ARTICLE OPEN

Check for updates

A quantum hamiltonian simulation benchmark

Yulong Dong[1,2], K. Birgitta Whaley[1,3,4] and Lin Lin[2,4,5]✉

Hamiltonian simulation is one of the most important problems in quantum computation, and quantum singular value transformation (QSVT) is an efficient way to simulate a general class of Hamiltonians. However, the QSVT circuit typically involves multiple ancilla qubits and multi-qubit control gates. In order to simulate a certain class of n-qubit random Hamiltonians, we propose a drastically simplified quantum circuit that we refer to as the minimal QSVT circuit, which uses only one ancilla qubit and no multi-qubit controlled gates. We formulate a simple metric called the quantum unitary evolution score (QUES), which is a scalable quantum benchmark and can be verified without any need for classical computation. Under the globally depolarized noise model, we demonstrate that QUES is directly related to the circuit fidelity, and the potential classical hardness of an associated quantum circuit sampling problem. Under the same assumption, theoretical analysis suggests there exists an 'optimal' simulation time $t^{\mathrm{opt}} \approx 4.81$, at which even a noisy quantum device may be sufficient to demonstrate the potential classical hardness.

npj Quantum Information (2022)8:131 ; https://doi.org/10.1038/s41534-022-00636-x

INTRODUCTION

Recent years have witnessed tremendous progress in quantum hardware and quantum algorithms. As near-term quantum devices become increasingly accessible, the need for holistic benchmarking of such devices is also rapidly growing. Indeed, while most of the frequently used quantum benchmarks, such as randomized benchmarking[1] and gateset tomography[2], still focus on the performance of one or a few qubits, over the past three years a number of 'whole machine' benchmarks have been proposed that aim at assessing the performance of quantum devices beyond a small number of qubits[3–10].

While results from such generic benchmarks certainly provide important characteristics of the quantum devices themselves, we are ultimately interested in applying the devices to carry out specific computational tasks. However, the circuit structure of quantum algorithms can be vastly different for different algorithms. Generic quantum benchmarks can miss structural information that is specific to a particular algorithm and which may amplify either quantum errors of certain types or errors amongst a certain group of qubits, and/or reduce errors elsewhere. In this work we address the benchmarking of quantum simulations for time-independent Hamiltonians. Such a simulation can be stated as follows: given an initial state $|\psi_0\rangle$ and a Hamiltonian $\mathfrak{H}$, evaluate the quantum state at time t according to $|\psi(t)\rangle = \exp(-it\mathfrak{H})|\psi_0\rangle$. Hamiltonian simulation is of immense importance in characterizing quantum dynamics for a diverse range of systems and situations in quantum physics, chemistry and materials science. Simulation of one quantum Hamiltonian by another quantum system was also one of the motivations of Feynman's 1982 proposal for design of quantum computers[11]. Hamiltonian simulation is also used as a subroutine in numerous other quantum algorithms, such as quantum phase estimation[12] and solving linear systems of equations[13].

Following the conceptualization of a universal quantum simulator using a Trotter decomposition of the time evolution operator $e^{-it\mathfrak{H}}$[14], a number of quantum algorithms for Hamiltonian simulation have been proposed[15–19]. Detailed assessment of these algorithms, with continued improvement of theoretical error bounds, has since emerged as a very active area of research[20–31]. In this context, one of the most significant developments in recent years is the quantum signal processing (QSP) method[17], and its generalization, the quantum singular value transformation (QSVT) method[32]. For sparse Hamiltonian simulation, the query complexity of QSVT matches the complexity lower bound with respect to all parameters[17,32]. The QSVT method also enjoys another advantage, namely that the quantum circuit is relatively simple, and requires very few ancilla qubits. QSVT allows one to use essentially the same parameterized quantum circuit to perform a wide range of useful computational tasks, including Hamiltonian simulation[33], solution of linear systems[32,34,35], and finding eigenstates of quantum Hamiltonians[36]. In this sense, it provides a 'grand unification' of a large class of known quantum algorithms[37].

Despite these advantages, QSVT is generally not viewed as a suitable technique for near-term quantum devices today. This is largely because these techniques rely on an input model called 'block encoding' which views the Hamiltonian $\mathfrak{H}$ as a submatrix of an enlarged unitary matrix $U_{\mathfrak{H}}$. For Hamiltonians arising from realistic applications (e.g., linear combination of products of Pauli or fermionic operators, and sparse matrices in general), the construction of $U_{\mathfrak{H}}$ often involves multiple ancilla qubits and multi-qubit control gates. Taken together, these requirements can make QSVT very difficult to implement with high fidelity and to date there has been no QSVT based Hamiltonian simulation on realistic devices.

RESULTS

Overview

In this work we remedy this situation by identifying and demonstrating an application for QSVT on near term quantum devices that allows benchmarking of Hamiltonian simulation for a

[1]Berkeley Center for Quantum Information and Computation, Berkeley, CA 94720, USA. [2]Department of Mathematics, University of California, Berkeley, CA 94720, USA. [3]Department of Chemistry, University of California, Berkeley, CA 94720, USA. [4]Challenge Institute of Quantum Computation, University of California, Berkeley, CA 94720, USA. [5]Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. ✉email: linlin@math.berkeley.edu

