Scalable Evaluation of Quantum-Circuit Error Loss Using Clifford Sampling

Zhen Wang, Yanzhu Chen, Zixuan Song, Dayue Qin, Hekang Li, Qiujiang Guo, H. Wang, Chao Song, and Ying Li

Phys. Rev. Lett. **126**, 080501 — Published 22 February 2021
DOI: 10.1103/PhysRevLett.126.080501

Scalable evaluation of quantum-circuit error loss using Clifford sampling

Zhen Wang[1,*], Yanzhu Chen[2,3,*], Zixuan Song[1], Dayue Qin[4], Hekang Li[1], Qiujiang Guo[1], H. Wang[1], Chao Song[1,†], and Ying Li[4,‡]

[1] *Interdisciplinary Center for Quantum Information, State Key Laboratory of Modern Optical Instrumentation, and Zhejiang Province Key Laboratory of Quantum Technology and Device, Department of Physics, Zhejiang University, Hangzhou 310027, China*

[2] *C. N. Yang Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, NY 11794-3840, USA*

[3] *Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, NY 11794-3800, USA*

[4] *Graduate School of China Academy of Engineering Physics, Beijing 100193, China*

A major challenge in developing quantum computing technologies is to accomplish high precision tasks by utilizing multiplex optimization approaches, on both the physical system and algorithm levels. Loss functions assessing the overall performance of quantum circuits can provide the foundation for many optimization techniques. In this paper, we use the quadratic error loss and the final-state fidelity loss to characterize quantum circuits. We find that the distribution of computation error is approximately Gaussian, which in turn justifies the quadratic error loss. It is shown that these loss functions can be efficiently evaluated in a scalable way by sampling from Clifford-dominated circuits. We demonstrate the results by numerically simulating ten-qubit noisy quantum circuits with various error models as well as executing four-qubit circuits with up to ten layers of two-qubit gates on a superconducting quantum processor. Our results pave the way towards the optimization-based quantum device and algorithm design in the intermediate-scale quantum regime.

Introduction.—In quantum computation, errors caused by decoherence and imperfect controls form the main obstacle to meaningful applications, such as solving integer factorization and quantum chemistry problems [1–4]. Evaluating the error severity in quantum computation is essential for improving the design of device [5–7], optimizing control parameters [8], and minimizing errors with mitigation protocols [9–11]. Various schemes of quantum system characterization have been developed. Randomized benchmarking [12–21] and quantum process tomography (QPT) [22–29] can measure the average gate fidelity and full information of a noisy quantum channel, respectively. These two methods are efficient in systems with a few qubits. Cross-entropy benchmarking [30, 31] and heavy output generation [32, 33] are used to verify a multi-qubit system but cannot be directly applied to circuits that are unsimulatable on classical computers. We can infer the performance of a large system by dividing it into tractable subsystems and characterizing each subsystem individually [11, 31, 34–37]. In this approach, given finite-size subsystems, not all crosstalk effects are incorporated. Cycle benchmarking is able to account for spatial correlation by repeating certain patterns of gates [38]. The temporal correlation of noise is another factor that usually limits the effectiveness of characterization techniques [29, 39–45].

Many quantum algorithms utilize multi-qubit and deep quantum circuits. Even for variational quantum computation, we need to implement hundreds of gates on tens of qubits [46–50]. In this paper, we propose an intuitive method that can efficiently characterize large quantum circuits, in the presence of both spatial and temporal error correlations. The resource cost of our method scales polynomially with the circuit size.

We take the quadratic loss function of computation error [51] as the measure of error severity, which is

$$L_{\mathbb{R}}(\boldsymbol{F}) \equiv \frac{1}{|\mathbb{R}|} \sum_{\boldsymbol{R}\in\mathbb{R}} \mathrm{Error}(\boldsymbol{F},\boldsymbol{R})^2. \tag{1}$$

Here $\mathrm{Error}(\boldsymbol{F},\boldsymbol{R}) \equiv \mathrm{com}(\boldsymbol{F},\boldsymbol{R}) - \mathrm{com}^{\mathrm{ef}}(\boldsymbol{F},\boldsymbol{R})$ is the computation error, $\mathrm{com}(\boldsymbol{F},\boldsymbol{R})$ and $\mathrm{com}^{\mathrm{ef}}(\boldsymbol{F},\boldsymbol{R})$ are respectively results (means of an observable, i.e., a function f of measurement outcomes) in the actual noisy computation and error-free computation, and $(\boldsymbol{F},\boldsymbol{R})$ specifies a quantum circuit. This loss function characterizes errors in a set of circuits with the same circuit frame $\boldsymbol{F}$ [see Fig. 1(a)] in the computation of f, which can be generalized to multiple circuit frames and observables by adding up error losses of different $\boldsymbol{F}$ and f. The circuit frame includes the qubit initialization, measurement and multi-qubit entangling gates (e.g., controlled-NOT and controlled-phase gates), which are usually more error-prone compared with single-qubit gates. We focus on the case that entangling gates are all Clifford. Frame operations are chosen to contain the noisiest components in the circuit and together with unitary single-qubit gates they form a universal set of operations for quantum computation. In the circuit set, each one contains a different configuration of single-qubit gates denoted by $\boldsymbol{R}$. $\mathbb{R}$ is the set of single-qubit gate configurations. When $\mathbb{R} = \mathbb{U}$, single-qubit gates can be any unitary transformations, and the summation should be taken as integration with respect to Haar measure [52]; when $\mathbb{R} = \mathbb{C}$, single-qubit gates are all Clifford. Our approach is particularly relevant to

