = (I \otimes \sqrt{T})P_2(I \otimes \sqrt{T^{-1}}). \tag{D7}$$

Consider noise channel $\Lambda_t = \Lambda_0 \circ \Lambda_1 \circ \Lambda_2$ in Section D.1 for the noisy controlled-(TX) gate. We randomly sample the Pauli fidelities of Λ_0 from a normal distribution $\mathcal{N}(\mu, \sigma)$,

denoted as a $\mathcal{N}(\mu, \sigma)$-Pauli channel, where μ and σ are the mean value and standard deviation. The parameters for Λ_1 are set to $\alpha_1 = \alpha_2 = 0.005$. The parameter for Λ_2 is set to $\beta_{12} = 0.01$. In the following discussions, we label the noise channel for the controlled-(TX) with $\Lambda_t(\mu, \sigma)$, since Λ_1 and Λ_2 remain the same in all the simulations.

We simulate the CAB and CCB protocols for the noisy controlled-(TX) gate with eight different noise channels $\{\Lambda_t(\mu_i, \sigma_i)\}$, and the noisy implementations of CTX and CTX^{-1} are given by

$$\tilde{\text{CTX}} = \text{CTX} \circ \Lambda_t(\mu, \sigma), \quad \tilde{\text{CTX}}^{-1} = \text{CTX}^{-1} \circ \Lambda_t(\mu, \sigma). \tag{D8}$$

The error parameters are taken as $\{(\mu_i, \sigma_i)\} = \{(0.995, 0.001), (0.990, 0.002), (0.980, 0.003), (0.970, 0.004), (0.960, 0.005), (0.950, 0.006), (0.940, 0.007), (0.930, 0.008)\}$. Take the $\mathcal{N}(0.998, 0.001)$-Pauli channel as the noise channel of the twirling gate set Λ_{ref} and then denote the noisy implementation of the twirling gate set as

$$\tilde{\mathsf{P}}_{\text{ctx}} = \Lambda_{\text{ref}} \circ \mathsf{P}_{\text{ctx}}. \tag{D9}$$

Take the $\mathcal{N}(0.998, 0.001)$-Pauli channel as the SPAM error channel Λ_{spam} and then denote the noisy implementations of the initial state $|\psi\rangle$ and measurement Q as

$$\rho_\psi = \Lambda_{\text{spam}}(|\psi\rangle\langle\psi|), \quad \tilde{Q} = Q \circ \Lambda_{\text{spam}}. \tag{D10}$$

The simulation procedures for CAB run as follows.

1. For each noise channel $\Lambda_t(\mu_i, \sigma_i)$, select a set of sequence length $\{m\} = \{1, 2, \cdots, m_{\max}\}$, where $m_{\max}$ satisfies $\mu_i^{m_{\max}} \approx \mu_i/3$.

2. For each sequence length m, sample $K = 50$ random gate sequences $\{(C, P_{\text{ctx}}^{(1)}, \cdots, P_{\text{ctx}}^{(2m)})\}$, where C and $P_{\text{ctx}}^{(i)}$ ($1 \le i \le 2m$) are sampled uniformly at random from $\mathsf{C}_1^{\otimes 2}$ and P_{ctx}, respectively. For each gate sequence, the noisy implementation in PTM is given by

$$\tilde{\mathcal{S}}_{\text{cab}} = \Lambda_{\text{ref}}\mathcal{C}^{-1}\Lambda_{\text{ref}}\mathcal{U}_{\text{inv}}\text{CTX}^{-1}\Lambda_t\Lambda_{\text{ref}}\mathcal{P}_{\text{ctx}}^{(2m)}...\mathcal{P}_{\text{ctx}}^{(2)}$$
$$\times \text{CTX}\Lambda_t\Lambda_{\text{ref}}\mathcal{P}_{\text{ctx}}^{(1)}\mathcal{C}, \tag{D11}$$

where the inverse gate is given by $\mathcal{U}_{\text{inv}} = (\text{CTX}^{-1}P_{\text{ctx}}^{(2m)}\cdots\text{CTX}^{-1}P_{\text{ctx}}^{(2)}\text{CTX}P_{\text{ctx}}^{(1)})^{-1}$.

3. Compute the survival probability over the $K = 50$ gate sequences for each measurement observable $Q_k \in \{II, IZ, ZI, ZZ\}$:

$$f_k(m) = \frac{1}{K}\sum_{\mathcal{S}_{\text{cab}}}\langle\langle\tilde{Q}_k|\tilde{\mathcal{S}}_{\text{cab}}|\rho_\psi\rangle\rangle, \tag{D12}$$

where $|\psi\rangle = |0\rangle^{\otimes 2}$.