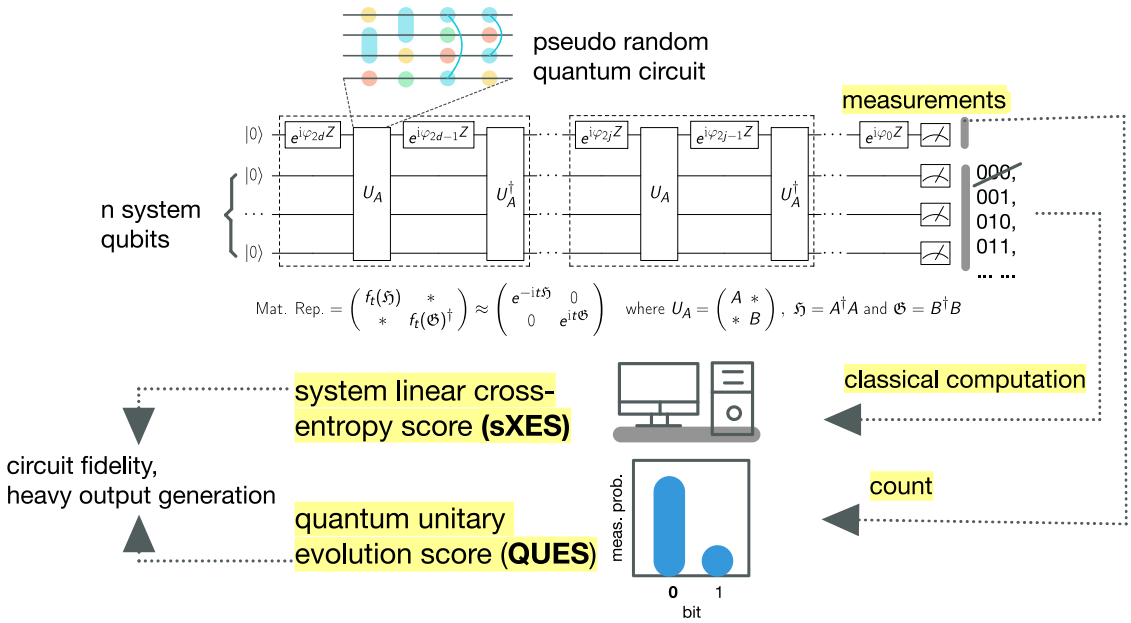

Fig. 1 Illustration of the minimal quantum singular value transformation (mQSVT) circuit for the Hamiltonian simulation benchmark. The overall circuit implements a complex matrix polynomial $f_t(\mathfrak{H})$ of degree d on the Hamiltonian $\mathfrak{H}$ that is defined in terms of a pseudo random quantum circuit U_A. The circuit acts on $n+1$ qubits, consisting of n system qubits and 1 ancilla qubit. After measuring the top ancilla qubit and post-selecting on the 0 outcome of this, the action on the bottom n system qubits accurately approximates $\exp(-it\mathfrak{H})|0^n\rangle$.

class of Hamiltonians that are relevant to recent efforts to demonstrate supremacy of quantum computation over classical computation[6]. This is the class of random Hamiltonians generated from block encoding of random unitary operators that correspond to random unitary circuits. We show that for this class of Hamiltonians it is possible to formulate a simple metric, called the quantum unitary evolution score (QUES), for the success of quantum unitary evolution. This metric is the primary output from the Hamiltonian simulation benchmark, and is directly related to the circuit fidelity. This allows verification of Hamiltonian simulation on near-term quantum devices without any need for classical computation, and the approach can be scaled to a large number of qubits.

The main result of this paper is a *very simple* quantum circuit (Fig. 1), called the minimal QSVT (mQSVT) circuit. With proper parameterization, the mQSVT circuit is able to propagate a certain class of random Hamiltonians $\mathfrak{H}$ to any given target accuracy. In fact, we argue that the mQSVT circuit is not only *the simplest* quantum circuit for carrying out a QSVT based Hamiltonian simulation, but that it is actually the simplest possible circuit for all tasks based on QSVT. Here $\mathfrak{H}$ is not a Hamiltonian corresponding to a given physical system, but a random Hamiltonian generated using a simple random unitary circuit, called a Hermitian random circuit block encoded matrix (H-RACBEM)[9]. However, for the purpose of benchmarking the capability of a quantum device to perform arbitrary Hamiltonian simulations, averaging over a distribution of the underlying arbitrary Hamiltonians is precisely what is required to generate a holistic benchmark protocol that samples from all possible instantiations.

The quantum circuit in Fig. 1 consists of two components: an arbitrary random unitary matrix U_A that implicitly defines the Hamiltonian $\mathfrak{H}$, together with its Hermitian conjugate $U_A^\dagger$ and a series of R_z gates with carefully chosen phase factors $\{\varphi_i\}_{i=0}^{2d}$ (see Supplementary Note 3). The mQSVT circuit makes d queries to U_A and $U_A^\dagger$, two of which are shown explicitly in Fig. 1. For an n-qubit matrix $\mathfrak{H}$, the total number of qubits needed is always $n+1$, i.e., only 1 ancilla qubit, hereafter referred to as the signal qubit, is required. This is even smaller than the simplest QSVT circuit[17], which requires at least 2 ancilla qubits. However, more important than the reduction of the number of qubits is the fact that Fig. 1 removes all two-qubit and multi-qubit gates outside of the unitary U_A. This means that one can choose any convenient entangling two-qubit gate (e.g. CZ, CNOT, $\sqrt{i}$SWAP) and any coupling map that is native to a quantum device to construct the random U_A. Combining this with the sequence of single qubit R_z gates then makes the resulting benchmarking quantum circuit of Fig. 1 readily executable.

Quantum unitary evolution score (QUES)

Figure 1 implements $f_t(\mathfrak{H})|0^n\rangle$ on the system qubits, where $f_t(\mathfrak{H})$ is a matrix polynomial (see Supplementary Note for details), with approximation error in the operator norm upper bounded by $\|f_t(\mathfrak{H}) - e^{-it\mathfrak{H}}\|_2 \le \epsilon$. Therefore in the absence of quantum errors, after applying the circuit to the input state $|0^{n+1}\rangle$, the probability $P_t(U_A) := \|f_t(\mathfrak{H})|0^n\rangle\|^2$ of measuring the top ancilla qubit with outcome 0 will be close to 1, indicating that the underlying Hamiltonian evolution is unitary.

From now on, we will primarily consider mQSVT circuits with a fixed set of phase factors $\{\varphi_i\}$ and hence fixed simulation time t. For notational simplicity, we will drop the t-dependence in quantities such as $P_t(U_A)$, unless specified otherwise.

On a real quantum device, the probability $P(U_A)$ should be replaced by $P_{\exp}(U_A)$, which is the experimentally measured probability. We define the quantum unitary evolution score (QUES) by

$$\mathrm{QUES}(n,d) := \mathbb{E}\big(P_{\exp}(U_A)\big), \tag{1}$$

where the expectation is taken over the ensemble of random quantum circuit instances U_A. The deviation of QUES from 1 then measures the average performance of the quantum computer under a Hamiltonian simulation task.

There is no unique prescription for constructing random quantum circuits. To fix the choice of U_A, we employ here the model random quantum circuit construction used to analyze the concept of quantum volume in[4]. Here, given a number of qubits n, U_A is constructed to contain n layers, each consisting of a random permutation of the qubit labels followed by random two-qubit gates between the n qubits. Given this construction, the QUES in

