ARTICLE OPEN Cross-platform comparison of arbitrary quantum processes

Congcong Zheng $1^{,2}$, Xutao Yu^{1,2,3 \boxtimes} and Kun Wang^{4 \boxtimes}

In this work, we present a protocol for comparing the performance of arbitrary quantum processes executed on spatially or temporally disparate quantum platforms using Local Operations and Classical Communication (LOCC). The protocol involves sampling local unitary operators, which are then communicated to each platform via classical communication to construct quantum state preparation and measurement circuits. Subsequently, the local unitary operators are implemented on each platform, resulting in the generation of probability distributions of measurement outcomes. The max process fidelity is estimated from the probability distributions, which ultimately quantifies the relative performance of the quantum processes. Furthermore, we demonstrate that this protocol can be adapted for quantum process tomography. We apply the protocol to compare the performance of five quantum devices from IBM and the "Qianshi" quantum computer from Baidu via the cloud. The experimental results unveil two notable aspects: Firstly, the protocol adeptly compares the performance of the quantum processes implemented on different quantum computers. Secondly, the protocol scales, although still exponentially, much more favorably with the number of qubits, when compared to the full quantum process tomography. We view our work as a catalyst for collaborative efforts in cross-platform comparison of quantum computers.

npj Quantum Information (2024)10:4; https://doi.org/10.1038/s41534-023-00797-3

INTRODUCTION

As the field of quantum computing and quantum information gains traction, more manufacturers are entering the market, developing their own quantum computers. However, the current generation of noisy intermediate-scale quantum (NISQ) computers, despite their potential, is still hindered by quantum noise¹. A great challenge is how to compare the performance of the quantum computers fabricated by different manufacturers and located in different laboratories, termed as cross-platform comparison. This task is especially relevant when we move towards regimes where comparing to classical simulations becomes computationally challenging, and therefore a direct comparison of quantum computers is necessary.

A standard method to achieve cross-platform comparison is quantum tomography², in which we first reconstruct the full information of guantum computers under investigation and then estimate their fidelity from the obtained matrices. However, quantum tomography is known to be time-consuming and computationally difficult; even learning a few-gubit guantum state poses experimental challenges^{3,4}. A more efficient way is to compare guantum computers directly without resorting to the full information. Indeed, a variety of certification and benchmarking tools^{5,6}, such as fidelity estimation^{7–12}, quantum verification^{13–16}, and quantum benchmarking^{17–21}, have been developed along this way. However, these methods commonly assume that one can access a known and theoretical target usually simulated by classical computers. They quickly become inaccessible for quantum computers containing several hundreds or even thousands of highly entangled qubits, due to the exponentially increasing time complexity in simulating classical computers. Therefore, there is a pressing need for directly comparing unknown quantum states and processes across distinct devices situated at various locations and temporal instances²².

Recently, Elben et al.²³ proposed the first cross-platform protocol for estimating the fidelity of guantum states, possibly generated by spatially and temporally separated quantum computers. This protocol requires only local measurements in randomized product bases and classical communication between quantum computers. Numerical simulation shows that its sample complexity scales exponentially as $\mathcal{O}(2^{bn})$ with $b \leq 1$, where *n* is the number of qubit. This is significantly less than full quantum state tomography which has an exponent $b \ge 2$. It is applicable in state-of-the-art quantum computers consisting of a few tens of qubits²⁴. Later on, Knörzer et al.²⁵ extended Elben's proposal to cross-platform comparison of quantum networks, achieving a linear scaling of sample complexity at the cost of requiring quantum links. Nevertheless, a quantum link transferring quantum states of many qubits with high accuracy between two distant quantum computers is far from reach in the near future.

In this work, by elaborating the core idea of²³, we present a protocol for cross-platform comparing spatially and temporally separated quantum processes. The protocol uses only single-qubit unitary gates and classical communication between quantum computers, without requiring quantum links or ancilla qubits. This approach allows for accurate estimation of the performance of quantum devices manufactured in separate laboratories and companies using different technologies. Furthermore, the protocol can be used to monitor the stable function of target quantum computers over time. We apply the protocol to compare the performance of five quantum devices from IBM and the "Qianshi" quantum computer from Baidu, all via the cloud. The experimental results reveal that our protocol can accurately compare the performance of arbitrary quantum processes. The numerical analysis shows that our protocol scales exponentially as $\mathcal{O}(2^{bn})$ with $b \approx 2$, where *n* is the number of qubits. This is much more favorable concerning the system size when compared to full quantum process tomography, whose scaling exponent is $b \ge 4^{26}$.

¹State Key Lab of Millimeter Waves, Southeast University, Nanjing, China. ²Frontiers Science Center for Mobile Information Communication and Security, Southeast University, Nanjing, China. ³Purple Mountain Laboratories, Nanjing, China. ⁴Institute for Quantum Computing, Baidu Research, Beijing, China. ^{Seemail}: yuxutao@seu.edu.cn; wangkun28@baidu.com

There are many more efficient process tomographic methods^{27–30}. However, they either require a specific structure or a priori knowledge of the system of interest, which are not directly applicable for the cross-platform comparison task. Overall, our protocol serves as an application of the randomized measurement toolbox³¹.