4. For each Q_k, fit $f_k(m)$ to the function

$$f_k(m) = A_k\mu_k^{2m}. \tag{D13}$$

5. Estimate the CAB fidelity as

$$F_{\text{cab}} = \frac{1}{16}(\mu_{II} + 3\mu_{IZ} + 3\mu_{ZI} + 9\mu_{ZZ}). \tag{D14}$$

The simulation procedures for CCB run as follows.

1. For each noise channel $\Lambda_t(\mu_i, \sigma_i)$, select a set of sequence length $\{m\} = \{1, 2, \cdots, m_{\max}\}$, where $m_{\max}$ satisfies $\mu_i^{m_{\max}} \approx \mu_i/3$.

2. Sample $M = 10$ operators $\{P_j\}$ uniformly at random from P_{ctx}.

3. For each P_j at each sequence length m, sample $K = 50$ random gate sequences $\{(P_{\text{ctx}}^{(0)}, P_{\text{ctx}}^{(1)}, \cdots, P_{\text{ctx}}^{(2m)})\}$, where $P_{\text{ctx}}^{(i)}$ are sampled uniformly at random from P_{ctx}. For each gate sequence, the noisy implementation in PTM is given by

$$\tilde{S}_{\text{cab}} = \Lambda_{\text{ref}}\mathcal{U}_{\text{inv}}\text{CTX}^{-1}\Lambda_t\Lambda_{\text{ref}}\mathcal{P}_{\text{ctx}}^{(2m)}\ldots\mathcal{P}_{\text{ctx}}^{(2)}\text{CTX}\Lambda_t\Lambda_{\text{ref}}\mathcal{P}_{\text{ctx}}^{(1)}\mathcal{P}_{\text{ctx}}^{(0)}, \tag{D15}$$

where the inverse gate is given by $\mathcal{U}_{\text{inv}} = (\text{CTX}^{-1}P_{\text{ctx}}^{(2m)}\cdots\text{CTX}^{-1}P_{\text{ctx}}^{(2)}\text{CTX}P_{\text{ctx}}^{(1)})^{-1}$.

4. Compute the survival probability over the $K = 50$ gate sequences for each measurement P_j,

$$f_j(m) = \frac{1}{K}\sum_{\mathcal{S}_{\text{ccb}}}\chi_j(P^{(0)})\langle\langle P_j|\Lambda_{\text{spam}}\tilde{S}_{\text{ccb}}|\rho_\psi\rangle\rangle, \tag{D16}$$

where $\chi_j(P^{(0)}) = 1$ if $P^{(0)}$ commutes with P_j and -1 otherwise, and $|\psi\rangle$ is the $+1$ eigenstate of P_j.

5. Fit $f_j(m)$ to the function

$$f_j(m) = A_j\lambda_j^{2m}. \tag{D17}$$

6. Estimate the CCB fidelity as

$$F_{\text{ccb}} = \frac{1}{M}\sum_{\{P_j\}}\lambda_j. \tag{D18}$$

3. Simulations for the Five-Qubit Error Correcting Circuit

Here, we take the five-qubit stabilizer encoding circuit shown in Fig. 6 as the target gate U, which only contains Clifford gates.

The noise channel we consider here has the form $\Lambda_t = \Lambda_0 \circ \Lambda_1 \circ \Lambda_2$, as presented in Section D.1. We set the $\mathcal{N}(0.98, 0)$-Pauli channel as Λ_0, which can be seen as a depolarizing channel, $\Lambda_{\text{dep}}(\rho) = p\rho + (1-p)I/d$ where $p = 0.98$. The error parameters $\{\alpha_i\}$ and $\{\beta_{ij}\}$ are sampled uniformly at random from the intervals $[0, 0.02]$ and $[0, 0.01]$, respectively. Then the noisy implementation of the target gate is given by

$$\tilde{\mathcal{U}} = \mathcal{U} \circ \Lambda_t. \tag{D19}$$

Since the five-qubit encoding circuit is a Clifford gate, we take the five-qubit Pauli group P_5 as the twirling gate set. For simplicity, we ignore the errors from the SPAM and twirling gates, $\Lambda_{\text{spam}} = I$ and $\Lambda_{\text{ref}} = I$.
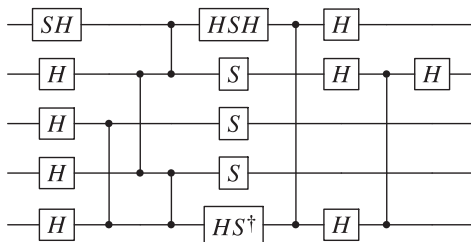
The simulation procedures for CAB run as follows.