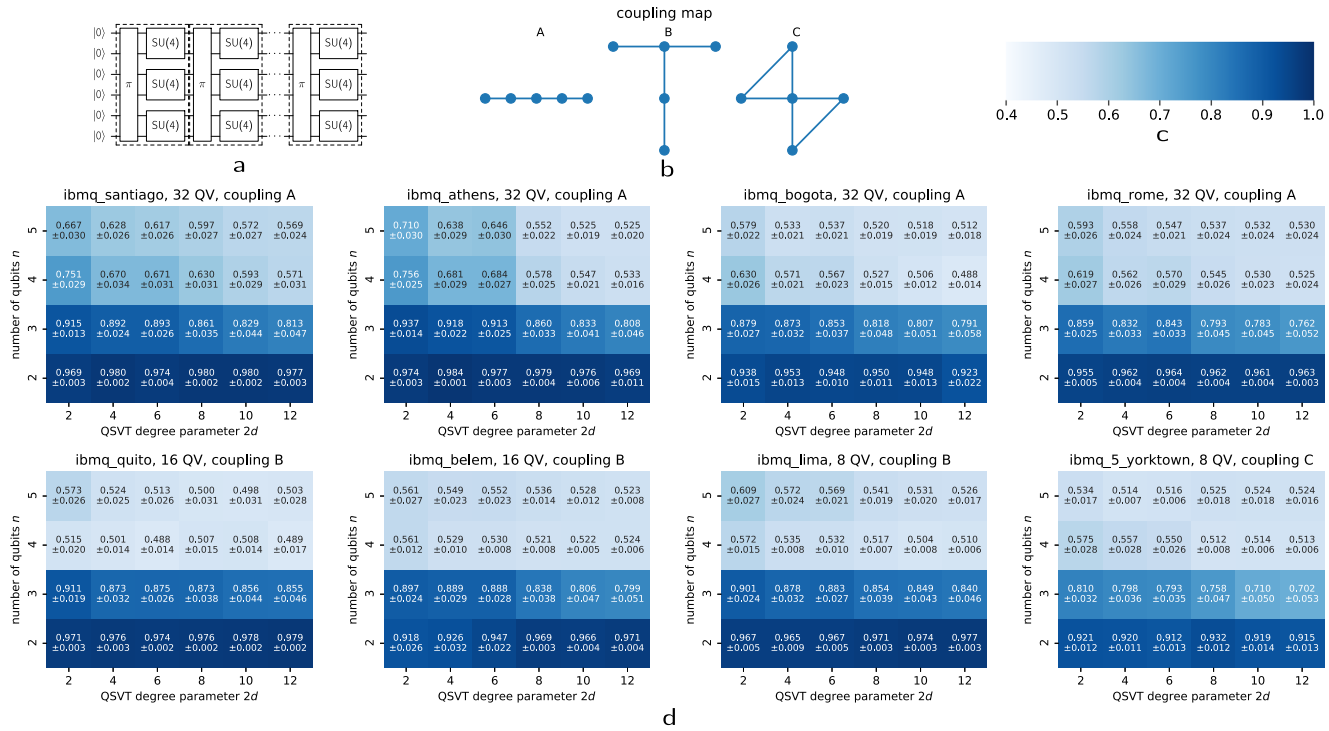

Fig. 2 Quantum unitary evolution score (QUES) of the 5-qubit quantum devices provided by the IBM Q platform. (**a**) Visualization of the quantum circuit U_A used in computing QUES. When the number of qubits is n, there are n layers of the dashed boxes consists of the random permutation of the qubits labels followed by random two-qubit gates. After calling the transpiler, the circuit U_A is decomposed with respect to the basic gate set $\Gamma = \{\mathrm{Rz}, \sqrt{\mathrm{X}}, \mathrm{X}, \mathrm{CNOT}\}$ and the coupling map which indicates the available qubit pairs on which CNOT can act. **b** Layouts of coupling maps. **c** Color bar of the heatmap. **d** Each heatmap displays the benchmarking result of a specific quantum device, with the title showing the name of the device, its quantum volume, and its coupling map. Each QUES is estimated from 50 circuit instances. Each circuit instance is measured with 1000 measurement shots. The number displayed in each heatmap is the QUES value and its 95% confidence interval.

Eq. (1) will then depend only on n and d, and the overall depth of the circuit is approximately $2d$ times the circuit depth of U_A. Note that given the basic quantum gate set of a particular quantum device, alternative constructions of U_A using random choices of specific one- and two-qubit gates are possible.

Figure 2 shows the results of computing the QUES across 8 different IBM Q quantum devices (https://quantum-computing.ibm.com), each having 5 qubits and one of three distinct coupling maps (panel b). When the number of qubits $n \le 3$, the QUES on all devices is relatively high ($\gtrsim 0.7$) but it decreases sharply for $n \ge 4$. In contrast, the QUES decreases only relatively mildly as d increases. This is particularly noticeable for $n = 2$, which may indicate that the quantum circuit transpiler provided by the IBM Q may be particularly effective for this device with very small qubit number. We emphasize that compared to generic benchmark measures such as the quantum volume, the QUES is specific to the computational task of the Hamiltonian simulation, and any information specific to this is not diluted by additional averaging over output distributions from other computational tasks. In particular, we find that even for quantum devices with relatively small quantum volume (8-QV), the performance in terms of QUES is only mildly worse than for those with a larger quantum volume (32-QV).

Circuit fidelity and system linear cross-entropy score (sXES)

The quality of a noisy implementation of a quantum circuit is often characterized by the circuit fidelity. Loosely speaking, the output quantum state of a noisy circuit can be characterized as a convex combination of the correct result and the result obtained under noise, i.e., 'output' $= \alpha \times$ 'correct result' $+ (1-\alpha) \times$ 'noise', where $0 \le \alpha \le 1$ is the circuit fidelity. Let $p(U_A, x)$ be the noiseless bitstring probability of measuring the mQSVT circuit with outcome 0 in the ancilla qubit and an n-bit binary string x in the n system qubits. Let $p_{\mathrm{exp}}(U_A, x)$ be the corresponding experimental bitstring probability, which can be estimated from the frequency of occurrence of the bitstring $0x$ in the measurement outcomes. Since the random circuit U_A is approximately drawn from the Haar measure, we make analogous assumptions to those in[3,6], and assume the following global depolarized error model:

$$p_{\mathrm{exp}}(U_A, x) = \alpha p(U_A, x) + \frac{1-\alpha}{2^{n+1}}. \tag{2}$$

We discuss the justification and potential generalization of such an error model in Supplementary Note 4.

Under the global depolarized error model, we now analyze the effect of noise on the circuit and show how measuring the QUES allows the circuit fidelity to be extracted. Prior work has made use of a combination of quantum and classical computation to obtain the circuit fidelity α. Such analysis relies on the possibility of evaluating the noiseless bitstring probability $p(U_A, x)$ classically, given U_A and x, e.g., via tensor network contraction[38]. This enabled the estimation of α from measurements of cross-entropy, referred to as XEB in this setting[3,6]. We adapt this approach to the Hamiltonian simulation problem by defining a system linear cross-entropy score (sXES):

$$\mathrm{sXES}\,(U_A) := \sum_{x \neq 0^n} p(U_A, x) p_{\mathrm{exp}}(U_A, x). \tag{3}$$

The prefix 'system' is added because the ancilla qubit is fixed to be the $|0\rangle$ state in the definition of $p(U_A, x), p_{\mathrm{exp}}(U_A, x)$, and the $|x\rangle$ state belongs to the system register. In order to connect to the problem of generating heavy weight samples later, our definition of sXES excludes the bitstring 0^n. This is necessary also since the

