nature physics

Article

https://doi.org/10.1038/s41567-025-03025-1

Certifying almost all quantum states with few single-qubit measurements

Received: 21 July 2024

Accepted: 6 August 2025

Published online: 08 September 2025

Check for updates

Hsin-Yuan Huang [1,2,3] ✉, John Preskill [2,4] ✉ & Mehdi Soleimanifar [2] ✉

Certifying that an n-qubit state synthesized in the laboratory is close to a given target state is a fundamental task in quantum information science. However, existing rigorous protocols applicable to general target states have potentially prohibitive resource requirements in the form of either deep quantum circuits or exponentially many single-qubit measurements. Here we prove that almost all n-qubit target states, including those with exponential circuit complexity, can be certified from only $O(n^2)$ single-qubit measurements. Given access to the target state's amplitudes, our protocol requires only $O(n^3)$ classical computation. This result is established by a technique that relates certification to the mixing time of a random walk. Our protocol has applications for benchmarking quantum systems, for optimizing quantum circuits to generate a desired target state and for learning and verifying neural networks, tensor networks and various other representations of quantum states using only single-qubit measurements. We show that such verified representations can be used to efficiently predict highly non-local properties of a synthesized state that would otherwise require an exponential number of measurements on the state. We demonstrate these applications in numerical experiments with up to 120 qubits and observe an advantage over existing methods such as cross-entropy benchmarking.

Our empirical knowledge of a quantum system often relies on statistical comparisons with a target model of its state. An essential challenge involves determining whether an n-qubit quantum state ρ prepared in the laboratory achieves high fidelity $\langle\psi|\rho|\psi\rangle$ with an intended target state $|\psi\rangle$. In many practical scenarios, our knowledge of the target state is encoded through an algorithmic representation that computes complex amplitudes $\langle x|\psi\rangle$ for any computational basis configuration $x \in \{0,1\}^n$, up to a global phase or normalization.

When the system size n remains modest, complete amplitude information can be stored directly in classical memory. For larger quantum systems, sophisticated computational frameworks, such as neural network quantum states or tensor network decompositions, provide efficient algorithmic access[1–11]. Many quantum states of theoretical and experimental interest, including phase states, coherent Gibbs states, graph states, Greenberger–Horne–Zeilinger (GHZ) states, W states and symmetric many-body configurations, admit compact algorithmic descriptions that directly enable efficient amplitude computation.

With algorithmic access to the target state $|\psi\rangle$ and experimental access to copies of the laboratory state ρ, quantum state certification seeks to establish that the fidelity $\langle\psi|\rho|\psi\rangle$ is sufficiently close to unity. Practical implementation demands protocols with minimal experimental overhead and computational complexity. We therefore concentrate on measurement schemes employing only single-qubit operations without preceding entangling circuits. Such measurements integrate seamlessly with diverse experimental architectures and generate classical data amenable to efficient analysis. The central question we address is: how many single-qubit measurements are required to reliably assess whether $\langle\psi|\rho|\psi\rangle$ is close to unity?

Extensive investigation has left unresolved whether certification can be achieved using (1) polynomially many single-qubit measurements

[1]Google Quantum AI, Venice, CA, USA. [2]California Institute of Technology, Pasadena, CA, USA. [3]Massachusetts Institute of Technology, Cambridge, MA, USA. [4]AWS Center for Quantum Computing, Pasadena, CA, USA. ✉e-mail: hsinyuan@caltech.edu; preskill@caltech.edu; mehdi@caltech.edu

on copies of an arbitrary n-qubit state ρ while (2) verifying overlap with a generic highly entangled target state $|\psi\rangle$. These requirements appear contradictory; since entanglement distributes quantum correlations non-locally across multiple qubits, local measurement statistics seemingly cannot capture global properties, such as fidelity $\langle\psi|\rho|\psi\rangle$, with highly entangled states $|\psi\rangle$.