Fig. 6. Five-qubit stabilizer encoding circuit.

1. For the five-qubit noise channel Λ_t, select a set of sequence lengths, $\{m\} = \{1, 2, \cdots, 20\}$.

2. For each sequence length m, sample K random gate sequences $\{(C, P^{(1)}, \cdots, P^{(2m)})\}$, where C and $P^{(i)}$ are sampled uniformly at random from $C_1^{\otimes 5}$ and P_5, respectively. For each gate sequence, the noisy implementation in PTM is given by

$$\tilde{S}_{\text{cab}} = \mathcal{C}^{-1}\mathcal{U}_{\text{inv}}\mathcal{U}^{-1}\Lambda_t\mathcal{P}^{(2m)}\ldots\mathcal{P}^{(2)}\mathcal{U}\Lambda_t\mathcal{P}^{(1)}\mathcal{C}, \tag{D20}$$

where the inverse gate is given by $\mathcal{U}_{\text{inv}} = (\mathcal{U}^{-1}P^{(2m)}\cdots\mathcal{U}^{-1}P^{(2)}\mathcal{U}P^{(1)})^{-1}$.

3. Compute the survival probability over the K gate sequences for each measurement observable $Q_k \in \{I, Z\}^{\otimes 5}$,

$$f_k(m) = \frac{1}{K}\sum_{\mathcal{S}_{\text{cab}}}\langle\langle\tilde{Q}_k|\tilde{S}_{\text{cab}}|\rho_\psi\rangle\rangle, \tag{D21}$$

where $|\rho_\psi\rangle = |0\rangle^{\otimes 5}$.

4. For each Q_k, fit $f_k(m)$ to the function

$$f_k(m) = A_k\mu_k^{2m}. \tag{D22}$$

5. Estimate the CAB fidelity as

$$F_{\text{cab}} = \frac{1}{4^5}\sum_k d_k\mu_k, \tag{D23}$$

where d_k is the dimension of the kth irrep, defined in Eq. (B31).

Note that the target gate is a Clifford gate. In the simulation, we obtain the inverse gate, $U_{\text{inv}} = \circ_{i=1}^{20} U^{-1}P_{2i}UP_{2i-1}$, by considering the permutation relation of a Clifford gate acting on Pauli operators. Furthermore, the survival probability for observable Q_k can be obtained from the Z basis measurement result. To be specific, Q_k can be expressed as the linear combination of $|z\rangle\langle z|$, $Q_k = \sum_z t_{kz}|z\rangle\langle z|$, where $z \in \{0,1\}^5$. Then the survival probability in Eq. (D21) is given by

$$\begin{aligned}f_k(m) &= \frac{1}{K}\sum_{\mathcal{S}_{\text{cab}}}\langle\langle\tilde{Q}_k|\tilde{S}_{\text{cab}}|\rho_\psi\rangle\rangle \\ &= \frac{1}{K}\sum_{\mathcal{S}_{\text{cab}}}\text{Tr}(Q_k\tilde{S}_{\text{cab}}(|0\rangle\langle 0|)) \\ &= \frac{1}{K}\sum_{\mathcal{S}_{\text{cab}}}\sum_z t_{kz}\text{Tr}(|z\rangle\langle z|\tilde{S}_{\text{cab}}(|0\rangle\langle 0|)),\end{aligned} \tag{D24}$$

where S_{cab} represents the channel of $\mathcal{S}_{\text{cab}}$.

For comparison, we also simulate the XEB protocol to benchmark the fidelity of the five-qubit error correcting circuit with the same noise channel. The simulation procedures for XEB run as follows.