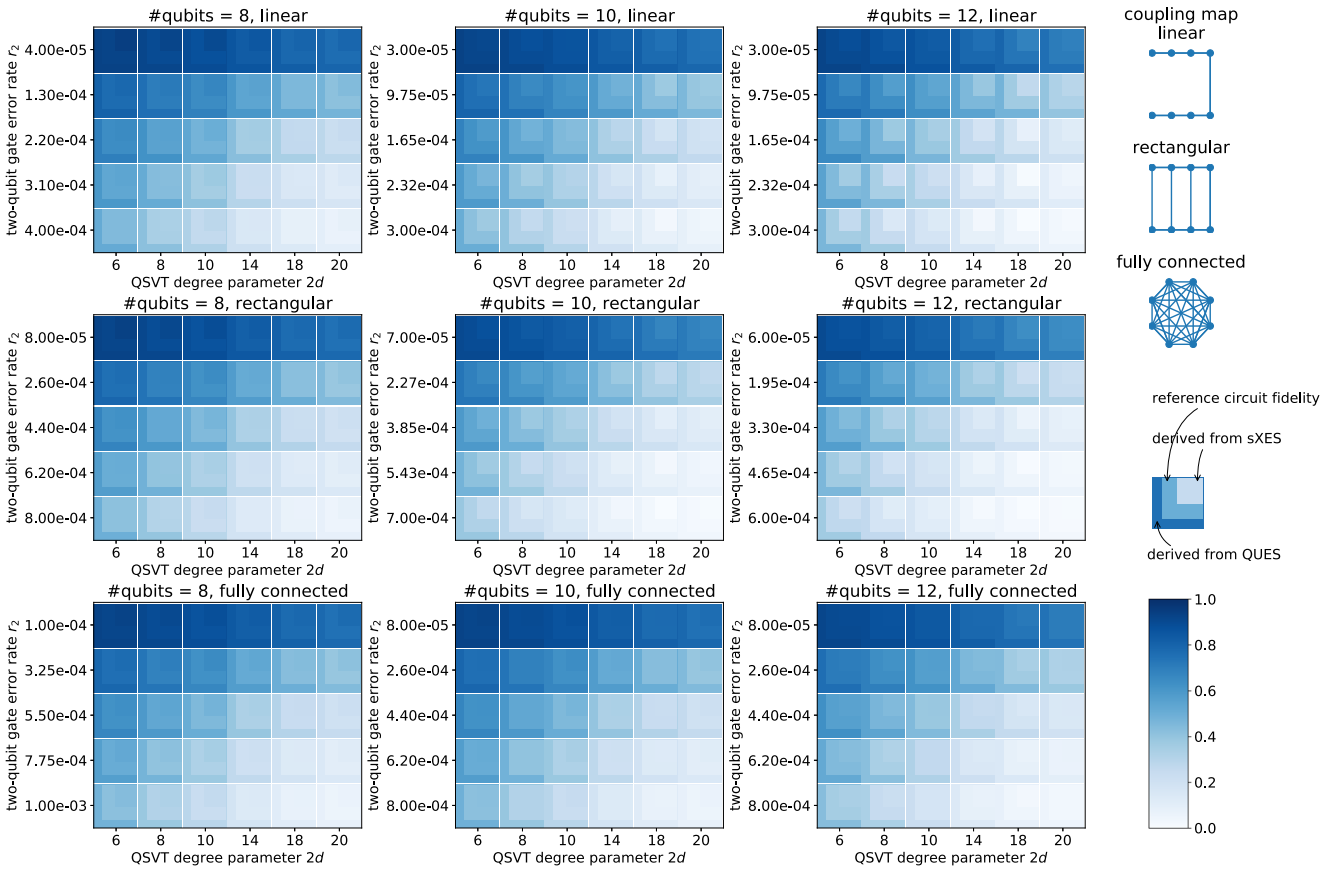

Fig. 3 Circuit fidelity estimated from the quantum Hamiltonian simulation benchmark. Colored grids represent the circuit fidelity estimated from ~100 circuit repetitions. The benchmarking is performed for circuits with a range of number of system qubits, having also variable types of couplings and a range of error parameters. The depth of the random circuit instances is set to the convergent depth deduced from the convergence to Haar measure (see Supplementary Note 7). The right column contains graphical depictions of the coupling maps, the layout of each grid, and the color bar.

statistical properties of the bitstring 0^n are different from those of the bitstrings in the system register. Taking the expectation with respect to the distribution of U_A, and rearranging Eq. (2) then gives an expression for the circuit fidelity:

$$\alpha = \frac{\mathbb{E}(\,\mathrm{sXES}\,(U_A)) - \frac{1}{2^{n+1}}\mathbb{E}\left(\sum_{x\neq 0^n} p(U_A,x)\right)}{\mathbb{E}\left(\sum_{x\neq 0^n} p(U_A,x)^2\right) - \frac{1}{2^{n+1}}\mathbb{E}\left(\sum_{x\neq 0^n} p(U_A,x)\right)}. \tag{4}$$

This expression holds for any ensemble of random matrices, and relies only on the assumption that the noise model is depolarizing.

Once the probability distribution of U_A is specified (e.g., the Haar measure[39]), the only term in α that requires a quantum computation is sXES(U_A), and all other terms in Eq. (4) can be evaluated classically. However, evaluation of the right-hand side of Eq. (4) often requires a significant amount of classical computation when n becomes large[3].

Inferring circuit fidelity from QUES

Based on the discussion so far, it might seem surprising that an alternative, very good approximation to the circuit fidelity can readily be obtained from the QUES metric in Eq. (1). This is arrived at by first defining $P_{\mathrm{exp}}(U_A) = \sum_x p_{\mathrm{exp}}(U_A, x)$, i.e., the average over all possible output bit strings x of the probability of measuring a given bit string as outcome of the action of U_A on the input state $|0^{n+1}\rangle$. Then summing both sides of Eq. (2) with respect to all bit strings x, further taking the expectation value of both sides over all possible U_A yields a fidelity estimate α_{QUES} that can be obtained directly from the measured QUES value, namely

$$\alpha_{\mathrm{QUES}} = 2\times \mathrm{QUES} - 1. \tag{5}$$

The approximation error ϵ is defined as the maximal error for simulating a bounded Hamiltonian using the mQSVT circuit, namely $\epsilon := \max_{\|\mathfrak{H}\|_2\leq 1} \|f_t(\mathfrak{H}) - e^{-it\mathfrak{H}}\|_2$. It determines the extent of deviation of α_{QUES} from α. Specifically, under the globally depolarized noise model, we have the following bound (Supplementary Note 6)

$$|\alpha_{\mathrm{QUES}} - \alpha| \leq 16\epsilon + \mathcal{O}(\epsilon^2). \tag{6}$$

Here the error bound is derived without including the Monte Carlo measurement error due to the finite number of measurement shots. The analysis of the resulting statistical error is given in Supplementary Note 12 C.

It is evident that, unlike Eq. (4), there is no classical overhead for evaluating α_{QUES} for any n. Since the circuit fidelity α should be non-negative, combining Eq. (5) and Eq. (6) also indicates that under the assumption of the depolarizing noise model, we have $\mathrm{QUES} \geq 0.5 - 8\epsilon + \mathcal{O}(\epsilon^2)$.

To numerically verify the relation between QUES and circuit fidelity, we make use of the digital error model of[3] in which each quantum gate in the circuit is subject to a depolarizing error channel with a certain error rate. We test the resulting noisy quantum circuit with different two-qubit gate error rates r_2 and set the one-qubit gate error rate to $r_1 = r_2/10$. We also discard the rotation gate with phase factor φ_{2d}, since this just adds a global phase to the exact Hamiltonian simulation. Then, given U_A with a

total of g_1 one-qubit gates and g_2 two-qubit gates, the reference value of the circuit fidelity can be set to $a_{\text{ref}} := (1-r_1)^{2d(g_1+1)}(1-r_2)^{2dg_2}$[3,6]. We assume U_A is Haar-distributed (numerically verified in Supplementary Note 7) to simplify the computation of classical expectations.