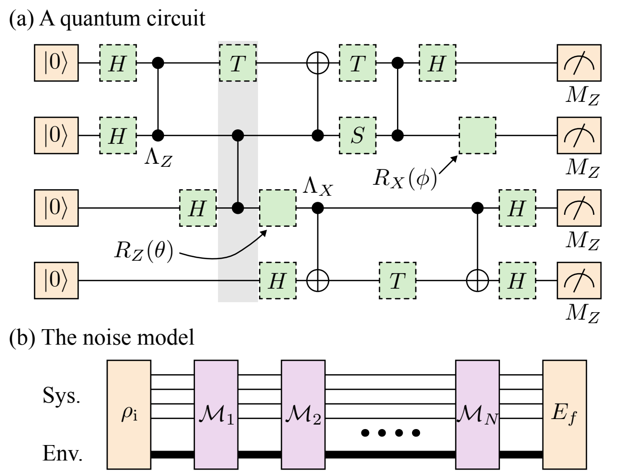

FIG. 1. (a) An example quantum circuit. The circuit $(\boldsymbol{F},\boldsymbol{R})$ consists of a circuit frame $\boldsymbol{F}$ and single-qubit unitary gates $\boldsymbol{R}$. Single-qubit gates are the green dashed squares, and all other operations form the frame. To evaluate $L_{\mathbb{R}}(\boldsymbol{F})$, the frame is always the same, but we sample random configurations of single-qubit gates. (b) The noise model. The system (Sys.) formed by qubits and the environment (Env.) are initialized in the state ρ_{i}. Following the initialization, there is a sequence of completely-positive maps applied. The map $\mathcal{M}_\tau$ describes the evolution applied at the time τ to realize a layer (column) of quantum gates, which actually acts on the system and environment because of imperfections. Finally, we implement the measurement on each qubit. The operator of observable is E_f, which acts on both the system and environment because of imperfections.

variational quantum algorithms which treat single-qubit gates as variables [3, 53, 54], and one can choose the circuit frame and observable accordingly. Loss functions in this form can be used to determine parameters in the learning-based quantum error mitigation [51, 55, 56].

In this paper, we demonstrate that the quadratic error loss $L_{\mathbb{U}}(\boldsymbol{F})$ (i.e., $\mathbb{R}=\mathbb{U}$) is a good objective function and can be efficiently evaluated when the circuit is large. By sampling random circuits, we study the statistics of $\mathrm{Error}(\boldsymbol{F},\boldsymbol{R})$ in experiments on a quantum device with four superconducting qubits and numerical simulations with up to ten qubits using various error models. We find that the error distribution is approximately Gaussian with zero mean when general-unitary circuits are uniformly sampled from $\boldsymbol{R}\in\mathbb{U}$ according to Haar measure, i.e. $L_{\mathbb{U}}(\boldsymbol{F})$ is the only value that we need for characterizing the error statistics [57]. Computing the error-free result $\mathrm{com}^{\mathrm{ef}}(\boldsymbol{F},\boldsymbol{R})$ is impractical for large general-unitary circuits but efficient for Clifford circuits, according to the Gottesman-Knill theorem [1, 58–60]. We prove $L_{\mathbb{U}}(\boldsymbol{F})=L_{\mathbb{C}}(\boldsymbol{F})$ under the assumption that errors in single-qubit gates are gate-independent, i.e., we can obtain $L_{\mathbb{U}}(\boldsymbol{F})$ by only sampling Clifford circuits. Note that we do not need any assumption on frame-operation errors. In addition to the analytical proof, the equivalence between unitary sampling and Clifford sampling is verified in both experiments and numerical simulations, by confirming that error losses from two sampling approaches are within statistical fluctuations of each other.

Formalism and Clifford sampling.—In a quantum circuit, we can draw gates applied in parallel in the same layer (column), see Fig. 1(a). For example, the gray box is the fourth layer, which contains a T gate and a controlled-phase gate. When gates are error-free, the overall map of the fourth layer is $\mathcal{M}_4^{\mathrm{ef}}=[T\otimes\Lambda_Z\otimes I]$, where an operator U acts on the density matrix ρ by $[U](\rho)=U\rho U^\dagger$, and I is the identity operator of a qubit. Because of the noise, the actual implementation leads to a different map $\mathcal{M}_4$, which acts on not only qubits but also the environment. This is a general formalism of errors in the quantum computation, including both spatial and temporal correlations. The temporal correlation is caused by the environment [61]. According to this formalism, we can express the actual computation result with error as $\mathrm{com}(\boldsymbol{F},\boldsymbol{R})=\mathrm{Tr}[E_f\mathcal{M}_N\cdots\mathcal{M}_2\mathcal{M}_1(\rho_{\mathrm{i}})]$ for an N-layer circuit [see Fig. 1(b)]. Here, ρ_{i}, E_f and $\mathcal{M}_\tau$ depend on $\boldsymbol{F}$ and $\boldsymbol{R}$. Qubits are measured in the computational basis, and the outcome is a binary vector $\boldsymbol{\mu}$. The corresponding measurement operator is $E_{\boldsymbol{\mu}}$. We consider the case that the computation result is the mean of a real function $f(\boldsymbol{\mu})$, then $E_f=\sum_{\boldsymbol{\mu}}f(\boldsymbol{\mu})E_{\boldsymbol{\mu}}$.