RESULTS

Background

Cross-platform comparison of quantum states. In quantum information, fidelity is an important metric that is widely used to characterize the closeness between quantum states. There are many different proposals for the definition of state fidelity³². In this work, we will concentrate on the max fidelity, formally defined as^{23,32}

$$F_{\max}(\rho_1, \rho_2) := \frac{\operatorname{Tr}[\rho_1 \rho_2]}{\max\{\operatorname{Tr}[\rho_1^2], \operatorname{Tr}[\rho_2^2]\}},\tag{1}$$

where ρ_i is an *n*-qubit quantum state produced by the quantum computer, i = 1, 2.

Elben et al.²³ proposed a randomized measurement protocol to estimate F_{max} , which process is as follows. First, we construct an *n*qubit unitary $U = \bigotimes_{k=1}^{n} U_k$, where each U_k is identically and independently sampled from a single-qubit set \mathcal{X}_2 satisfying unitary 2-design^{33,34}. This information will be classically communicated to the quantum computers, possibly spatially or temporally separated, that produce the quantum states ρ_1 and ρ_2 , respectively. Then, each quantum computer executes the unitary U, performs a computational basis measurement, and records the measurement outcome s. Repeating the above procedure for fixed U a number of times, we are able to obtain two probability distributions over the outcomes of the form $Pr_U^{(1)}, Pr_U^{(2)}$, where the superscript *i* indicates that the distribution is obtained from quantum state ρ_i . Next, we repeat the whole procedure for many different random unitaries U, yielding a set of probability distributions $\{Pr_{U}^{(1)}, Pr_{U}^{(2)}\}_{U}$. From the experimental data, we estimate the overlap between ρ_i and ρ_i as²

$$\operatorname{Tr}[\rho_i \rho_j] = 2^n \sum_{\mathbf{s}, \mathbf{s}' \in \{0,1\}^n} (-2)^{-\mathcal{D}[\mathbf{s}, \mathbf{s}']} \overline{\operatorname{Pr}_U^{(i)}[\mathbf{s}]} \operatorname{Pr}_U^{(j)}[\mathbf{s}']},$$
(2)

where $\overline{\cdots}$ denotes the ensemble average over the sampled unitaries U and $\mathcal{D}[\mathbf{s}, \mathbf{s}']$ denotes the hamming distance between two bitstrings \mathbf{s} and \mathbf{s}' . Specially, $\text{Tr}[\rho_1\rho_2]$ can be estimated from Eq. (2) by setting i = 1 and j = 2, whereas the purities $\text{Tr}[\rho_1^2]$ and $\text{Tr}[\rho_2^2]$ can be obtained by setting i = j = 1 and i = j = 2, respectively. Using the above estimated quantities, we successfully compute the max fidelity $F_{\max}(\rho_1, \rho_2)$.

Using experimental data from³⁵, Elben et al. showcased the experiment-theory fidelities and experiment-experiment fidelities of highly entangled quantum states prepared via quench dynamics in a trapped ion quantum simulator as a proof of principle²³. Recently, Zhu et al. reported a thorough cross-platform comparison of quantum states in four ion-trap and five superconducting quantum platforms, with detailed analysis of the results and an intriguing machine learning approach to explore the data²⁴.

Quantum process performance metric. A quantum process, also known as a quantum operation or a quantum channel, is a mathematical description of the evolution of a quantum system. It is mathematically formulated as a completely positive and trace-preserving (CPTP) linear map on the quantum states³⁶. The Choi-Jamiołkowski isomorphism provides a unique way to represent quantum processes as quantum states in a larger Hilbert space. Formally, the Choi state of an *n*-qubit quantum process \mathcal{E} is defined as³⁷

$$\eta_{\mathcal{E}} := (\mathcal{I} \otimes \mathcal{E}) |\psi_{+}\rangle \langle \psi_{+}|, \tag{3}$$

where \mathcal{I} is the identity channel and $|\psi_+\rangle := 1/\sqrt{2^n}\sum |ii\rangle$ is a maximally entangled state of a bipartite quantum system composed of two *n*-qubit subsystems.

One lesson we can learn from the cross-platform state comparison protocol is that we must choose a process metric before comparing two quantum processes. Gilchrist et al.³⁸ introduced a systematic way to generalize a metric originally defined on quantum states to a corresponding metric on quantum processes, utilizing the Choi-Jamiołkowski isomorphism. Specifically, the max fidelity between two *n*-qubit quantum processes \mathcal{E}_1 and \mathcal{E}_2 , implemented on different quantum platforms, is defined as

$$F_{\max}(\mathcal{E}_1, \mathcal{E}_2) := F_{\max}(\eta_{\mathcal{E}_1}, \eta_{\mathcal{E}_2}),\tag{4}$$

where $\eta_{\mathcal{E}}$ is the Choi state of quantum process \mathcal{E} . Thanks to the Choi-Jamiołkowski isomorphism, this metric fulfills the axioms for process fidelities^{32,38}, i.e., (i) $F_{\max}(\mathcal{E}_1, \mathcal{E}_2) \leq 1$ with $F_{\max}(\mathcal{E}_1, \mathcal{E}_2) = 1$ if and only if $\mathcal{E}_1 = \mathcal{E}_2$; and (ii) $F_{\max}(\mathcal{E}_1, \mathcal{E}_2) = F_{\max}(\mathcal{E}_2, \mathcal{E}_1)$. Thus, it can be used to verify whether two quantum devices have implemented the same quantum process.

In this work, we propose an experimentally efficient protocol to estimate this metric. This protocol makes use of only single-qubit unitaries and classical communication, thus can be executed in spatially and temporally separated quantum devices. This enables cross-platform comparison of arbitrary quantum processes.