Figure 3 summarizes the estimated circuit fidelity for random quantum circuits with different depth parameter d, variable coupling maps, and a range of error parameters. In all cases, we find that the derived circuit fidelity from QUES (a_{QUES}), the circuit fidelity a obtained from sXES, and the reference value a_{ref} are generally consistent with each other. Numerical results also show that a_{QUES} exhibits a trend that slightly overestimates the value of the fidelity a (see Supplementary Table 1 for numerical values of the fidelities). We also see that for a given set of error rates r_1, r_2, the circuits with highest connectivity show the best performance. This is because random circuits on these architectures converge faster to the Haar measure, which reduces the circuit depth (see Supplementary Note 7).

In the next two subsections we show how to assess and evaluate whether the Hamiltonian simulation with the mQSVT circuit can be a classically hard task. We first define the analog of XHOG for Hamiltonian simulation, which we refer to as sXHOG, and give conditions for the hardness of this. We then show that potential classical hardness can be inferred directly from the value of the circuit fidelity obtained from the QUES, i.e. from a_{QUES}.

Classical hardness and system linear cross-entropy heavy output generation (sXHOG)

The complexity-theoretic foundation of the Google claim of 'quantum supremacy' in[6] is based on a computational task called linear cross-entropy heavy output generation (XHOG) with Haar-distributed unitaries[3,6,40,41]. Specifically, given a number $b>1$ and a random n-qubit unitary U, the task is to generate k nonzero bitstrings $x_1, x_2, \cdots, x_k \in \{0,1\}^n$ such that $\frac{1}{k}\sum_{j=1}^{k} q(U,x_j) \geq b \times 2^{-n}$, where $q(U,x) = |\langle x|U|0^n\rangle|^2$. Here we use U without the subscript to distinguish the XHOG problem and the sXHOG problem which will be defined later. For k randomly generated bitstrings, we expect that $\frac{1}{k}\sum_{j=1}^{k} q(U,x_j) \approx 2^{-n}$. Therefore any value $b>1$ will correspond to a 'heavy weight' output. When k is large enough, successful solution of the XHOG problem is considered to be classically hard for every value $b>1$[40,41]. This holds for every circuit fidelity estimate $a>0$ obtained from the XEB metric, leading to the claim of supremacy in[6] based on extraction of a value $a \approx 0.002$ from the experiments.

For the Hamiltonian simulation benchmark, we can define an analogous linear cross-entropy heavy output generation problem for the n system qubits. Note that the heavy weight samples are now defined only for the system qubits. We shall refer to this heavy output generation problem for Hamiltonian simulation as the sXHOG problem, to emphasize this important feature and the difference from the standard XHOG problem. Specifically, given a number $b>1$, a Hamiltonian simulation benchmark circuit with sufficiently small approximation error ϵ, and a random $(n+1)$-qubit unitary U_A defining a random Hamiltonian on the n qubits, the task is to generate k nonzero bitstrings $x_1, x_2, \cdots, x_k \in \{0,1\}^n \backslash \{0^n\}$ such that $\frac{1}{k}\sum_{j=1}^{k} p(U_A,x_j) \geq b \times 2^{-n}$. Now for the case of Hamiltonian simulation, $p(U_A,x) = \mathcal{O}(2^{-n})$ for any $x \neq 0^n$ at all t, but $p(U_A,0^n)$ can be much larger (for more details see Supplementary Figure 5(b) in Supplementary Note 11). The state 0^n is then by definition 'heavy' and we must therefore exclude this from the measure in order to avoid a trivial outcome. This is what distinguishes the sXHOG problem from the original XHOG problem.

The potential classical hardness of the XHOG problem is justified by a reduction to a complexity-theoretic conjecture, called linear cross-entropy quantum threshold assumption (XQUATH)[40]. For completeness, we give a similar variant of the reduction of sHOG problem to a conjecture named system linear cross-entropy quantum threshold assumption (sXQUATH) in Theorem 6 of Supplementary Note 9. The concept of sXQUATH directly parallelizes that of XQUATH, with a similar restriction as above to exclude the output bit string 0^n (for more details see Supplementary Note 9). Similar to the construction in ref. [40], the classically efficient solution to sXHOG problem yields a violation to sXQUATH, which assumes that $p(U_A,x)$ for $x \neq 0^n$ cannot be efficiently estimated on classical computers to sufficient precision.

Inferring classical hardness from QUES

In order to decide whether a noisy implementation of the Hamiltonian simulation benchmark is potentially in the classically hard regime, we need to establish whether or not the sXHOG problem can be solved for $b>1$.

Under the assumption that U_A is drawn from the Haar measure, and that the approximation error ϵ of the mQSVT circuit is sufficiently small, we derive the following relation between b and the circuit fidelity a:

$$b = 1 + \frac{\gamma(a-a^*)}{a+1}. \tag{7}$$

Here a^* is a fidelity threshold (not the complex conjugation of a) and γ a constant. Explicit expressions for the threshold value a^* and the constant γ are given in Supplementary Note 10. Both quantities are independent of the circuit fidelity a and depend only on the number of system qubits n and the simulation time t. Eq. (7) thus shows that when $\gamma>0$ and $a>a^*$, we will have $b>1$ so that the sXHOG problem solved by the mQSVT circuit might be classically hard. This is qualitatively different from the situation for XEB experiments, for which every $a>0$ implies $b>1$[6].

Using the relation between QUES and a in Eqs. (5) and (6), and assuming that ϵ is sufficiently small, we immediately arrive at the conclusion that when

$$\text{QUES} \geq (1+a^*)/2, \quad \gamma>0, \tag{8}$$

the corresponding sXHOG problem might be classically hard for a sufficiently large value of n. This is a surprising result, since as noted above, the estimation of QUES does not require intensive classical computation. In fact it is not even necessary to actually generate any heavy weight samples - instead we just need to measure the value of QUES, Eq. (1), which is readily done by repeatedly running the circuit in Fig. 1 with random circuit parameters as described above. Of course, should one wish to actually solve the sXHOG problem itself, the heavy weight samples would need to be generated using a quantum computer and intensive classical computation for computation of $\frac{1}{k}\sum_{j=1}^{k} p(U_A,x_j)$ would then also be required. But in order to demonstrate the potential regime of classical hardness for Hamiltonian simulation, i.e., the minimal values of n and d to reach this regime, this is not required.

To further investigate the implications of Eq. (7), we now explicitly indicate the time dependence of all quantities (i.e., we employ the notation $\gamma \to \gamma_t, a^* \to a_t^*$). In Fig. 4 we plot the values of γ_t, a_t^* according to the expressions given in Supplementary Note 10 as a function of the simulation time t, for $n=4,8,12$ qubits. Figure 4 shows that $\gamma_t>0$ for all t, so then we only need to determine whether it is possible to have fidelity $a \geq a_t^*$. It is evident from Fig. 4 that both a_t^* (panel (a)) and γ_t (panel (b)) show oscillatory behavior. We now analyze this behavior to identify an optimal time at which the potential classical hardness of Hamiltonian simulation in this random Hamiltonian setting can be demonstrated for a sufficiently large number of qubits n.