We can express the error-free map $\mathcal{M}_\tau^{\mathrm{ef}}$ as a product of frame gates $\mathcal{G}_\tau^{\mathrm{ef}}$ and single-qubit gates $\mathcal{R}_\tau^{\mathrm{ef}}$, i.e., $\mathcal{M}_\tau^{\mathrm{ef}}=\mathcal{G}_\tau^{\mathrm{ef}}\mathcal{R}_\tau^{\mathrm{ef}}$. For example, we have $\mathcal{G}_4^{\mathrm{ef}}=[I\otimes\Lambda_Z\otimes I]$ and $\mathcal{R}_4^{\mathrm{ef}}=[T\otimes I^{\otimes 3}]$. The actual map can always be expressed in the form $\mathcal{M}_\tau=\mathcal{J}_\tau\left(\mathcal{R}_\tau^{\mathrm{ef}}\otimes[\mathbb{1}_{\mathrm{E}}]\right)\mathcal{K}_\tau$. Here, $\mathbb{1}_{\mathrm{E}}$ is the identity operator of the environment, and $\mathcal{J}_\tau$ and $\mathcal{K}_\tau$ are maps on both the system and environment. $\mathcal{M}_\tau$ in this form is a linear map for matrix entries of $\mathcal{R}_\tau^{\mathrm{ef}}$. Therefore, we have the tensor form of the quantum computation $\mathrm{com}(\boldsymbol{F},\boldsymbol{R})=\mathrm{Tr}[(\overline{R}\otimes\overline{R}^*)F]$, where $\overline{R}$ is the tensor product of error-free single-qubit gates [e.g., $\overline{R}=H^{\otimes 3}\otimes T\otimes R_Z(\theta)\otimes\cdots$ in Fig. 1(a), in which gates are listed from top to bottom then left to right], and F is a tensor describing the effect of frame operations (see Ref. [62]). Errors are single-qubit-gate-independent (i.e., $\boldsymbol{R}$-independent) if ρ_{i}, E_f, $\mathcal{J}_\tau$ and $\mathcal{K}_\tau$ (for all τ) are constants. Then $\mathrm{Error}(\boldsymbol{F},\boldsymbol{R})^2$ is a homogeneous polynomial of degree 2 in both matrix elements of single-qubit gates and their Hermitian conjugates. The Clifford group is a unitary 2-design [14, 63, 64], and therefore $L_{\mathbb{U}}(\boldsymbol{F})=L_{\mathbb{C}}(\boldsymbol{F})$. We remark that, not only the second- but also the first- and third-order moments of the error distribution in unitary sampling can be evaluated using the Clifford sampling, because the Clifford group is also a 1-design and 3-design [65, 66].

In the Clifford sampling, we uniformly sample each single-qubit gate in the circuit from the Clifford group.

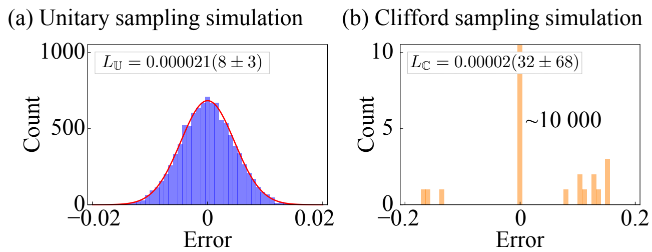

FIG. 2. Numerical results. (a) The unitary sampling of $\mathrm{Error}(\boldsymbol{F},\boldsymbol{R})$. The red curve denotes the Gaussian distribution $\mathcal{N}(0, L_{\mathbb{U}})$. (b) The Clifford sampling of $\mathrm{Error}(\boldsymbol{F},\boldsymbol{R})$. A ten-qubit circuit with a hundred two-qubit gates is used to compute the mean of a Pauli operator. The error rate per gate is 0.002.

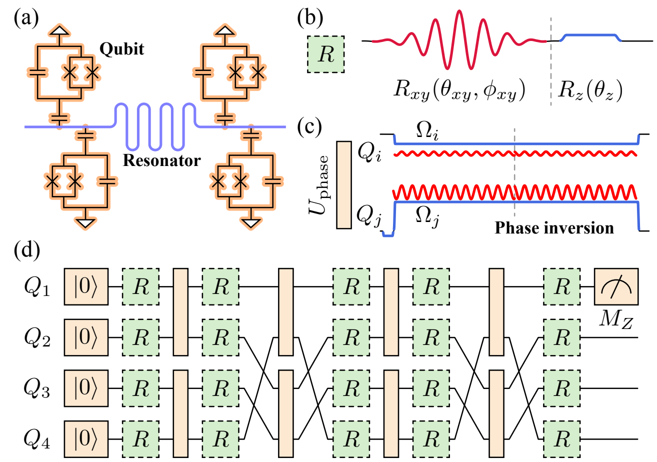

FIG. 3. Experiment setup. (a) Diagram of four Xmon qubits coupled to a central bus resonator. (b) A single-qubit gate is realized using a R_{xy} gate followed by a R_z gate. The R_{xy} gate has two parameters θ_{xy} and ϕ_{xy}, which are controlled by the amplitude and phase of XY pulse (red). The pulse length is 40 ns. The R_z gate has only one parameter θ_z, which is controlled by the amplitude of the Z pulse (blue), whose length is 10 ns. (c) Pulse sequences (square Z pulse in blue and sinusoidal microwave in red for each qubit) for the two-qubit gate U_{phase}. A short Z pulse is applied on one of the qubits before the long Z pulse to align the x-axes of their Bloch spheres. We note that the driving field (Ω_i on the qubit Q_i) is different in U_{phase} gates on different qubits. The length of U_{phase} is around 300 ns. Phases of driving fields are inverted at the middle of the gate. (d) The quantum circuit used to demonstrate the Clifford sampling.