Theories

We first provide a simple example to illustrate the necessity of cross-platform comparison. Then, we introduce a protocol for estimating the max process fidelity that is conceptually straightforward yet experimentally challenging. Next, we propose a modification to the protocol that employs randomized input states and provide a detailed explanation of the approach. Our protocol is motivated by the observation that even identical quantum computers cannot produce identical outcomes on each run due to the intrinsic randomness of quantum mechanics, but they do generate identical probability distributions from a statistical perspective.

Cross-platform comparison of quantum computers is essential for at least two reasons. Firstly, comparing the actual implementation with an idealized theoretical simulation can be challenging, especially as classical simulations become computationally demanding with an increasing number of gubits. Secondly, due to the presence of varying forms of quantum noise across different quantum platforms, the actual implementation of quantum processes can vary significantly, even if they maintain the same process fidelity with respect to the ideal target. To illustrate this point, consider the following example. Suppose Alice has a superconducting quantum computer and Bob has a trapped-ion quantum computer. They independently implement the single-qubit Hadamard gate $\mathcal{H}(\rho) = H\rho H^{\dagger}$ on their respective quantum computers. However, Alice's implementation \mathcal{E}_1 suffers from the depolarizing noise, vielding $\mathcal{E}_1(\rho) = (1 - p_1)H\rho H^{\dagger} + p_1 1/2$, where $p_1 = 7/30$ and 1 is the identity matrix. On the other hand, Bob's implementation \mathcal{E}_2 suffers from the dephasing noise. such that $\mathcal{E}_2(\rho) = (1 - p_2) H \rho H^{\dagger} + p_2 \Delta(\rho)$, where $p_2 = 1/5$ and $\Delta(\cdot)$ is the dephasing operation. After simple calculations, we obtain $F_{\max}(\mathcal{E}_1, \mathcal{H}) = F_{\max}(\mathcal{E}_2, \mathcal{H}) \approx 0.808$ and $F_{\max}(\mathcal{E}_1, \mathcal{E}_2) \approx 0.978$. Despite \mathcal{E}_1 and \mathcal{E}_2 having the same fidelity level when compared to the ideal target \mathcal{H}_r a discernible difference exists between them. Therefore, solely comparing the fidelity of a quantum process to an ideal reference is insufficient, and a direct comparison between quantum processes is warranted.

Ancilla-assisted cross-platform comparison. Firstly, we propose a conceptually simple approach for estimating F_{max} defined in



Fig. 1 Two protocols to estimate the max process fidelity F_{max} between quantum processes implemented on different quantum platforms. a Ancilla-assisted protocol: Prepare the maximally entangled state, execute the target quantum process, and perform the randomized measurements given by $\bigotimes_{k=1}^{n} U_{1}^{(k)} \otimes \bigotimes_{k=1}^{n} U_{2}^{(k)}$. b Ancilla-free protocol: Randomly sample a computational basis $|s\rangle$, execute the unitaries $\bigotimes_{k=1}^{n} U_{1}^{(k)T}$, execute the target quantum process, and perform the randomized measurements given by $\bigotimes_{k=1}^{n} U_{2}^{(k)}$. c Run the quantum circuits constructed in **a** or **b** on platform S_i to obtain the probability distribution $\Pr_{U}^{(i)}[s, k]$. The max process fidelity $F_{max}(\mathcal{E}_i, \mathcal{E}_j)$ is inferred from the probability distributions (see text).

Eq. (4). The key observation is that F_{max} can be seen as the max state fidelity between the Choi states of the corresponding quantum processes. Hence, we can generalize the quantum state comparison protocol in the section "Cross-platform comparison of quantum states" to achieve quantum process comparison. This approach has been independently proposed in²⁵.

In the first step, we construct the Choi states of the target quantum process $\mathcal E$ in two quantum platforms as follows. First, we introduce an additional *n*-qubit clean auxiliary system in each platform. Then, we prepare a 2*n*-qubit maximally entangled state $|\psi_+\rangle$ using the auxiliary system. Finally, we apply $\mathcal E$ to half of the whole system, successfully preparing the Choi state of $\mathcal E$. In the second step, we estimate the max state fidelity of these two Choi states using the procedure introduced in the section "Cross-platform comparison of quantum states". The complete protocol is

illustrated in Fig. 1a–c.

We refer to this protocol as the ancilla-assisted cross-platform comparison because it requires additional clean ancilla qubits to prepare the Choi state of the quantum process. To perform this protocol, a maximally entangled state is required as input, resulting in a two-fold overhead when comparing 2*n*-qubit states instead of *n*-qubit states. Consequently, this protocol may not be practical in scenarios with limited quantum computing resources. Furthermore, preparing high-fidelity maximally entangled states can be experimentally challenging, which may negatively impact the accuracy of the protocol.

Ancilla-free cross-platform comparison. To overcome the limitations of the ancilla-assisted protocol, we propose an efficient and ancilla-free approach for estimating the max process fidelity F_{max} . Our protocol does not require any additional qubits or the preparation of maximally entangled states. The key observation is that the auxiliary system in the ancilla-assisted protocol only needs to perform randomized measurements. After the measurement, the auxiliary system collapses to one eigenstate of the sampled measurement operator. Based on the identity $(|u\rangle\langle u|\otimes 1)|\psi_+\rangle = |u\rangle\otimes |u^*\rangle$, where 1 is the identity matrix, and the deferred measurement principle³⁹, we can eliminate the auxiliary system by preparing computational states and applying the transposed unitary operator on the main system. Please refer to Supplementary Note 1 for a detailed analysis.