Current rigorous certification methods face considerable limitations, namely, they either necessitate complex quantum operations before measurement[12–17], require exponentially scaling resources[12,18–29] or remain restricted to specialized state classes[25,27,29–37]. Linear cross-entropy benchmarking (XEB)[38–41] provides a popular heuristic for estimating fidelity with complex entangled states. Nevertheless, XEB maintains theoretical validity only under white noise assumptions, where the laboratory state equals the target state corrupted by global depolarizing channels. Under other noise models, XEB fails to provide reliable estimates and may indicate perfect fidelity even when actual overlap vanishes, particularly under dephasing or coherent error processes. See Supplementary Section A for a detailed review of prior work. We resolve this challenge by introducing shadow overlap, a fidelity surrogate that can certify almost all quantum states with few single-qubit measurements.

Main results

We introduce a simple procedure for certifying whether an n-qubit (possibly mixed) laboratory state ρ is close to a target pure state $|\psi\rangle$. Our protocol (detailed in the next section and illustrated in Fig. 1) estimates the shadow overlap, a quantity that serves as a faithful surrogate for fidelity $\langle\psi|\rho|\psi\rangle$, while being efficiently estimable through single-qubit Pauli measurements. Each measurement collapses a single copy of the laboratory state ρ and produces a single number ω. The expectation $\mathbb{E}[\omega]$, which we refer to as the shadow overlap, exhibits a precise relationship with fidelity $\langle\psi|\rho|\psi\rangle$ characterized by

$$\mathbb{E}[\omega] \geq 1-\epsilon \quad \Rightarrow \quad \langle\psi|\rho|\psi\rangle \geq 1-\tau\epsilon, \tag{1}$$

$$\langle\psi|\rho|\psi\rangle \geq 1-\epsilon \quad \Rightarrow \quad \mathbb{E}[\omega] \geq 1-\epsilon. \tag{2}$$

The parameter τ denotes the relaxation time of a random walk on the Boolean hypercube designed to sample from the computational basis probability distribution $\pi(x) := |\langle x|\psi\rangle|^2$ determined by the target state $|\psi\rangle$. This random walk construction is formalized in the next section. The relaxation time τ quantifies how rapidly the random walk converges to its stationary distribution, which can be measured in terms of the number of random bit flips, and relates to the spectral gap of the associated transition matrix. When τ is polynomially bounded, shadow overlap provides an efficient proxy for global fidelity through local measurement statistics.

This connection between shadow overlap and fidelity enables efficient quantum certification for target states with polynomial relaxation times, as formalized in our first main result:

Theorem 1. *(Certification of quantum states, informal) Given an n-qubit target state $|\psi\rangle$ with a relaxation time $\tau \geq 1$, there is a certification procedure that performs single-qubit Pauli measurements on $T = \mathcal{O}(\tau^2/\epsilon^2)$ samples of an unknown n-qubit state ρ and, with high probability, outputs failed if the fidelity is low $\langle\psi|\rho|\psi\rangle < 1-\epsilon$ and outputs are certified if the fidelity is high $\langle\psi|\rho|\psi\rangle \geq 1-\frac{\epsilon}{2\tau}$. When one allows more general single-qubit measurements on the unknown n-qubit state ρ, the sample complexity can be improved to $T = \mathcal{O}(\tau/\epsilon)$.*

Our certification protocol requires $\mathcal{O}(T)$ queries to the algorithmic representation of the target state $|\psi\rangle$. When both $\tau \leq \mathrm{poly}(n)$ and efficient algorithmic access are available, the entire procedure is computationally efficient. We establish polynomial relaxation time bounds $\tau \leq \mathrm{poly}(n)$ for diverse structured quantum families, making our certification scheme practically efficient for these states. Specifically, phase states and GHZ-like states achieve $\tau = \mathcal{O}(n)$ (demonstrated in Supplementary Sections G and I), while ground states of gapped sign-free κ-local Hamiltonians as well as any states that match the probability distribution $|\langle x|\psi\rangle|^2$ for these ground states satisfy $\tau = \mathcal{O}(n^\kappa)$ (proven in Supplementary Section H).