We compute the error loss using the Monte Carlo summation method. There are two approaches. In the mean-value approach, we run each random circuit for multiple times on the actual quantum computer and in the simulation on a classical computer to estimate $\mathrm{com}(\boldsymbol{F},\boldsymbol{R})$ and $\mathrm{com}^{\mathrm{ef}}(\boldsymbol{F},\boldsymbol{R})$, respectively. Then, we can compute the error loss directly according to its definition. This approach is used in our experiments and numerical simulations. In the single-run approach, each sampled random circuit only runs for once or twice (see Ref. [62] for the detailed procedure). Then, the variance (due to finite sampling) of the $L_{\mathbb{C}}(\boldsymbol{F})$ estimator is upper bounded by $4\|E_f\|^4/N_{\mathrm{s}}$, where $4N_s$ circuit runs are implemented on the quantum computer and $4N_s$ circuit runs are classically simulated. We remark that usually the observable E_f scales polynomially with the qubit number in quantum simulation algorithms [3, 67].

Generalizations.—The error loss can be readily generalized to the multi-observable case, where one takes average of the quadratic error loss functions with different observables. One application is for characterizing the probability distribution of measurement outcomes in the computational basis by taking measurement operators $E_{\boldsymbol{\mu}}$ as observables. We can use this approach to assess the severity of error for any algorithm that qubits are measured in the computational basis. An alternative approach is the final-state fidelity loss [51]. We leave the detailed discussion in Ref. [62].

We conclude that $L_{\mathbb{U}}(\boldsymbol{F}) = L_{\mathbb{C}}(\boldsymbol{F})$ by assuming single-qubit-gate errors are gate-independent. If errors are gate-dependent, the equality does not hold in general. Then we can more reliably estimate $L_{\mathbb{U}}(\boldsymbol{F})$ by sampling hybrid circuits, in which not all single-qubit gates are Clifford and a few of them are general unitaries. In the first-order correction, $L_{\mathbb{U}}(\boldsymbol{F}) = \sum_{i=1}^{N_R} L_{\mathbb{H}_i} - (N_R - 1)L_{\mathbb{C}}$, where $\mathbb{H}_i$ denotes that only the i-th gate in the total N_R single-qubit gates is general unitary. We remark that Clifford-dominated circuits can be efficiently simulated using classical computer: The cost scales polynomially with circuit size given a fixed number of non-Clifford gates [68]. See Ref. [62] for more details.

Numerical results.—We implement numerical simulations of quantum circuits using QuESTlink [69, 70] for various error models, circuit frames, obervables and up to ten qubits, see Ref. [62]. Here, we only show results of the depolarizing model. In Fig. 2(a), we plot the distribution of $\mathrm{Error}(\boldsymbol{F},\boldsymbol{R})$ in the unitary sampling, and we can find that the distribution is approximately Gaussian. This conclusion holds in all our numerical simulations and experiments.

The error distribution is non-Gaussian in the Clifford sampling: the distribution is concentrated at several values of the error, and most of the probability is concentrated at zero, as shown in Fig. 2(b). We can understand this result as follows [51]. For Clifford circuits, if the observable to be measured E_f is a Pauli operator as in our case, $\mathrm{com}^{\mathrm{ef}}(\boldsymbol{F},\boldsymbol{R})$ takes three values 0 or ± 1. For most of the cases, $\mathrm{com}^{\mathrm{ef}}(\boldsymbol{F},\boldsymbol{R}) = 0$, and we always have $\mathrm{com}(\boldsymbol{F},\boldsymbol{R}) = 0$ if errors are Pauli, i.e., $\mathrm{Error}(\boldsymbol{F},\boldsymbol{R}) = 0$. Therefore, for Pauli errors, many Clifford circuits are error-insensitive [71], which is a cause of the non-Gaussian distribution.

Experimental results.—To demonstrate the feasibility and usefulness of Clifford sampling in an actual quantum computer, we implement it on a superconducting quantum device, which is illustrated in Fig. 3(a). Four frequency-tunable Xmon qubits ($Q_1 \sim Q_4$) are coupled to a central bus resonator, which mediates the effective interaction between qubits for implementing two-qubit gates. For every single-qubit gate $R \in \mathrm{U}(2)$, we can decompose it into two experimentally feasible gates $R = e^{i\alpha} R_z(\theta_z) R_{xy}(\theta_{xy}, \phi_{xy})$, as shown in Fig. 3(b), where α, θ_z, θ_{xy} and ϕ_{xy} are real numbers. For two-qubit gates, we use the Clifford dressed-state gate U_{phase}, which is generated by tuning two qubits into near resonance with microwave driving field being applied on each qubit dynamically [see Fig. 3(c)] and can be expressed as $\mathrm{diag}(1, -i, -i, 1)$ in the basis $\{|{++}\rangle, |{+-}\rangle, |{-+}\rangle, |{--}\rangle\}$ [72]. We can implement the gate U_{phase} between any pair of qubits, therefore we have six gate setups for four qubits. See Ref. [62] for device parameters and detailed implementation of gates.