We refer to this protocol as the ancilla-free cross-platform comparison and it works as follows. We consider two *n*-qubit quantum processes \mathcal{E}_1 and \mathcal{E}_2 realized on different quantum platforms \mathcal{S}_1 and \mathcal{S}_2 , whose Choi states are η_1 and η_2 , respectively. The protocol, illustrated in Fig. 1b–c, consists of three main steps: sampling unitaries, running circuits, and post-processing.

- 1. Sampling unitaries: Construct two *n*-qubit unitaries $U_i = \bigotimes_{k=1}^{n} U_i^{(k)}$, i = 1, 2, where each $U_i^{(k)}$ is identically and independently sampled from a single-qubit set \mathcal{X}_2 satisfying unitary 2-design^{33,34}. The information of U_i is then communicated to both platforms via classical communication.
- 2. Running circuits: After receiving the information of the sampled unitaries, each platform S_i (i = 1, 2) initializes its quantum system to the computational states $|\mathbf{s}\rangle$ and applies the first unitary U_1 to $|\mathbf{s}\rangle$. Subsequently, S_i implements the quantum process \mathcal{E}_i and applies the second unitary U_2 . Finally, S_i performs the projective measurement in the computational basis and obtains an outcome \mathbf{k} . Repeating the above procedure many times, we obtain two probability distributions $\Pr_{\mathbf{k}|\mathbf{s},U_1,U_2}^{(1)}$ and $\Pr_{\mathbf{k}|\mathbf{s},U_1,U_2}^{(2)}$ over the measurement outcomes \mathbf{k} for the fixed computational state $|\mathbf{s}\rangle$ and unitaries U_1 and U_2 . By exhausting the computational states and repeatedly sampling the unitaries, we obtain two probability distributions $\Pr_{\mathbf{k},\mathbf{S},|U_1,U_2}^{(i)}$ with respect to the

sampled unitaries and computational state inputs. For simplicity, we abbreviate $\Pr_{K,S|U_1,U_2}^{(i)}$ to $\Pr_U^{(i)}$.

3. Post-processing: From the experimental data, we estimate the overlap between the Choi states η_i and η_j for i, j = 1, 2 as

$$\operatorname{Tr}[\eta_{i}\eta_{j}] = 4^{n} \sum_{\boldsymbol{s}, \boldsymbol{s}', \boldsymbol{k}, \boldsymbol{k}' \in \{0, 1\}^{n}} (-2)^{-\mathcal{D}[\boldsymbol{s}, \boldsymbol{s}'] - \mathcal{D}[\boldsymbol{k}, \boldsymbol{k}']} \times \overline{\operatorname{Pr}_{U}^{(i)}[\boldsymbol{s}, \boldsymbol{k}] \operatorname{Pr}_{U}^{(j)}[\boldsymbol{s}', \boldsymbol{k}']}.$$
(5)

where $\overline{\cdots}$ denotes the ensemble average over the sampled unitaries U_1 and U_2 . This is proven in Supplementary Note 1. By setting i = 1 and j = 2, we can estimate the overlap $\text{Tr}[\eta_1 \eta_2]$ from the above equation, which is the second-order cross-correlation of the probabilities $\Pr_U^{(1)}$ and $\Pr_U^{(2)}$. We can obtain the purities $\text{Tr}[\eta_1^2]$ and $\text{Tr}[\eta_2^2]$ by setting i = j = 1 and i = j = 2, respectively. These are the second-order autocorrelations of the probabilities. Using the estimated quantities, we compute the max process fidelity $F_{\text{max}}(\mathcal{E}_1, \mathcal{E}_2)$ in Eq. (4).

There are several important points to note about our protocol. First, when classical simulation is available, the protocol can be used to compare the experimentally implemented process to the theoretical simulation, providing a useful tool for experiment-theory comparison. Second, our protocol can also estimate the process purity $Tr[\eta_{\mathcal{E}}^2]$ of a quantum process \mathcal{E} , which measures the extent to which \mathcal{E} preserves the purity of the quantum state. This is an important measure for characterizing quantum processes, and our protocol provides an efficient way to estimate it. Finally, it is worth noting that the definition of max process fidelity is not unique, and different approaches exist^{38,40}. Our protocol, based on statistical correlations of randomized inputs and measurements, can be readily extended to any metric that depends solely on the process overlap $Tr[\eta_{\mathcal{E}_1}\eta_{\mathcal{E}_2}]$ and the process purities $Tr[\eta_{\mathcal{E}_1}^2]$ and $Tr[\eta_{\mathcal{E}_2}^2]$. This makes our protocol highly versatile and applicable to a wide range of quantum computing scenarios.



Fig. 2 The performance matrices for the single-qubit H and two-qubit CNOT gates generated from seven different quantum platforms. The entry in the *i*-th row and *j*-th column of the matrix represents the max process fidelity between platform-*i* and platform-*j*. The entries in the upper right corner are visualized in pie chart format. **a** The performance matrix of the H gate. Each entry is inferred from $2^1 \cdot N_U = 20$ random circuits and each circuit is repeated $M_{\text{shots}} = 500$ times. **b** The performance matrix of the CNOT gate. Each entry is inferred from $2^2 \cdot N_U = 20$ random circuits and each circuit is repeated $M_{\text{shots}} = 500$ times.