For very short times, i.e., when $t \approx 0$, we have $a_t^* > 1$. This means that we cannot have $b>1$ for any value of the circuit fidelity $0 \leq a \leq 1$. To see why this is the case, consider the limit $t=0$. Here $p_t(U_A,0^n)=1$, and $p_t(U_A,x)=0$ for any $x \neq 0^n$. By continuity, when t

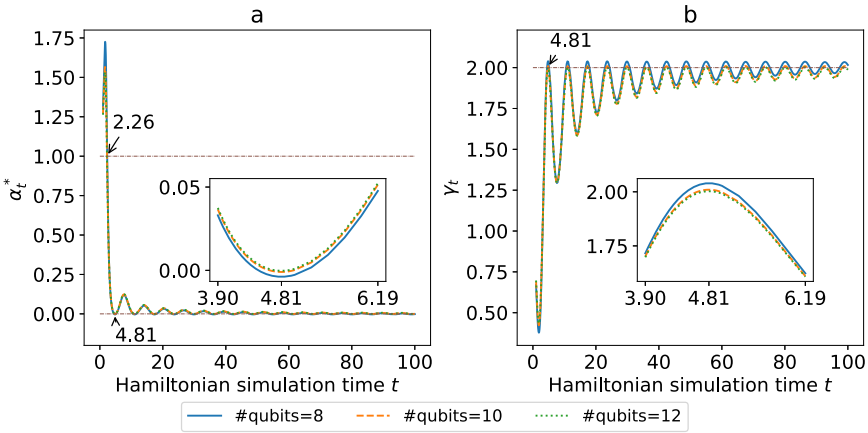

Fig. 4 Quantities relevant to the system linear cross-entropy heavy output generation (sXHOG) problem, evaluated using the explicit expressions given in Supplementary Note 10. a The threshold fidelity α_t^* as a function of Hamiltonian simulation time t. The upper value noted on the plot indicates the time value $t^{\text{thr}} \approx 2.26$ where $\alpha^*(t^{\text{thr}}) = 1$. The lower value noted on the plot indicates the regime at finite time $t^{\text{opt}} \approx 4.81$ with the first minimal value of threshold fidelity. **b** The parameter γ_t as a function of Hamiltonian simulation time t. The value noted on the plot indicates the value $\gamma_t \approx 2$ at the optimal time $t^{\text{opt}} \approx 4.81$. The averages in (**a**) and (**b**) are estimated numerically from ~100 instances of the mQSVT circuit encoding random Hamiltonians drawn from the Haar measure. Insets in each panel show the behavior of α_t^* and γ_t near the optimal time $t^{\text{opt}} \approx 4.81$.

is very small, the magnitude of $p_t(U_A, x)$ for most bitstrings $x \neq 0^n$ is still very small and cannot reach the heavy output regime. Figure 4(a) also shows that there is a critical simulation time $t^{\text{thr}} \approx 2.26$, for which $\alpha_t^* < 1$ for any $t > t^{\text{thr}}$.

When $t > t^{\text{thr}}$, ideally we would like to have $\alpha_t^* \approx 0$, so that a very low experimental circuit fidelity α is sufficient to reach the heavy output regime. To this end we investigate what happens at the vanishing circuit fidelity, i.e., $\alpha = 0$. Detailed analysis shows that in the large n limit, we have $\gamma_t \alpha_t^* = \mathbb{E}(p_t(U_A, 0^n))$, and Eq. (7) can be simplified as (see Supplementary Note 10)

$$b|_{\alpha=0} = 1 - \mathbb{E}(p_t(U_A, 0^n)), \tag{9}$$

where the expectation value is taken with respect to the random unitaries U_A as before. Thus when the expectation value is positive, i.e., $\mathbb{E}(p_t(U_A, 0^n)) > 0$, in the large n limit we have $b|_{\alpha=0} < 1$ and the task should not be classically hard. Moreover, since b is a continuous function of α, even if we now have finite circuit fidelity α, when this is small enough we can still find $b < 1$. This provides an alternative explanation of Eq. (7), namely, that the circuit fidelity α needs to be larger than the finite positive threshold value $\alpha_t^* > 0$ for *most* values of $t > t^{\text{thr}}$.

As a result of these considerations, when n is large enough, it is important to focus on the regimes where the expectation value $\mathbb{E}(p_t(U_A, 0^n)) \approx 0$, which from Eq. (7) implies that the threshold fidelity $\alpha_t^* \approx 0$. The numerical results shown in Fig. 4 indicate that this can happen in two different scenarios. The first is when the simulation time $t \to \infty$ (see the analytic justification of this statement in Supplementary Note 12 D). Of course this requires a very large circuit depth and is a physically 'trivial' limit that is impractical on near-term quantum devices. The second scenario, which is much more relevant in practice, is when α_t^* reaches its first minimum, which defines an optimal time $t = t^{\text{opt}}$. In the large n limit, the value of t^{opt} can be rationalized as the first node of the Bessel function $J_0(t/2)$ (see Supplementary Note 11). Figure 4(a) shows that for $t^{\text{opt}} \approx 4.81$, we already have $\mathbb{E}(p_t(U_A, 0^n)) \approx 0$ and $\alpha_t^* \approx 0$. Therefore simulating to the time $t = t^{\text{opt}}$ is highly desirable, since this is a relatively short time at which the Hamiltonian simulation benchmark is nevertheless now guaranteed to solve the sXHOG problem even for a very small circuit fidelity. Our numerical results indicate that the values of t^* and t^{opt} depend only weakly on n, and their values are nearly converged for n as small as 12. Therefore this value of t^{opt} can be used in a future quantum simulation in the heavy output regime.

DISCUSSION

We have presented a quantum benchmark for Hamiltonian simulation on quantum computers. The Hamiltonian simulation problem is solved using a minimal quantum singular value transformation (mQSVT) circuit. The primary output of the Hamiltonian simulation benchmark is a single number called QUES, which can be verified without any classical computation, even in the regime that is potentially hard for classical computation. Therefore the Hamiltonian simulation benchmark provides a scalable benchmark of the circuit fidelity under the global depolarized error model, and can be executed and verified on future quantum devices with a large number of qubits.

As the current quantum computing technologies advance, the possibility of implementing some error correction is improving[42]. Here the highly structured mQSVT circuit provides useful indications of where best to implement error correction under limited resources for this. Recall that the mQSVT circuit consists of a series of repetitions of a random circuit U_A and its conjugate $U_A^\dagger$, interleaved with single-qubit Z rotation operators characterized by carefully selected phase factors. Thus given a specific random Hamiltonian block encoded in U_A, the time dependent evolution operator for this Hamiltonian is defined entirely in terms of the phase angles for the single-qubit Z rotation operators. Since these phases should moreover be precisely determined, this suggests that on near-term quantum devices that may allow for some error correction but have overall limited resources, quantum error correction for these single-qubit rotations should be prioritized.