Before applying Clifford sampling, we benchmarked fidelities of single-qubit gates and two-qubit gates. Fidelities of six U_{phase} setups measured using QPT are $95\% \sim 97\%$. In some circuits, two U_{phase} are applied in parallel. Because of crosstalk, gate fidelities are changed slightly in parallel operations. We use randomized benchmarking to measure gate fidelities of R_{xy} and R_z, which is implemented on each qubit individually as well as simultaneously on all qubits. Both approaches yield no less than 99.3% fidelities for X_π, Y_π, $X_{\pi/2}$, and Z_π gates, where $P_\theta = e^{-i\frac{\theta}{2}P}$ with P representing the Pauli operator. The average error rate of single-qubit gates is at least an order of magnitude lower than two-qubit gates, thus we can safely infer that most of the noise is introduced by U_{phase}. The gate performance can be improved by optimization based on Clifford sampling. See Ref. [62] for benchmarking and optimization data.

We use the circuit in Fig. 3(d) as an example to implement the Clifford sampling. The observable to be measured is the probability of Q_1 being in $|0\rangle$, i.e., $E_f = |0\rangle\langle 0|_1 = (I_1 + Z_1)/2$. Given a specific circuit $(\boldsymbol{F}, \boldsymbol{R})$, we run the circuit for 1000 times in order to estimate the probability in $|0\rangle$. The probability obtained in the experiment is P_0^{exp}, and its error-free value computed using the classical computer is P_0^{ef}. We note that P_0^{exp} has been corrected for readout errors [73]. Then, the computation error is $\text{Error} = P_0^{\text{exp}} - P_0^{\text{ef}}$.

Both unitary sampling and Clifford sampling are implemented in the experiment. For each case, 20000 random configurations of single-qubit gates $\boldsymbol{R}$ are generated. In the unitary sampling, the error distribution is Gaussian as shown in Fig. 4(a), the same as in numerical simulations. However, in the Clifford sampling, the error distribution is continuous as shown in Fig. 4(b), which is obviously different from the discrete distribution concentrated at several peaks shown in Fig. 2(b). In Ref. [62], we give numerical results of a composite error

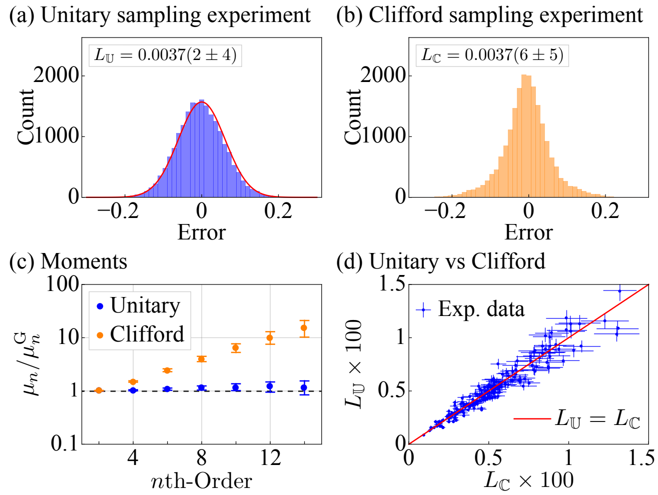

FIG. 4. Experimental results. (a) The unitary sampling of $\text{Error}(\boldsymbol{F}, \boldsymbol{R})$. The red curve denotes the Gaussian distribution $\mathcal{N}(0, L_{\mathbb{U}})$. (b) The Clifford sampling of $\text{Error}(\boldsymbol{F}, \boldsymbol{R})$. (c) Moments of two sampling approaches. $\mu_n = \mathrm{E}[\text{Error}^n]$ is the nth-order moment, and μ_n^{G} is the moment of $\mathcal{N}(0, L_{\mathbb{U}})$. The quantum circuit used to generate data in (a), (b) and (c) is in Fig. 3(d). (d) Error losses of unitary sampling versus Clifford sampling for 50 randomly generated frame operation configurations with one to ten layers of two-qubit gates. Each layer has one or two U_{phase} gates. When only one gate is applied, we protect the other two qubits from dephasing by applying dynamically drive fields [72]. The observable is $|0\rangle\langle 0|$ for one of four qubits. See Ref. [62] for details. The error bar denotes the standard deviation due to finite sampling.

model (a combination of coherent and amplitude damping errors) and the experimentally-measured model [74] (from QPT). The error distribution in Clifford sampling for these two models are in qualitative agreement with the experimental result. We plot moments up to the 14th-order in Fig. 4(c): Moments of the unitary sampling are consistent with the Gaussian distribution, and moments of the Clifford sampling are obviously larger beyond the second-order moment [71]. Although two distributions are different, their 2nd-order moments are the same up to the sampling noise.

In addition to the circuit in Fig. 3(d), we implemented the experiment for 50 randomly generated circuit frames $\boldsymbol{F}$. The error loss ($L_{\mathbb{U}}$ or $L_{\mathbb{C}}$) is estimated by sampling 500 single-qubit gate configurations $\boldsymbol{R}$ for each $\boldsymbol{F}$. The result is plotted in Fig. 4(d). Almost all data points are within 2σ (95.45% confidence interval) from the diagonal line, which represents $L_{\mathbb{U}} = L_{\mathbb{C}}$.