Experiments

In this section, we report experimental results on cross-platform comparison of quantum processes across various spatially and temporally separated quantum devices. First, we demonstrate the efficacy of our protocol in comparing the H and CNOT gates implemented on different platforms with their ideal counterparts are obtained from classical simulation. Next, we use our protocol to monitor the stability of the "Qianshi" quantum computer from Baidu over a week. Finally, we conduct an extensive numerical analysis to determine the expected number of experimental runs required to obtain reliable results. All the experiments are conducted using the Quantum Error Processing toolkit developed on the Baidu Quantum Platform⁴¹.

Before presenting the experimental results, we summarize the working procedure shared among different experiments. For a target *n*-qubit quantum process under investigation, the protocol in the section "Ancilla-free cross-platform comparison" constitutes three steps:

- 1. Randomly sample N_U number of *n*-qubit unitaries composed of single-qubit unitaries. Note that the random Pauli basis measurements {*X*, *Y*, *Z*} are equivalent to randomized measurements with a single qubit Clifford group^{24,42}. The Clifford group is a unitary 2-design group and it can be employed to achieve complete process tomography of *n*qubit quantum processes. This equivalence enables us to sample directly from the 3^{*n*} Pauli preparation and 3^{*n*} Pauli measurement unitaries in our experiments.
- 2. Prepare 2^n number of computational basis states as inputs, evolve them with the randomly sampled unitary and the target quantum process, and measure the final state in the computational basis measurement with M_{shots} number of shots.
- 3. From the measurement statistics, we estimate the fidelity using Eq. (5).

The sample complexity of the above procedure is captured by the total number of experimental runs, $2^n \times N_U \times M_{shots}$.

Comparing spatially separated quantum processes. We utilize our ancilla-free cross-platform comparison protocol to assess the performance of H and CNOT gates implemented on seven distinct platforms that are freely accessible to the public over the internet. They include six superconducting quantum computers, namely *ibmq_quito* (IBM_1), *ibmq_oslo* (IBM_2), *ibmq_lima* (IBM_3), *ibm_nairobi* (IBM_4), *ibmq_manila* (IBM_5), and *baidu_qianshi* (BD_1), as well as the *baidu ideal simulator* (IDEAL), which is for experiment-theory comparisons.

We set $N_U = 10$ and $M_{shots} = 500$ to compare the performance of the single-qubit H gate across the above seven quantum platforms. Since the CNOT gate is a two-qubit gate, we set $N_U = 100$ and $M_{shots} = 500$ to compare the performance of CNOT gate across the above seven quantum platforms. The performance matrices for the H and CNOT gates are presented in Fig. 2. The experimental results make it clear that, while some quantum devices may achieve fidelities that are comparable to those of the ideal simulator, there remains a significant discrepancy between them. This emphasizes the importance of directly comparing the performance of quantum devices with each other, rather than relying solely on comparisons to an ideal simulator, as such comparisons may not be adequate.

Comparing temporally separated quantum processes. Our protocol is also useful for monitoring the stable performance of quantum devices over time. To this end, we employ the ancillafree cross-platform comparison protocol to assess the stability of H and CNOT gates implemented on Baidu's "Qianshi" quantum computer (BD_1) over the course of one week. The experiment is conducted every day at 14:00 and lasts about 4 hours, which is asynchronous with the daily calibration procedure usually performed at midnight. The experimental settings for the H and CNOT gates are identical to those used in the previous section. Specifically, for the single-qubit H gate, we set $N_U = 10$ and $M_{shots} = 500$. For the two-qubit CNOT gate, we set $N_U = 100$ and $M_{shots} = 500$. The performance matrices of the H and CNOT gates generated from the daily data of "Qianshi" are shown in Fig. 3.



Fig. 3 The performance matrices of the single-qubit H and two-qubit CNOT gates generated from the daily data of Baidu's "Qianshi" quantum computer for one week. The entry in the *i*-th row and *j*-th column of the matrix represents the max process fidelity between platform-*i* and platform-*j*. The entries in the upper right corner are visualized in pie chart format. **a** The performance matrix of the H gate. Each entry is inferred from $2^1 \cdot N_U = 20$ random circuits and each circuit is repeated $M_{shots} = 500$ times. **b** The performance matrix of the CNOT gate. Each entry is inferred from $2^2 \cdot N_U = 400$ random circuits and each circuit is repeated $M_{shots} = 500$ times.



Fig. 4 Average statistical error $|\tilde{F}_{max} - 1|$ as a function of the total number of experimental runs $2^n N_U M_{shots}$. The target quantum process is taken to be the H gate for n = 1 and the CNOT gate for n = 2. The green lines obey ~ $1/(2^n N_U M_{shots})$ and are guides for the eye. The data are obtained via numerical simulation.

After analyzing the cross-platform fidelities presented in Fig. 3, we discover several noteworthy features. First, we observe that the stability of the H gate is considerably higher than that of the CNOT gate on "Qianshi," which aligns with the expectation that twoqubit gates are harder to implement and maintain in a superconducting quantum computer than single-qubit gates. Additionally, on the last day of the week (DAY_7), there is a significant drop in the performance of the CNOT gate. After consulting with researchers from Baidu's Quantum Computing Hardware Laboratory, it is determined that the instability is caused by the sudden halt of the dilution cooling system. After the system is restarted, all native quantum gates have to be re-calibrated to achieve optimal performance. Furthermore, it is observed that the temperature variation had a negligible impact on the H gate. This observation might be helpful for the experimenters to identify potential hardware issues.