It is also useful to consider here the applicability of this Hamiltonian simulation approach to general Hamiltonians, i.e., not restricted to random Hamiltonians, on near-term quantum computers. Unfortunately it appears that for current quantum technologies there is potentially a large gap between the feasible simulation of a H-RACBEM given in this work and that of a general Hamiltonian relevant to e.g., molecular or solid-state physics. The main reason is that the block encoding of most Hamiltonians of practical interest will involve significant numbers of ancilla qubits, as well as multi-qubit control gates, all of which are extremely expensive on near-term quantum devices. In contrast to this general situation, the construction of H-RACBEM uses only whatever one-qubit and two-qubit gates are available for a given quantum device and is thus considerably easier. Nevertheless, it is possible that undertaking Hamiltonian simulation with H-RACBEM may also yield interesting physical applications to the various

settings in which quantum chaotic dynamics are relevant. One immediate possibility in this direction is to use H-RACBEM to simulate the dynamics of quantum scrambling or quantum chaos in strongly interacting quantum systems. Scrambling dynamics can be studied by simulating out-of-time-order correlators (OTOCs) for effective Hamiltonians that can be defined implicitly in terms of a random circuit for time t (see e.g., ref. [43]). We note that one can easily perform a Hamiltonian simulation backward in time, merely by reversing the sign of t, so the mQSVT circuit for an OTOC at any time t of a random Hamiltonian encoded in H-RACBEM can be readily constructed by adding local operators between forward and backward implementations of the mQSVT. Evaluation of the circuit at different times t can be implemented either by reevaluating the phase factors (which may required building a longer circuit depending on the accuracy required). The circuits can also be adapted to Hamiltonian simulation at finite temperatures and hence also to scrambling at finite temperatures. From a theoretical perspective it would also be useful to explore to what extent the structure of the H-RACBEM influences the speed of scrambling[44].

Our theoretical analysis of the sensitivity of the Hamiltonian benchmarking scheme in this work was based on a fully depolarized noise model, which is often assumed to be a good model for superconducting qubits[6]. In general the Pauli stochastic noise model on which is based may be biased or non-uniform across qubits. In addition, thermal noise and coherent errors are important for some qubit architectures. It will be useful to extend the current analysis to more general noise models, and some of these aspects are discussed in Supplementary Note 4.

Finally, we note that while this Hamiltonian simulation benchmark is restricted to the specific class of random Hamiltonians, it might also provide information relevant to more general Hamiltonian simulations. Efforts to analyze the complexity of analog Hamiltonian simulations have often focused on the relation of such simulations to classical sampling tasks[45–47], and are closely related to the cross-entropy analysis for sampling of random quantum circuits of[3,6]. As noted recently[47], the classical hardness can be shown for certain classes of analog quantum Hamiltonian simulation[48,49]. Note that the potential classical hardness of the original XHOG problem corresponding to Google's supremacy experiment is justified by a reduction to a complexity-theoretic conjecture called XQUATH[40]. However, a recent paper[50] that appeared after submission of the current work has provided evidence that can refute XQUATH, at least for some classes of quantum circuits. Therefore it is possible that our sXQUATH assumption can be refuted on the same basis. It could be useful to explore generalizations of other classical sampling tasks to the QSVT setting, as was done here for the cross-entropy heavy output generation, to help guide the search for Hamiltonians whose simulation by QSVT can exhibit quantum advantages. Finally, the current approach of analysis of alternative fidelity measures under Hamiltonian simulation using mQSVT may provide useful for analysis of recent fidelity based experimental studies of analog Hamiltonian simulations that followed the emergent random nature of a projected ensemble of states[51].

METHODS

Details of numerical simulations

All numerical tests are implemented in `python3.7` and `Qiskit`[52]. Quantum circuits in Fig. 2 are optimized by the `transpiler` provided by `Qiskit` before being executed on a real quantum device. The number of measurements (shots) is fixed to be 1000 for the experiments on real quantum devices in Fig. 2, and it is set to 1,000,000 for those on classical simulators in Fig. 3. The classical generation of Haar random unitaries in Fig. 4 is performed by QR factorization to random complex matrices with i.i.d. Gaussian entries according to the recipe in ref. [53].

DATA AVAILABILITY

The experimental data that support the finding are available from the authors upon request.

CODE AVAILABILITY

The codes that support the finding are available from the authors upon request.

Received: 15 September 2021; Accepted: 26 September 2022;
Published online: 09 November 2022

REFERENCES

1. Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.* **106**, 180504 (2011).
2. Blume-Kohout, R. et al. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nat. Commun.* **8**, 1–13 (2017).
3. Boixo, S. et al. Characterizing quantum supremacy in near-term devices. *Nat. Phys.* **14**, 595–600 (2018).
4. Cross, A. W., Bishop, L. S., Sheldon, S., Nation, P. D. & Gambetta, J. M. Validating quantum computers using randomized model circuits. *Phys. Rev. A* **100**, 032328 (2019).
5. Erhard, A. et al. Characterizing large-scale quantum computers via cycle benchmarking. *Nat. Commun.* **10**, 5347 (2019).
6. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
7. Proctor, T., Rudinger, K., Young, K., Nielsen, E. & Blume-Kohout, R. Measuring the capabilities of quantum computers. *Nat. Phys.* **18**, 75–79 (2022).
8. Cornelissen, A., Bausch, J. & Gilyén, A. Scalable benchmarks for gate-based quantum computers. Preprint at https://arxiv.org/abs/2104.10698 (2021).
9. Dong, Y. & Lin, L. Random circuit block-encoded matrix and a proposal of quantum linpack benchmark. *Phys. Rev. A* **103**, 062412 (2021).
10. Proctor, T., Seritan, S., Rudinger, K., Nielsen, E., Blume-Kohout, R. & Young, K. Scalable Randomized Benchmarking of Quantum Computers Using Mirror Circuits. *Phys. Rev. Lett.* **129**, 150502 (2022).
11. Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).
12. Kitaev, A. Y. Quantum measurements and the abelian stabilizer problem. Preprint at https://arxiv.org/abs/quant-ph/9511026 (1995).
13. Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.* **103**, 150502 (2009).
14. Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).
15. Berry, D. W., Ahokas, G., Cleve, R. & Sanders, B. C. Efficient quantum algorithms for simulating sparse Hamiltonians. *Commun. Math. Phys.* **270**, 359–371 (2007).
16. Berry, D. W., Childs, A. M., Cleve, R., Kothari, R. & Somma, R. D. Simulating Hamiltonian dynamics with a truncated Taylor series. *Phys. Rev. Lett.* **114**, 090502 (2015).
17. Low, G. H. & Chuang, I. L. Optimal hamiltonian simulation by quantum signal processing. *Phys. Rev. Lett.* **118**, 010501 (2017).
18. Low, G. H. & Wiebe, N. Hamiltonian simulation in the interaction picture. Preprint at https://arxiv.org/abs/1805.00675 (2019).
19. Campbell, E. Random compiler for fast Hamiltonian simulation. *Phys. Rev. Lett.* **123**, 070503 (2019).
20. Berry, D. W. & Childs, A. M. Black-box Hamiltonian simulation and unitary implementation. *Quantum Inf. Comput.* **12**, 29–62 (2012).
21. Berry, D. W., Cleve, R. & Gharibian, S. Gate-efficient discrete simulations of continuous-time quantum query algorithms. *Quantum Inf. Comput.* **14**, 1–30 (2014).
22. Berry, D. W., Childs, A. M., & Kothari, R. Hamiltonian simulation with nearly optimal dependence on all parameters. In *Proc 56th IEEE Symposium on Foundations of Computer Science* (792–809) (Institute of Electrical and Electronics Engineers; Piscataway; New Jersey; United States, 2015).
23. Childs, A. M., Maslov, D., Nam, Y., Ross, N. J. & Su, Y. Toward the first quantum simulation with quantum speedup. *Proc. Nat. Acad. Sci. USA* **115**, 9456–9461 (2018).
24. Childs, A. M., Ostrander, A. & Su, Y. Faster quantum simulation by randomization. *Quantum* **3**, 182 (2019).