Discussion.—We propose to characterize quantum circuits executed on a noisy device by evaluating the quadratic error loss and fidelity loss using the Clifford sampling method. In these two loss functions, all the temporal and spatial correlations are automatically taken into account by treating the entire circuit as a whole. We demonstrate the Clifford sampling method with both numerical simulations and experiments on a superconducting device. We prove that fully-Clifford sampling is sufficient as long as the noise is independent of single-qubit gates. This conclusion holds as well in the presence of correlated readout error [75], see Ref. [62] for experimental result of a four-qubit observable. Weak gate dependence can be tackled using hybrid sampling. Experimental results do not show the significant effect of gate dependence. We observe a continuous distribution of the computation error in Clifford sampling in the experiment, whereas some simple error models such as Pauli error models predict a discrete distribution. This result suggests that these models cannot correctly describe the noise in our experiment. One can verify an error model and determine its parameters by modifying the error loss and replacing the ideal computation result with the result of error model.

In addition to characterizing quantum circuits, our method can find application in optimizing their performance. We experimentally implemented the optimization of a Rabi frequency driving two-qubit gates in a set of four-qubit four-depth circuits. The error losses decrease by more than 10% by using Clifford sampling, agreeing well with infidelity measured by QPT [76]. In the single-run approach, the sampling cost scales polynomially with the circuit size. Therefore, our method is promising in the multi-parameter optimization for large-scale quantum circuits. Other than optimizing parameters, our method can provide ground for choosing circuits. The circuit for a computation task may not be unique. Given a noisy quantum device, one can select a working circuit among theoretically equivalent circuits based on our loss functions. A similar idea was proposed in Ref. [56]. Compared to metrics assessing the general performance of the device, such as the quantum volume [32, 33, 48], our scheme is more application-oriented, i.e., each characterization experiment reflects the likelihood that the device performs well in solving a particular problem. Consequently, the optimization based on our characterization is tailored for specific problems and corresponding circuits.

We thank Wuxing Liu and Xu Zhang for discussions and technical support. YC thanks Tzu-Chieh Wei and Jason P. Kestner for insightful discussions. We are grateful to the authors of QuESTlink for making their package public. We acknowledge the support of the National Natural Science Foundation of China (No. 11725419 and No. 11875050), the National Key Research and Development Program of China (Grants No. 2017YFA0304300, No. 2019YFA0308100), the Zhejiang Province Key Research and Development Program (Grant No. 2020C01019) and the Basic Research Funding of Zhejiang University. YC is supported by the National Science Foundation (Grant No. PHY 1915165) and BNL LDRD #19-002. DYQ and YL are also supported by NSAF (Grant No. U1930403).

* Z. W. and Y. C. contributed equally to this work.
† chaosong@zju.edu.cn
‡ yli@gscaep.ac.cn

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Pr., 2010).
[2] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," Proceedings 35th Annual Symposium on Foundations of Computer Science (1994), 10.1109/sfcs.1994.365700.
[3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," Nat. Commun. **5**, 4213 (2014).
[4] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, "Quantum computational chemistry," Rev. Mod. Phys. **92**, 015003 (2020).
[5] 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).
[6] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver, "Noise spectroscopy through dynamical decoupling with a superconducting flux qubit," Nat. Phys. **7**, 565–570 (2011).
[7] J. Braumüller, L. Ding, A. P. Vepsäläinen, Y. Sung, M. Kjaergaard, T. Menke, R. Winik, D. Kim, B. M. Niedzielski, A. Melville, J. L. Yoder, C. F. Hirjibehedin, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "Characterizing and optimizing qubit coherence based on squid geometry," Phys. Rev. Applied **13**, 054079 (2020).
[8] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, "Optimal quantum control using randomized benchmarking," Phys. Rev. Lett. **112**, 240504 (2014).
[9] Y. Li and S. C. Benjamin, "Efficient variational quantum simulator incorporating active error minimization," Phys. Rev. X **7**, 021050 (2017).
[10] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," Phys. Rev. Lett. **119**, 180509 (2017).
[11] S. Endo, S. C. Benjamin, and Y. Li, "Practical quantum error mitigation for near-future applications," Phys. Rev. X **8**, 031027 (2018).
[12] J. Emerson, R. Alicki, and K. Życzkowski, "Scalable noise estimation with random unitary operators," J. Opt. B: Quantum Semiclassical Opt. **7**, S347–S352 (2005).
[13] 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).
[14] C. Dankert, R. Cleve, J. Emerson, and E. Livine, "Exact and approximate unitary 2-designs and their application to fidelity estimation," Phys. Rev. A **80**, 012304 (2009).
[15] E. Magesan, J. M. Gambetta, and J. Emerson, "Scal-

able and robust randomized benchmarking of quantum processes," Phys. Rev. Lett. **106**, 180504 (2011).

[16] J. M. Epstein, A. W. Cross, E. Magesan, and J. M. Gambetta, "Investigating the limits of randomized benchmarking protocols," Phys. Rev. A **89**, 062321 (2014).

[17] S. Kimmel, M. P. da Silva, C. A. Ryan, B. R. Johnson, and T. Ohki, "Robust extraction of tomographic information via randomized benchmarking," Phys. Rev. X **4**, 011050 (2014).