Scaling of the required number of experimental runs. In practice, the accuracy of the estimated fidelity is unavoidably subject to statistical error, as a result of the finite number of random circuits $(2^n \times N_U)$ and the finite number of computational basis measurements (M_{shots}) performed per random circuit. Therefore, it is experimentally crucial to consider the scaling of the total number of experimental runs $2^n \times N_U \times M_{shots}$, which represents the measurement budget, in order to effectively suppress the statistical error to a prespecified threshold ϵ when evaluating the performance of an *n*-qubit quantum process. In the following, we present numerical simulation to investigate this behavior.

In Fig. 4, numerical results for the average statistical error as a function of the measurement budget $2^n \times N_U \times M_{shots}$ are presented, and the scaling of the measurement budget with respect

to the system size *n* is derived. In order to keep consistent with previous experiments, we choose the H gate when n = 1 and the CNOT gate when n = 2 in the simulation. Note that in this case the ideal fidelity $F_{max} = 1$ is known. We repeat our protocol on the ideal simulator 5 times for each point in the figure and record the mean of the statistical errors $|\tilde{F}_{max} - 1.0|$. We find that the statistical error scales as $|\tilde{F}_{max} - 1.0| \sim 1/(2^n N_U M_{shots})$, where \tilde{F}_{max} is the estimated max process fidelity via simulation.

Now we investigate the scaling of the required number of experimental runs, $2^n \times M_{\text{shots}}$, per unitary to estimate the max fidelity F_{max} within an average statistical error of $\epsilon = 0.05$ while fixing N_U to 100. We employ our protocol to two very different types of quantum processes with different numbers of qubits n: (i) a highly entangled quantum process corresponding to an *n*-qubit GHZ state preparation circuit (Entangled) and (ii) a completely local quantum process composed of *n* single-qubit rotation gates (Non-Entangled). The numerical results are presented in Fig. 5. From the fitted data, we observe that $2^n \times M_{\text{shots}} \sim 2^{bn}$, where $b = 2.05 \pm 2e-5$ for the entangled case and $b = 1.89 \pm 2e-4$ for the non-entangled case. This scaling analysis closely aligns with the previous findings in cross-platform state comparisons²³ and reveals that the sample complexity of our ancilla-free crossplatform protocol scales as $2^n \times N_U \times M_{shots} \sim \mathcal{O}(2^{bn})$ with $b \approx 2$. This scaling, although exponential, is significantly better than full quantum process tomography (QPT), which has an exponent $b \ge 4^{26}$.

DISCUSSION

We have proposed an ancilla-free cross-platform protocol that enables the performance comparison of arbitrary quantum



Fig. 5 Scaling of the minimal number of required experimental runs $2^n M_{shots}$ to estimate \tilde{F}_{max} up to a fixed statistical error of 0.05 as a function of the number of qubits *n*. The number of random unitaries is fixed to $N_U = 100$. The target quantum process is taken to be the *n*-qubit GHZ state preparation circuit for the entangled case and the rotation circuit composed of *n* single-qubit rotation gates for the nonentangled case. The data is obtained via numerical simulation and the error bars are generated from five independent experiments.

processes using only single-qubit unitaries and classical communication. This protocol is suitable for comparing quantum processes that are independently manufactured over different times and locations, built by different teams using different technologies. We have experimentally demonstrated the crossplatform protocol on six remote quantum computers fabricated by IBM and Baidu, and monitored the stable functioning of Baidu's "Qianshi" quantum computer over one week. The experimental results reveal that our protocol accurately compares the performance of different quantum computers with significantly fewer measurements than quantum process tomography. Additionally, we have shown that our protocol applies to quantum process tomography.

However, some problems must be further explored to make the cross-platform protocols more practical. Firstly, the sample complexity of these protocols lacks theoretical guarantees, necessitating the empirical selection of experimental parameters. References^{43,44} might be a good starting point for addressing this challenge, where the authors gave analytically scaling law of the statistical errors with global random unitaries. Secondly, it is vital to make the protocols robust against state preparation and measurement errors. One possible solution is to apply quantum error mitigation methods^{45–49} to alleviate quantum errors and increase the estimation accuracy. We suggest that ideas from randomized benchmarking²¹ and quantum gateset tomography⁵⁰ might also be helpful for designing error robust cross-platform protocols.

METHODS

Randomized quantum process tomography

Here we show that the protocols proposed in the "Results" section can be utilized to achieve full quantum process tomography. This idea is motivated by Ref. ⁵¹, in which they proposed a method for

performing full quantum state tomography using randomized measurements.