Although not every quantum state in a fixed computational basis admits polynomial relaxation times, our rigorous analysis employing sophisticated random walk techniques on the Boolean hypercube $\{0,1\}^n$ establishes that $\tau \leq \tau^* = \mathcal{O}(n^2)$ for almost all n-qubit pure states drawn from the Haar measure. This fundamental result is proven in Supplementary Section D. Combined with Theorem 1, we demonstrate that almost all Haar-random quantum states, including those with extremely high entanglement and exponential circuit depth, can be certified using polynomially many single-qubit measurements.

Theorem 2. *(Certification of almost all quantum states, informal) All except an exponentially small fraction $2^{-\Omega(n)}$ of n-qubit pure states $|\psi\rangle$ sampled from the Haar measure can be certified using $\mathcal{O}(n^2/\epsilon)$ single-qubit measurements. The protocol outputs failed with high probability when $\langle\psi|\rho|\psi\rangle < 1-\epsilon$ and were certified when $\langle\psi|\rho|\psi\rangle \geq 1-\frac{\epsilon}{2\tau^*}$, where $\tau^* = \mathcal{O}(n^2)$.*

Theorem 2 establishes that our protocol reliably determines whether an arbitrary laboratory state ρ exhibits high fidelity with the target state $|\psi\rangle$, independent of the underlying noise mechanisms affecting ρ. This noise independence contrasts sharply with XEB, which operates reliably only under global depolarizing noise assumptions and can erroneously certify laboratory states that have neither entanglement nor high fidelity with the target state (Fig. 3). The robustness and generality of our approach enable diverse applications that we examine in 'Applications' section, including machine learning-based quantum tomography, experimental quantum device benchmarking and quantum circuit optimization for state preparation.

Certification procedure

Our certification protocol, illustrated in Fig. 1, operates through a simple two-phase process. During the first phase, we obtain a single copy of the laboratory state ρ and randomly select one qubit from the n available, which we label as $k \in \{1, \ldots, n\}$. All qubits except the chosen qubit k are measured in the computational basis, yielding measurement outcomes denoted collectively as $z \in \{0,1\}^{n-1}$. Subsequently, we randomly select one of the three Pauli measurements (X, Y or Z) with equal probability and perform the measurement on qubit k. We denote the resulting post-measurement state of this qubit as $|s\rangle$.

After collecting measurement data from the laboratory state ρ, our protocol transitions to the second phase involving queries to the algorithmic representation of the target state $|\psi\rangle$. We assume access to a computational model that encodes the target state $|\psi\rangle = \sum_{x\in\{0,1\}^n} \psi(x)|x\rangle$ through its computational basis amplitudes $\psi(x)$. When provided with any input string $x \in \{0,1\}^n$, this model returns a complex value proportional to the corresponding amplitude $\psi(x)$. Importantly, we do not require these outputs to be properly normalized. Since computing the global normalization factor $\sum_{x\in\{0,1\}^n} |\psi(x)|^2$ is generally computationally intractable, this flexibility broadens the practical applicability of our approach.

During the query phase, we interrogate the model twice using carefully constructed input strings $z^{(0)}$ and $z^{(1)}$. Each query string $z^{(a)}$ (where $a \in \{0,1\}$) has its kth bit set to the value a while all remaining $n-1$ bits match the measurement outcomes z obtained in the first phase. These query results enable us to construct the conditional post-measurement single-qubit state as follows:

$$|\Psi_{k,z}\rangle := \frac{\Psi(z^{(0)})|0\rangle + \Psi(z^{(1)})|1\rangle}{\sqrt{|\Psi(z^{(0)})|^2 + |\Psi(z^{(1)})|^2}}. \tag{3}$$