[18] 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).

[19] I. Roth, R. Kueng, S. Kimmel, Y.-K. Liu, D. Gross, J. Eisert, and M. Kliesch, "Recovering quantum gates from few average gate fidelities," Phys. Rev. Lett. **121**, 170502 (2018).

[20] T. J. Proctor, A. Carignan-Dugas, K. Rudinger, E. Nielsen, R. Blume-Kohout, and K. Young, "Direct randomized benchmarking for multiqubit devices," Phys. Rev. Lett. **123**, 030503 (2019).

[21] D. C. McKay, S. Sheldon, J. A. Smolin, J. M. Chow, and J. M. Gambetta, "Three-qubit randomized benchmarking," Phys. Rev. Lett. **122**, 200502 (2019).

[22] 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).

[23] J. F. Poyatos, J. I. Cirac, and P. Zoller, "Complete characterization of a quantum process: The two-bit quantum gate," Phys. Rev. Lett. **78**, 390–393 (1997).

[24] G. M. D'Ariano and P. Lo Presti, "Quantum tomography for measuring experimentally the matrix elements of an arbitrary quantum operation," Phys. Rev. Lett. **86**, 4195–4198 (2001).

[25] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, "Ancilla-assisted quantum process tomography," Phys. Rev. Lett. **90**, 193601 (2003).

[26] M. Mohseni and D. A. Lidar, "Direct characterization of quantum dynamics," Phys. Rev. Lett. **97**, 170501 (2006).

[27] M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, A. OâĂŹConnell, H. Wang, A. N. Cleland, and J. M. Martinis, "Process tomography of quantum memory in a josephson-phase qubit coupled to a two-level state," Nat. Phys. **4**, 523–526 (2008).

[28] D. Greenbaum, "Introduction to quantum gate set tomography," (2015), arXiv:1509.02921.

[29] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, "Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography," Nat. Commun. **8**, 14485 (2017).

[30] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, "Characterizing quantum supremacy in near-term devices," Nat. Phys. **14**, 595–600 (2018).

[31] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, *et al.*, "Quantum supremacy using a programmable superconducting processor," Nature **574**, 505–510 (2019).

[32] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, "Validating quantum computers using randomized model circuits," Phys. Rev. A **100**, 032328 (2019).

[33] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, L. Capelluto, O. GÃijnlÃijk, T. Itoko, N. Kanazawa, A. Kandala, G. A. Keefe, K. Krsulich, W. Landers, E. P. Lewandowski, D. T. McClure, G. Nannicini, A. Narasgond, H. M. Nayfeh, E. Pritchett, M. B. Rothwell, S. Srinivasan, N. Sundaresan, C. Wang, K. X. Wei, C. J. Wood, J.-B. Yau, E. J. Zhang, O. E. Dial, J. M. Chow, and J. M. Gambetta, "Demonstration of quantum volume 64 on a superconducting quantum computing system," (2020), arXiv:2008.08571.

[34] L. C. G. Govia, G. J. Ribeill, D. Ristè, M. Ware, and H. Krovi, "Bootstrapping quantum process tomography via a perturbative ansatz," Nat. Commun. **11**, 1084 (2020).

[35] J. Cotler and F. Wilczek, "Quantum overlapping tomography," Phys. Rev. Lett. **124**, 100401 (2020).

[36] M. R. Geller and M. Sun, "Efficient correction of multiqubit measurement errors," (2020), arXiv:2001.09980v1.

[37] K. E. Hamilton, T. Kharazi, T. Morris, A. J. McCaskey, R. S. Bennink, and R. C. Pooser, "Scalable quantum processor noise characterization," (2020), arXiv:2006.01805v1.

[38] 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).

[39] J. J. Wallman and S. T. Flammia, "Randomized benchmarking with confidence," New J Phys **16**, 103032 (2014).

[40] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak, "Nonexponential fidelity decay in randomized benchmarking with low-frequency noise," Phys. Rev. A **92**, 022326 (2015).

[41] H. Ball, T. M. Stace, S. T. Flammia, and M. J. Biercuk, "Effect of noise correlations on randomized benchmarking," Phys. Rev. A **93**, 022303 (2016).

[42] S. Mavadia, C. L. Edmunds, C. Hempel, H. Ball, F. Roy, T. M. Stace, and M. J. Biercuk, "Experimental quantum verification in the presence of temporally correlated noise," npj Quantum Inf. **4**, 7 (2018).

[43] K. Rudinger, T. Proctor, D. Langharst, M. Sarovar, K. Young, and R. Blume-Kohout, "Probing context-dependent errors in quantum processors," Phys. Rev. X **9**, 021045 (2019).

[44] A. Veitia and S. J. van Enk, "Testing the context-independence of quantum gates," (2018), arXiv:1810.05945v1.

[45] M. Huo and Y. Li, "Self-consistent tomography of temporally correlated errors," (2018), arXiv:1811.02734v3.

[46] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, "Gate-count estimates for performing quantum chemistry on small quantum computers," Phys. Rev. A **90**, 022305 (2014).

[47] N. Moll, A. Fuhrer, P. Staar, and I. Tavernelli, "Optimizing qubit resources for quantum chemistry simulations in second quantization on a quantum computer," J. Phys. A: Math. Theor. **49**, 295301 (2016).

[48] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M.