1. For the five-qubit noise channel Λ_t, select a set of sequence lengths, $\{m\} = \{1, 2, \cdots, 20\}$.

2. For each sequence length m, sample K random gate sequences $\{(C^{(1)}, \cdots, C^{(2m)})\}$, where $C^{(i)}$ are sampled uniformly at random from $C_1^{\otimes 5}$. For each gate sequence, the noisy implementation in PTM is given by

$$\tilde{S}_{\text{xeb}} = \mathcal{U}\Lambda_t\mathcal{C}^{(2m)}\cdots\Lambda_t\mathcal{C}^{(2)}\mathcal{U}\Lambda_t\mathcal{C}^{(1)}. \tag{D25}$$

3. Compute the Z basis measurement result for each gate sequence $\mathcal{S}_{\text{xeb}}$, for $z \in \{0,1\}^5$,

$$f_{\text{xeb}}(m, z) = \langle\langle z|\tilde{\mathcal{S}}_{\text{xeb}}|\rho_\psi\rangle\rangle, \qquad \text{(D26)}$$

where $|\psi\rangle = |0\rangle^{\otimes 5}$.

4. Compute the ideal Z basis measurement result if the gate is ideal for each gate sequence, for $z \in \{0,1\}^5$,

$$f'_{\text{xeb}}(m, z) = \langle\langle z|\mathcal{U}\mathcal{C}^{(2m)}\cdots\mathcal{C}^{(2)}\mathcal{U}\mathcal{C}^{(1)}|\rho_\psi\rangle\rangle. \qquad \text{(D27)}$$

5. Compute the average XEB fidelity over the K gate sequences:

$$f_{\text{xeb}}(m) = \frac{1}{K}\sum_{\mathcal{S}_{\text{xeb}}}\frac{2^n\sum_z f_{\text{xeb}}(m, z)f'_{\text{xeb}}(m, z) - 1}{2^n\sum_z f_{\text{xeb}}(m, z)^2 - 1}. \qquad \text{(D28)}$$

6. Fit $f_{\text{xeb}}(m)$ to the function

$$f_{\text{xeb}}(m) = A_{\text{xeb}}p^{2m} + B_{\text{xeb}}. \qquad \text{(D29)}$$

The estimation of process fidelity with XEB is given by $p + (1 - p)/d^2$.

4. Comparison between CAB and ICRB for Benchmarking a CZ Gate

In this part, we present the comparison of CAB and ICRB for benchmarking a CZ gate. The noise model is the same as the one in Appendix D.1, including a Pauli channel, amplitude damping channel, and a qubit–qubit correlation channel. Specifically, we set the $\mathcal{N}(0.99,0)$-Pauli channel as Λ_0. For the amplitude damping channel Λ_1 and the qubit–qubit correlation channel Λ_2, the error parameters $\{\alpha_i\}$ and $\{\beta_{ij}\}$ are sampled uniformly at random from the intervals $[0, 0.01]$ and $[0, 0.01]$, respectively. Then the noisy implementation of CZ gate is set as

$$\widetilde{\text{CZ}} = \text{CZ} \circ \Lambda_2 \circ \Lambda_1 \circ \Lambda_0. \qquad \text{(D30)}$$

The fidelity of the generated noise channel $\Lambda_2 \circ \Lambda_1 \circ \Lambda_0$ is 0.9864 in our simulation, matching the fidelity of a real CZ gate in experiments. We consider an error probability of 0.02 for state preparation. That is, for each qubit, the actual prepared state is $0.98|0\rangle\langle 0| + 0.02|1\rangle\langle 1|$ for $|0\rangle$ and is $0.02|0\rangle\langle 0| + 0.98|1\rangle\langle 1|$ for $|1\rangle$. For simplicity, we set the measurement to be perfect computational basis measurement.

The simulation of the CAB process is similar to previous cases. The only difference is that we consider the effect of finite number of measurements in this case. That means, after computing the computational basis measurement probability for one sequence, we use this probability to generate a measurement frequency associated with a fixed number of single-shot measurements. Then the measurement frequency is used to compute the survival probability as well as fidelity. For simulating ICRB [28,36], we consider the same measurement setting and take the same number of single-shot measurements for one-sample sequence.