The final step combines measurement and query data to evaluate the local overlap as

https://doi.org/10.1038/s41567-025-03025-1

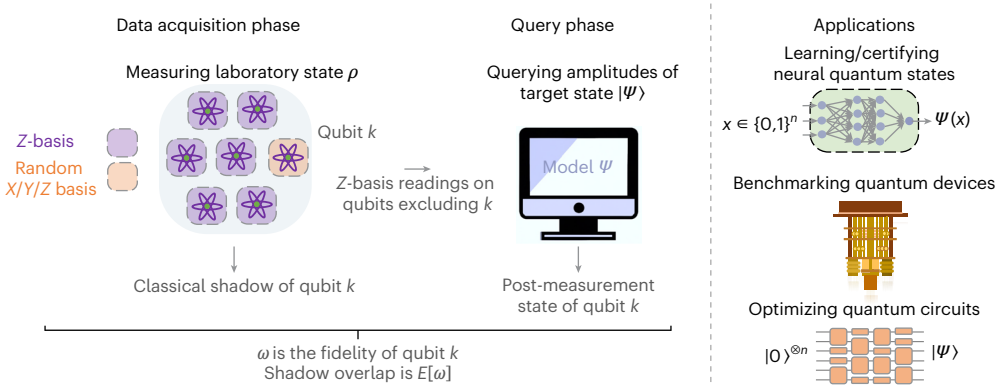

Fig. 1 | Estimating the shadow overlap. The data collection phase: for each copy of the laboratory state ρ, a random qubit k is selected. All qubits except k are measured in the Z basis. Qubit k is measured in a random X, Y or Z basis to obtain its classical shadow. Query phase: by querying the amplitudes of the target state $|\psi\rangle$ twice, the ideal post-measurement state $|\psi_{k,z}\rangle$ of qubit k is found. Using the classical shadow of qubit k from the laboratory state, its overlap ω with $|\psi_{k,z}\rangle$ is evaluated. Finally, the shadow overlap E[ω] is estimated by averaging ω across all copies.

$$\omega := \langle\Psi_{k,z}|\,(3\,|s\rangle\,\langle s| - \mathbb{1})\,|\Psi_{k,z}\rangle\,. \tag{4}$$

When both query results vanish, that is, $\Psi(z^{(0)}) = \Psi(z^{(1)}) = 0$, we set $\omega = 0$. We execute this complete two-phase protocol across T independent copies of the laboratory state ρ, generating a sequence of local overlaps $\omega_1, \ldots, \omega_T$. Our final estimate of the shadow overlap between ρ and the target state is the empirical average $\hat{\omega} = \frac{1}{T}\sum_{t=1}^{T}\omega_t$. This basic protocol admits natural generalizations. In the Methods, we develop an extended framework that measures $m > 1$ qubits in random Pauli bases rather than restricting to a single qubit, providing additional flexibility for specialized applications.

Analysis overview

The detailed analysis of our certification protocol is presented in the Methods and Supplementary Section C. Here, we provide intuitive insight into the underlying principles of our approach. To understand the core mechanism, consider an idealized variant of our protocol where qubit k is measured using the projective basis $\{|\Psi_{k,z}\rangle\langle\Psi_{k,z}|, \mathbb{1} - |\Psi_{k,z}\rangle\langle\Psi_{k,z}|\}$ rather than randomized Pauli measurements, with $|\Psi_{k,z}\rangle$ defined as in equation (3). In this idealized scenario, we assign $\omega = 1$ when the measurement projects onto $|\Psi_{k,z}\rangle\langle\Psi_{k,z}|$ and $\omega = 0$ otherwise. When the laboratory state is perfect, that is, $\rho = |\psi\rangle\langle\psi|$, and we measure $n-1$ qubits in the computational basis yielding outcome z, the conditional state of qubit k precisely matches $|\Psi_{k,z}\rangle\langle\Psi_{k,z}|$. Consequently, every measurement yields $\omega = 1$, resulting in an estimated shadow overlap $\hat{\omega} = \frac{1}{T}\sum_{t=1}^{T}\omega_t = \langle\psi|\,\rho\,|\psi\rangle = 1$.

The protocol described earlier employs randomized Pauli measurements instead of this idealized projective measurement, computing the local overlap through expression (4). The crucial insight is that the operator $(3\,|s\rangle\langle s| - 1)$ reproduces the desired projector $|\Psi_{k,z}\rangle\langle\Psi_{k,z}|$ in expectation over random Pauli basis choices and measurement outcomes. This property, detailed in Supplementary Section B, constitutes the fundamental principle behind classical shadow tomography[12]. For perfect laboratory state, that is, $\rho = |\psi\rangle\langle\psi|$, this ensures that $\mathbb{E}[\omega] = 1$. The empirical average $\hat{\omega}$ is close to unity when ρ is close to $|\psi\rangle\langle\psi|$ and sufficient samples T are collected. This analysis confirms that our protocol will successfully certify high-fidelity laboratory states.