For an *n*-qubit quantum process \mathcal{E} , we can first construct the Choi state of \mathcal{E} and then use the proposed protocol to obtain the full information of the Choi state $\eta_{\mathcal{E}}$. However, as previously mentioned, this method is inefficient and is impractical due to the imperfect preparation of maximally entangled states and the requirement for an additional *n*-qubit auxiliary system. Likewise, we may use the randomized input states trick introduced in the section "Ancilla-free cross-platform comparison" to overcome the above issues. Specifically, based on the experimental data \Pr_U collected in the section "Ancilla-free cross-platform comparison", the complete information of an unknown *n*-qubit quantum process \mathcal{E} can be obtained via

$$\eta_{\mathcal{E}} = 4^{n} \sum_{\boldsymbol{s}, \boldsymbol{s}', \boldsymbol{k}, \boldsymbol{k}' \in \{0,1\}^{n}} (-2)^{-\mathcal{D}[\boldsymbol{s}, \boldsymbol{s}'] - \mathcal{D}[\boldsymbol{k}, \boldsymbol{k}']} \times \overline{\Pr_{\boldsymbol{U}}[\boldsymbol{s}, \boldsymbol{k}] \boldsymbol{U}^{\dagger} | \boldsymbol{s}' \boldsymbol{k}' \rangle \langle \boldsymbol{s}' \boldsymbol{k}' | \boldsymbol{U}},$$
(6)

where $U = U_1 \otimes U_2$ and $\overline{\cdots}$ denotes the ensemble average over the sampled unitaries U_1 and U_2 as before. This is proven in Supplementary Note 2.

Comparison with previous works

Knörzer et al.²⁵ have recently introduced a set of protocols that enable pair-wise comparisons between distant nodes in a quantum network. Their contributions are twofold. First, they proposed four cross-platform state comparison schemes as alternatives to Elben's protocol²³, each relying on the presence of quantum links. Second, based on these state comparison protocols, they designed three cross-platform quantum process comparison protocols M1, M2, and M3.

We explain how our process comparison protocols differ from M1, M2, and M3. While M1 involves an ancilla-assisted comparison protocol that we have rephrased in the section "Ancilla-assisted

cross-platform comparison", our protocol does not rely on ancilla qubits. M3 involves a series of entanglement tests that are fundamentally different from our protocol. M2 and our protocol do share similarities, such as the absence of ancilla qubits and the need to sample random unitaries and computational basis states, but there are notable differences. Specifically, our protocol only samples from a single-qubit unitary 2-design and accurately estimates the max fidelity. On the other hand, M2(i) estimates the average gate fidelity and requires sampling from a multi-qubit unitary 2-design, which can be resource-intensive as the number of qubits increases. M2(iii) is conceptually straightforward but its estimated quantity is basisdependent, while the other fidelities investigated in this paper are all basis-independent. Additionally, M2(i) and M2(iii) are limited to comparing the performance of unitary quantum processes.

DATA AVAILABILITY

Data that support the plots and other findings of this study are available from the corresponding authors upon reasonable request.

CODE AVAILABILITY

Code that supports the findings of this study is available in the Quantum Error Processing toolkit developed on the Baidu Quantum Platform and from the corresponding authors upon reasonable request.

Received: 25 March 2023; Accepted: 8 December 2023; Published online: 03 January 2024

REFERENCES

- 1. Preskill, J. Quantum computing in the Nisq era and beyond. *Quantum* **2**, 79 (2018).
- O'Donnell, R. & Wright, J. Efficient quantum tomography. In Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing, STOC 16, 899–912 (Association for Computing Machinery, New York, NY, USA, 2016).
- Häffner, H. et al. Scalable multiparticle entanglement of trapped ions. Nature 438, 643–646 (2005).
- Carolan, J. et al. On the experimental verification of quantum complexity in linear optics. Nat. Photonics 8, 621–626 (2014).
- Eisert, J. et al. Quantum certification and benchmarking. Nat. Rev. Phys. 2, 382–390 (2020).
- Kliesch, M. & Roth, I. Theory of quantum system certification. PRX Quantum 2, 010201 (2021).
- Hofmann, H. F. Complementary classical fidelities as an efficient criterion for the evaluation of experimentally realized quantum operations. *Phys. Rev. Lett.* 94, 160504 (2005).
- Bendersky, A., Pastawski, F. & Paz, J. P. Selective and efficient estimation of parameters for quantum process tomography. *Phys. Rev. Lett.* **100**, 190403 (2008).
- 9. Flammia, S. T. & Liu, Y.-K. Direct fidelity estimation from few pauli measurements. *Phys. Rev. Lett.* **106**, 230501 (2011).
- da Silva, M. P., Landon-Cardinal, O. & Poulin, D. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.* **107**, 210404 (2011).
- Reich, D. M., Gualdi, G. & Koch, C. P. Optimal strategies for estimating the average fidelity of quantum gates. *Phys. Rev. Lett.* **111**, 200401 (2013).
- 12. Greenaway, S., Sauvage, F., Khosla, K. E. & Mintert, F. Efficient assessment of process fidelity. *Phys. Rev. Res.* **3**, 033031 (2021).
- Pallister, S., Linden, N. & Montanaro, A. Optimal verification of entangled states with local measurements. *Phys. Rev. Lett.* **120**, 170502 (2018).
- Zhu, H. & Hayashi, M. Efficient verification of pure quantum states in the adversarial scenario. *Phys. Rev. Lett.* **123**, 260504 (2019).
- Wang, K. & Hayashi, M. Optimal verification of two-qubit pure states. *Phys. Rev. A* 100, 032315 (2019).
- Yu, X.-D., Shang, J. & Gühne, O. Statistical methods for quantum state verification and fidelity estimation. Adv. Quantum Technol. 5, 2100126 (2022).
- Magesan, E. et al. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.* **109**, 080505 (2012).
- Proctor, T., Rudinger, K., Young, K., Sarovar, M. & Blume-Kohout, R. What randomized benchmarking actually measures. *Phys. Rev. Lett.* **119**, 130502 (2017).
- 19. Erhard, A. et al. Characterizing large-scale quantum computers via cycle benchmarking. *Nat. Commun.* **10**, 5347 (2019).