In this simulation, the circuit depth of CAB is set to $\{1,2,5,10,20,50\}$, corresponding to an overall CZ gate number $\{2,4,10,20,40,100\}$. To ensure that the numbers of implemented CZ gates in the two protocols are identical, the circuit depth of ICRB is set to $\{2,4,10,20,40,100\}$. The sample complexity for each circuit depth is taken from $\{5,10,25,50,100,200\}$, and the number of single shots for one sample sequence is taken from $\{100,200\}$. Then the amount of single-shot measurements for CAB is

the number of circuit depths

$\times$ sample complexity for one depth

$\times$ single shots for one sequence.   **(D31)**

Due to the additional sample complexity brought by the character gate, the amount of single-shot measurements for ICRB is

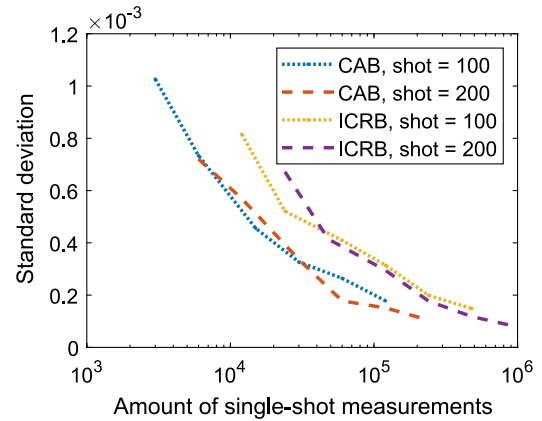$2^n \times$ the number of circuit depths

$\times$ sample complexity for one depth

$\times$ single shots for one sequence.   **(D32)**

It can be seen that the factor 2^n associated with the character gate is an exponential overhead. We separately simulate 50 experiments for each setting and compute the mean and standard deviation of the fidelity of these 50 experiments. The simula-
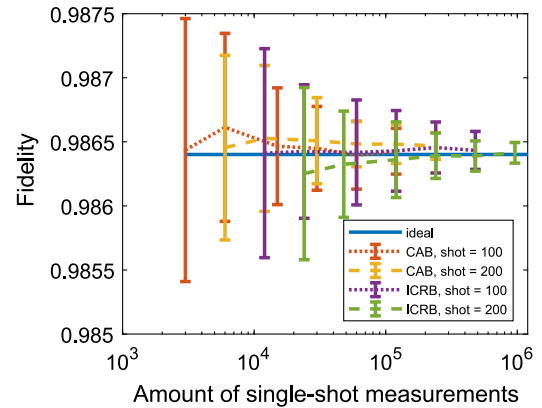
(a)

(b)

Fig. 7. Simulation results of CAB and ICRB for benchmarking a CZ gate. (a) The error bar of the fidelity obtained from 50 experiments via CAB and ICRB protocols. The x-axis represents the amount of single-shot measurements, and the y-axis represents the fidelity. The term "shot" represents the number of single shots associated with one sample sequence. One can see that CAB and ICRB can both obtain an accurate estimation of the process fidelity of a CZ gate. (b) The standard deviation of the fidelity obtained from 50 experiments via CAB and ICRB protocols. The x-axis represents the amount of single-shot measurements, and the y-axis represents the standard deviation of the fidelity. The figure shows that CAB has an advantage over ICRB with regard to standard deviation.

tion results of the two protocols are presented in Fig. 7. The results show that CAB and ICRB have a similar accuracy in benchmarking CZ, and CAB is a little better than ICRB in the sense of standard deviation. That means, to estimate the process fidelity of CZ within a given precision, the amount of single-shot measurements for CAB is less than that for ICRB, showing an advantage of CAB in benchmarking experiments.

Disclosures.   The authors declare no conflicts of interest.