However, the critical question remains whether the protocol be deceived into certifying low-fidelity states. Addressing this concern requires recognizing that $\mathbb{E}[\omega] = \mathrm{Tr}\,[L\rho]$ for a specific observable L that can be constructed through queries to the target state's algorithmic representation. Ideally, we want this observable to be the projector $|\psi\rangle\langle\psi|$ onto the target state $|\psi\rangle$. Our analysis in Supplementary Section C demonstrates that the actual observable L satisfies $L\,|\psi\rangle = |\psi\rangle$ and $\langle\psi^{\perp}|\,L\,|\psi^{\perp}\rangle \le 1 - \frac{1}{\tau}$ for any orthogonal state $|\psi^{\perp}\rangle$, where $\tau \ge 1$ is a state-dependent parameter. This makes L an approximate projector onto the target state $|\psi\rangle$. Consequently, we can differentiate between high-fidelity and low-fidelity laboratory states ρ by measuring $T = \mathcal{O}(\tau^2)$ samples of ρ.

A key insight, proven in Supplementary Section C, is that the observable L shares identical eigenvalues with the normalized transition matrix P of a suitably constructed random walk (or Markov chain) on the Boolean hypercube $\{0,1\}^n$. The parameter $1/\tau$ corresponds to the spectral gap of P, which is the difference between its largest and second-largest eigenvalues. This implies that τ is the relaxation time of this random walk. This connection is useful because it allows us to exploit the rich literature on Markov chain relaxation times to establish upper bounds on τ and, consequently, on our certification protocol's sample complexity. We emphasize that this random walk serves purely as an analytical tool and is not implemented as part of the protocol itself.

The transition matrix P is determined by the computational basis measurement distribution $\pi(x) = |\langle x|\psi\rangle|^2$ of the target state $|\psi\rangle$. For a general state $|\psi\rangle = \sum_{x\in\{0,1\}^n}\sqrt{\pi(x)}e^{i\phi(x)}\,|x\rangle$, the random walk transitions between vertices $x, y \in \{0,1\}^n$ according to

$$P(x,y) = \begin{cases} \frac{1}{n}\frac{\pi(y)}{\pi(x)+\pi(y)} & x \sim y, \\ \frac{1}{n}\sum\limits_{x':x'\sim x}\frac{\pi(x)}{\pi(x)+\pi(x')} & x = y, \\ 0 & \text{otherwise}, \end{cases} \tag{5}$$

where adjacent vertices $x \sim y$ differ in exactly one bit position. This construction ensures that the stationary distribution assigns probability $\pi(x)$ to vertex x. For the uniform distribution $\pi(x) = \frac{1}{2^n}$, this reduces to the standard lazy random walk on the Boolean hypercube, where each step either remains at the current vertex with probability 1/2 or moves to a uniformly chosen neighbour with probability 1/2.

By exploiting results on random walks, we can analyse our certification protocol's performance. We identify numerous quantum state families achieving $\tau \le \mathrm{poly}(n)$, rendering our certification approach efficient. These include generic Haar-random states (Supplementary Section D), structured entangled families such as quantum phase states (Supplementary Section G), gapped ground states (Supplementary Section H) and GHZ-like states (Supplementary Section I).