- Proctor, T., Rudinger, K., Young, K., Nielsen, E. & Blume-Kohout, R. Measuring the capabilities of quantum computers. *Nat. Phys.* 18, 75–79 (2022).
- Helsen, J., Roth, I., Onorati, E., Werner, A. & Eisert, J. General framework for randomized benchmarking. *PRX Quantum* 3, 020357 (2022).
- Carrasco, J., Elben, A., Kokail, C., Kraus, B. & Zoller, P. Theoretical and experimental perspectives of quantum verification. *PRX Quantum* 2, 010102 (2021).
- 23. Elben, A. et al. Cross-platform verification of intermediate scale quantum devices. *Phys. Rev. Lett.* **124**, 010504 (2020).
- 24. Zhu, D. et al. Cross-platform comparison of arbitrary quantum states. *Nat. Commun.* **13**, 6620 (2022).
- Knörzer, J., Malz, D. & Cirac, J. I. Cross-platform verification in quantum networks. Phys. Rev. A 107, 062424 (2023).
- Mohseni, M., Rezakhani, A. T. & Lidar, D. A. Quantum-process tomography: resource analysis of different strategies. *Phys. Rev. A* 77, 032322 (2008).
- Shabani, A. et al. Efficient measurement of quantum dynamics via compressive sensing. *Phys. Rev. Lett.* **106**, 100401 (2011).
- Flammia, S. T., Gross, D., Liu, Y.-K. & Eisert, J. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New J. Phys.* 14, 095022 (2012).
- Kim, Y. et al. Universal compressive characterization of quantum dynamics. *Phys. Rev. Lett.* **124**, 210401 (2020).
- Teo, Y. S. et al. Objective compressive quantum process tomography. *Phys. Rev. A* 101, 022334 (2020).
- 31. Elben, A. et al. The randomized measurement toolbox. *Nat. Rev. Phys.* 5, 9–24 (2023).
- Liang, Y.-C. et al. Quantum fidelity measures for mixed states. *Rep. Prog. Phys.* 82, 076001 (2019).
- Dankert, C., Cleve, R., Emerson, J. & Livine, E. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A* 80, 012304 (2009).
- Gross, D., Audenaert, K. & Eisert, J. Evenly distributed unitaries: on the structure of unitary designs. J. Math. Phys. 48, 052104 (2007).
- Brydges, T. et al. Probing rényi entanglement entropy via randomized measurements. Science 364, 260–263 (2019).
- Watrous, J. The Theory of Quantum Information (Cambridge University Press, Cambridge, 2018).
- Jamiołkowski, A. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.* 3, 275–278 (1972).
- Gilchrist, A., Langford, N. K. & Nielsen, M. A. Distance measures to compare real and ideal quantum processes. *Phys. Rev. A* 71, 062310 (2005).
- Nielsen, M. A. & Chuang, I. L. Quantum Computation and Quantum Information: 10th Anniversary Edition (Cambridge University Press, 2010).
- Korzekwa, K., Czachórski, S., Puchała, Z. & Życzkowski, K. Coherifying quantum channels. *New J. Phys.* 20, 043028 (2018).
- 41. Baidu Research. Quantum Error Processing Toolkit (QEP). https://quantumhub.baidu.com/qep/.
- Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* 16, 1050–1057 (2020).
- Vermersch, B., Elben, A., Dalmonte, M., Cirac, J. I. & Zoller, P. Unitary n-designs via random quenches in atomic hubbard and spin models: application to the measurement of rényi entropies. *Phys. Rev. A* 97, 023604 (2018).
- Anshu, A., Landau, Z. & Liu, Y. Distributed quantum inner product estimation. In Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2022, 44-51 (Association for Computing Machinery, New York, NY, USA, 2022).
- Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
- Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* 567, 491–495 (2019).
- Wang, K., Chen, Y.-A. & Wang, X. Mitigating quantum errors via truncated Neumann series. Sci. China Inf. Sci. 66, 180508 (2023).
- Tang, S., Zheng, C. & Wang, K. Detecting and eliminating quantum noise of quantum measurements. Preprint at https://arxiv.org/abs/2206.13743 (2023).
- Cai, Z. et al. Quantum error mitigation. Preprint at https://arxiv.org/abs/ 2210.00921 (2023).
- 50. Nielsen, E. et al. Gate set tomography. Quantum 5, 557 (2021).
- Elben, A., Vermersch, B., Roos, C. F. & Zoller, P. Statistical correlations between locally randomized measurements: a toolbox for probing entanglement in manybody quantum states. *Phys. Rev. A* 99, 052323 (2019).

ACKNOWLEDGEMENTS

Part of this work was done when C.Z. was a research intern at Baidu Research. We acknowledge the use of IBM Quantum services for this work. We expressed the views, and do not reflect the official policy or position of IBM or the IBM Quantum team. This

8

work was partially supported by the National Science Foundation of China (Nos. 61871111 and 61960206005).

AUTHOR CONTRIBUTIONS

K.W. formulated the idea. K. W. and C.Z. designed the protocol. C.Z. conducted the experiments and performed the analysis. All authors contributed to the preparation of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-023-00797-3.

Correspondence and requests for materials should be addressed to Xutao Yu or Kun Wang.

Reprints and permission information is available at http://www.nature.com/ reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons ۲ Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http:// creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024