Data Availability.   Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. J. Preskill, "Quantum computing in the NISQ era and beyond," Quantum **2**, 79 (2018).
2. L. M. K. Vandersypen and I. L. Chuang, "NMR techniques for quantum control and computation," Rev. Mod. Phys. **76**, 1037–1069 (2005).
3. S. Chu, "Cold atoms and quantum control," Nature **416**, 206–210 (2002).
4. B. M. Terhal, "Quantum error correction for quantum memories," Rev. Mod. Phys. **87**, 307–346 (2015).
5. E. T. Campbell, B. M. Terhal, and C. Vuillot, "Roads towards fault-tolerant universal quantum computation," Nature **549**, 172–179 (2017).
6. I. L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box," J. Mod. Opt. **44**, 2455–2467 (1997).
7. D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, "Quantum state tomography via compressed sensing," Phys. Rev. Lett. **105**, 150401 (2010).
8. S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, "Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators," New J. Phys. **14**, 095022 (2012).
9. S. T. Flammia and Y.-K. Liu, "Direct fidelity estimation from few Pauli measurements," Phys. Rev. Lett. **106**, 230501 (2011).
10. J. Emerson, R. Alicki, and K. Zyczkowski, "Scalable noise estimation with random unitary operators," J. Opt. B **7**, S347–S352 (2005).
11. J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, "Symmetrized characterization of noisy quantum processes," Science **317**, 1893–1896 (2007).
12. 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).
13. E. Magesan, J. M. Gambetta, and J. Emerson, "Scalable and robust randomized benchmarking of quantum processes," Phys. Rev. Lett. **106**, 180504 (2011).
14. E. Magesan, J. M. Gambetta, and J. Emerson, "Characterizing quantum gates via randomized benchmarking," Phys. Rev. A **85**, 042311 (2012).
15. S. T. Flammia and J. J. Wallman, "Efficient estimation of Pauli channels," ACM Trans. Quantum Comput. **1**, 1–32 (2020).
16. R. Harper, W. Yu, and S. T. Flammia, "Fast estimation of sparse quantum noise," PRX Quantum **2**, 010322 (2021).
17. 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).
18. J. P. Gaebler, A. M. Meier, T. R. Tan, R. Bowler, Y. Lin, D. Hanneke, J. D. Jost, J. P. Home, E. Knill, D. Leibfried, and D. J. Wineland, "Randomized benchmarking of multiqubit gates," Phys. Rev. Lett. **108**, 260503 (2012).
19. O. Moussa, M. P. da Silva, C. A. Ryan, and R. Laflamme, "Practical experimental certification of computational quantum gates using a twirling procedure," Phys. Rev. Lett. **109**, 070504 (2012).
20. R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, "Superconducting quantum circuits at the surface code threshold for fault tolerance," Nature **508**, 500–503 (2014).
21. D. Lu, H. Li, D.-A. Trottier, J. Li, A. Brodutch, A. P. Krismanich, A. Ghavami, G. I. Dmitrienko, G. Long, J. Baugh, and R. Laflamme, "Experimental estimation of average fidelity of a Clifford gate on a 7-qubit quantum processor," Phys. Rev. Lett. **114**, 140505 (2015).
22. C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol, and D. M. Lucas, "High-fidelity quantum logic gates using trapped-ion hyperfine qubits," Phys. Rev. Lett. **117**, 060504 (2016).
23. J. P. Gaebler, T. R. Tan, Y. Lin, Y. Wan, R. Bowler, A. C. Keith, S. Glancy, K. Coakley, E. Knill, D. Leibfried, and D. J. Wineland, "High-fidelity universal gate set for 9Be+ ion qubits," Phys. Rev. Lett. **117**, 060505 (2016).
24. T. Proctor, S. Seritan, K. Rudinger, E. Nielsen, R. Blume-Kohout, and K. Young, "Scalable randomized benchmarking of quantum computers using mirror circuits," Phys. Rev. Lett. **129**, 150502 (2021).
25. E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, "Efficient measurement of quantum gate error by interleaved randomized benchmarking," Phys. Rev. Lett. **109**, 080505 (2012).
26. J. J. Wallman, "Randomized benchmarking with gate-dependent noise," Quantum **2**, 47 (2018).
27. S. T. Merkel, E. J. Pritchett, and B. H. Fong, "Randomized benchmarking as convolution: Fourier analysis of gate dependent errors," Quantum **5**, 581 (2021).
28. J. Helsen, X. Xue, L. M. K. Vandersypen, and S. Wehner, "A new class of efficient randomized benchmarking protocols," NPJ Quantum Inf. **5**, 71 (2019).
29. A. Erhard, J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt, "Characterizing large-scale quantum computers via cycle benchmarking," Nat. Commun. **10**, 5347 (2019).
30. F. Arute, K. Arya, and R. Babbush et al., "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).
31. W. Fulton and J. Harris, Representation Theory (Springer, 2004).
32. M. Horodecki, P. Horodecki, and R. Horodecki, "General teleportation channel, singlet fraction, and quasidistillation," Phys. Rev. A **60**, 1888–1898 (1999).
33. G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities, 2nd ed. (Cambridge University, 1952), Vol. **XII**, p. 324.
34. S. Bernstein, The Theory of Probabilities (Gastehizdat Publishing House, 1946).
35. W. Hoeffding, "Probability inequalities for sums of bounded random variables," J. Am. Stat. Assoc. **58**, 13–30 (1963).
36. X. Xue, T. F. Watson, J. Helsen, D. R. Ward, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, S. Wehner, and L. M. K. Vandersypen, "Benchmarking gate fidelities in a Si/SiGe two-qubit device," Phys. Rev. X **9**, 021011 (2019).