Gambetta, M. Ganzhorn, A. Kandala, A. Mezzacapo, P. Mï£ · ller, W. Riess, G. Salis, *et al.*, "Quantum optimization using variational algorithms on near-term quantum devices," Quantum Sci. Technol. **3**, 030503 (2018).

[49] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, "Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer," Quantum Sci. Technol. **4**, 045005 (2019).

[50] B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes, "Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm," npj Quantum Inf. **6**, 10 (2020).

[51] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, "Learning-based quantum error mitigation," (2020), arXiv:2005.07601v1.

[52] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1. 0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," Nat. Methods **17**, 261–272 (2020).

[53] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," Nature **549**, 242–246 (2017).

[54] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, "Supervised learning with quantum-enhanced feature spaces," Nature **567**, 209–212 (2019).

[55] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, "Error mitigation with clifford quantum-circuit data," (2020), arXiv:2005.10189v1.

[56] L. Cincio, K. Rudinger, M. Sarovar, and P. J. Coles, "Machine learning of noise-resilient quantum circuits," (2020), arXiv:2007.01210v1.

[57] We observe that the mean is approximately zero in all numerical simulations and experiments, in which the observable is a Pauli operator. In general, the mean can be measured using Clifford sampling and Hybrid sampling in the same way as the loss function.

[58] D. Gottesman, "The heisenberg representation of quantum computers," (1998), arXiv:quant-ph/9807006v1.

[59] S. Aaronson and D. Gottesman, "Improved simulation of stabilizer circuits," (2004), arXiv:quant-ph/0406196.

[60] S. Anders and H. J. Briegel, "Fast simulation of stabilizer circuits using a graph-state representation," Phys. Rev. A **73**, 022334 (2006).

[61] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2007).

[62] See Supplemental Material.

[63] D. Gross, K. Audenaert, and J. Eisert, "Evenly distributed unitaries: On the structure of unitary designs," J. Math. Phys. **48**, 052104 (2007).

[64] A. Roy and A. J. Scott, "Unitary designs and codes," Des. Codes Cryptogr. **53**, 13–31 (2009).

[65] Z. Webb, "The clifford group forms a unitary 3-design," (2015), arXiv:1510.02769v3.

[66] H. Zhu, "Multiqubit clifford groups are unitary 3-designs," Phys. Rev. A **96**, 062336 (2017).

[67] S. Lloyd, "Universal quantum simulators," Science **273**, 1073–1078 (1996).

[68] S. Bravyi and D. Gosset, "Improved classical simulation of quantum circuits dominated by clifford gates," Phys. Rev. Lett. **116**, 250501 (2016).

[69] T. Jones and S. Benjamin, "QuESTlink—mathematica embiggened by a hardware-optimised quantum emulator," Quantum Sci. Technol. **5**, 034012 (2020).

[70] T. Jones, A. Brown, I. Bush, and S. C. Benjamin, "QuEST and high performance simulation of quantum computers," Sci. Rep. **9**, 10736 (2019).

[71] These circuits may also be insensitive to non-Pauli errors, which may explain the non-Gaussian distribution of Clifford sampling in the experiment. This observation suggests that we can reduce the statistical fluctuation in evaluating $L_{\mathbb{C}}(\boldsymbol{F})$ using the importance sampling by selecting error-sensitive Clifford circuits.

[72] Q. Guo, S.-B. Zheng, J. Wang, C. Song, P. Zhang, K. Li, W. Liu, H. Deng, K. Huang, D. Zheng, X. Zhu, H. Wang, C.-Y. Lu, and J.-W. Pan, "Dephasing-insensitive quantum information storage and processing with superconducting qubits," Phys. Rev. Lett. **121**, 130501 (2018).

[73] See Supplemental Material [62] for the details of readout correction, which includes Refs. [77, 78].

[74] See Supplemental Material [62] for the details of constructing the experimentally-measured model, which includes Ref. [79].

[75] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, "Mitigating measurement errors in multi-qubit experiments," (2020), arXiv:2006.14044.

[76] See Supplemental Material [62] for the details of optimization procedure and results, which includes Ref. [80].

[77] C. Song, K. Xu, W. Liu, C.-p. Yang, S.-B. Zheng, H. Deng, Q. Xie, K. Huang, Q. Guo, L. Zhang, P. Zhang, D. Xu, D. Zheng, X. Zhu, H. Wang, Y.-A. Chen, C.-Y. Lu, S. Han, and J.-W. Pan, "10-qubit entanglement and parallel logic operations with a superconducting circuit," Phys. Rev. Lett. **119**, 180511 (2017).

[78] C. Song, K. Xu, H. Li, Y.-R. Zhang, X. Zhang, W. Liu, Q. Guo, Z. Wang, W. Ren, J. Hao, H. Feng, H. Fan, D. Zheng, D.-W. Wang, H. Wang, and S.-Y. Zhu, "Generation of multicomponent atomic schrödinger cat states of up to 20 qubits," Science **365**, 574–577 (2019).

[79] P. Pechukas, "Reduced dynamics need not be completely positive," Phys. Rev. Lett. **73**, 1060–1062 (1994).

[80] S. Li, A. D. Castellano, S. Wang, Y. Wu, M. Gong, Z. Yan, H. Rong, H. Deng, C. Zha, C. Guo, L. Sun, C. Peng, X. Zhu, and J.-W. Pan, "Realisation of high-fidelity nonadiabatic CZ gates with superconducting qubits," npj Quantum Inf. **5**, 84 (2019).