25. Low, G. H. Hamiltonian simulation with nearly optimal dependence on spectral norm. In *Proc 51st Annual ACM SIGACT Symposium on Theory of Computing* (491–502) (Association for Computing Machinery; New York; NY; United States, 2019).
26. Childs, A. M. & Su, Y. Nearly optimal lattice simulation by product formulas. *Phys. Rev. Lett.* **123**, 050503 (2019).
27. Childs, A. M., Su, Y., Tran, M. C., Wiebe, N. & Zhu, S. Theory of Trotter error with commutator scaling. *Phys. Rev. X* **11**, 011020 (2021).
28. Chen, C.-F., Huang, H.-Y., Kueng, R. & Tropp, J. A. Concentration for random product formulas. *Pys. Rev. X Quantum* **2**, 040305 (2021).
29. Şahinoğlu, B. & Somma, R. D. Hamiltonian simulation in the low-energy subspace. *npj Quantum Inf.* **7**, 1–5 (2021).
30. An, D., Fang, D. & Lin, L. Time-dependent unbounded hamiltonian simulation with vector norm scaling. *Quantum* **5**, 459 (2021).
31. Su, Y., Berry, D. W., Wiebe, N., Rubin, N. & Babbush, R. Fault-tolerant quantum simulations of chemistry in first quantization. *Phys. Rev. X Quantum* **2**, 040332 (2021).
32. Gilyén, A., Su, Y., Low, G. H. & Wiebe, N. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proc 51st Annual ACM SIGACT Symposium on Theory of Computing* (193–204) (Association for Computing Machinery; New York; NY; United States, 2019).
33. Dong, Y., Meng, X., Whaley, K. B. & Lin, L. Efficient phase factor evaluation in quantum signal processing. *Phys. Rev. A* **103**, 042419 (2021).
34. Lin, L. & Tong, Y. Optimal quantum eigenstate filtering with application to solving quantum linear systems. *Quantum* **4**, 361 (2020).
35. Tong, Y., An, D., Wiebe, N. & Lin, L. Fast inversion, preconditioned quantum linear system solvers, fast green's-function computation, and fast evaluation of matrix functions. *Phys. Rev. A* **104**, 032422 (2021).
36. Lin, L. & Tong, Y. Near-optimal ground state preparation. *Quantum* **4**, 372 (2020).
37. Martyn, J. M., Rossi, Z. M., Tan, A. K. & Chuang, I. L. Grand unification of quantum algorithms. *Phys. Rev. X Quantum* **2**, 040203 (2021).
38. Villalonga, B. et al. Establishing the quantum supremacy frontier with a 281 pflop/s simulation. *Quantum Sci. Technol.* **5**, 034003 (2020).
39. Mehta, M. L. *Random Matrices* (Elsevier, 2004).
40. Aaronson, S. & Gunn, S. On the classical hardness of spoofing linear cross-entropy benchmarking. Preprint at https://arxiv.org/abs/1910.12085 (2019).
41. Aaronson, S. & Chen, L. Complexity theoretic foundations of quantum supremacy experiments. Preprint at https://arxiv.org/abs/1612.05903 (2016).
42. Google Quantum AI Exponential suppression of bit or phase errors with cyclic error correction. *Nature* **595**, 383–387 (2021).
43. Mi, X. et al. Information scrambling in quantum circuits. *Science* **374**, 1479–1483 (2021).
44. Brown, W. & Fawzi, O. Decoupling with Random Quantum Circuits. *Commun. Math. Phys.* **340**, 867–900 (2015).
45. Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing* (333–342) (Association for Computing Machinery; New York; NY; United States, 2011).
46. Bremner, M. J., Montanaro, A. & Shepherd, D. J. Average-case complexity versus approximate simulation of commuting quantum computations. *Phys. Rev. Lett.* **117**, 080501 (2016).
47. Haferkamp, J. et al. Closing gaps of a quantum advantage with short-time hamiltonian dynamics. *Phys. Rev. Letters* **125**, 250501 (2020).
48. Gao, X., Wang, S.-T. & Duan, L.-M. Quantum supremacy for simulating a translation-invariant ising spin model. *Phys. Rev. Lett.* **118**, 040502 (2017).
49. Bermejo-Vega, J., Hangleiter, D., Schwarz, M., Raussendorf, R. & Eisert, J. Architectures for quantum simulation showing a quantum speedup. *Phys. Rev. X* **8**, 021010 (2018).
50. Gao, X. et al. Limitations of linear cross-entropy as a measure for quantum advantage. Preprint at https://arxiv.org/abs/2112.01657 (2021).
51. Choi, J. et al. Emergent randomness and benchmarking from many-body quantum chaos. Preprint at https://arxiv.org/abs/2103.03535 (2021).
52. Abraham, H. et al. Qiskit: An Open-source Framework for Quantum Computing (Zenodo, 2019).
53. Mezzadri, F. How to generate random matrices from the classical compact groups. *Notices of the American Mathematical Society* **54**, 592–604 (2007).

ACKNOWLEDGEMENTS

This work was partially supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator (B.W., L.L.) and a Google Quantum Research Award (Y.D., B.W., L.L.), by the Department of Energy under grant DE-SC0017867, and by the Department of Energy under the Center for Advanced Mathematics for Energy Research Applications (CAMERA) program (L.L.). We thank András Gilyén, Xun Gao, Yunchao Liu, Murphy Niu, and Jiahao Yao for helpful discussions, and in particular thank Timothy Proctor for extended discussions on noise models.

AUTHOR CONTRIBUTIONS

Y.D. and L.L. designed the Hamiltonian simulation benchmark and proved its theoretical properties. Y.D., B.W. and L.L. designed the experiments. Y.D. carried out classical simulations and IBM-Q experiments. All authors contributed to the discussion of results and writing 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/s41534-022-00636-x.

Correspondence and requests for materials should be addressed to Lin Lin.

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.