Applications

We stated in Theorem 2 that almost all quantum states drawn from the Haar measure can be certified using shadow overlaps, which can

https://doi.org/10.1038/s41567-025-03025-1

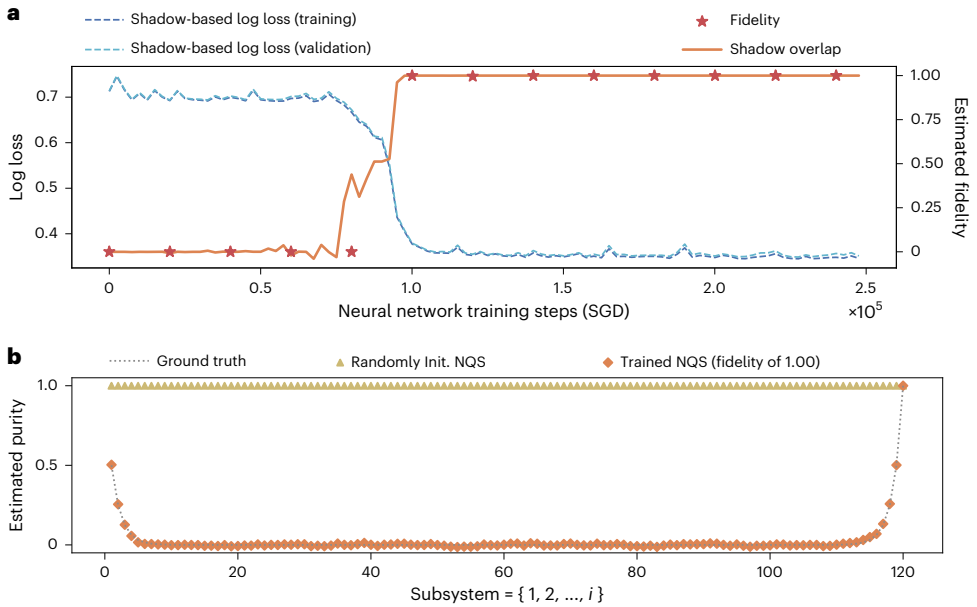

Fig. 2 | Neural network quantum state tomography: training and certifying a neural quantum state with the shadow overlap. a, A dual-input neural network is trained to learn a quantum phase state as in equation (6) with random phases $\phi(x)$ on $n = 120$ qubits using single-qubit measurements. The neural network admits two inputs $x_0, x_1 \in \{0,1\}^n$, which differ only in one bit, and computes $\frac{\langle x_0|\psi\rangle}{\langle x_1|\psi\rangle}$ as its output. Applying this neural network architecture n times allow us to compute the amplitude $\langle x|\psi\rangle$ for a given x. The dual-input neural quantum states can be trained using a shadow-based log loss, leveraging data acquired by single-qubit measurements as prescribed in the shadow overlap protocol. A shadow-based loss function trains the model on 50,000 measurement data acquired in the shadow overlap estimation, which consist of tuples $(x_0, x_1, |\phi(x_0) - \phi(x_1)|)$, representing the phase difference between two adjacent strings x_0 and x_1. The log loss is minimized via SGD. The model is then certified using fidelity and shadow overlap on a separate data set of size 10,000. **b**, The trained neural quantum state is used to estimate the subsystem purity of the random phase state, exhibiting a high degree of entanglement compared with a randomly initialized neural quantum state (Randomly Init. NQS).

be reliably estimated with few single-qubit measurements. In what follows, we give an overview of various interesting applications of the shadow overlap formalism.

Neural network quantum state tomography

The development of quantum technologies increasingly relies on classical computational models capable of capturing complex quantum phenomena while remaining tractable for numerical analysis. Among the most promising approaches are machine learning representations that provide algorithmic access to quantum state amplitudes. These models enable efficient computation of amplitudes $\psi(x) \in \mathbb{C}$ for any computational basis element $x \in \{0,1\}^n$, up to a global phase.

Contemporary approaches include neural quantum states and tensor network decompositions, which have been extensively studied in recent literature[1,4–11,42,43]. Polynomial-size neural networks can efficiently evaluate quantum amplitudes $\psi(x)$ for represented n-qubit states. Similarly, tensor networks with tractable contraction schemes, including matrix product states[2] and tree tensor networks[3], provide efficient amplitude computation of the represented n-qubit state.

Learning ML models of quantum states

Our certification scheme yields an algorithm for learning ML models of quantum states with rigorous sample complexity guarantees. This is achieved using learning by hypothesis selection, which can generally be applied to a set of models $\{\Psi_1, \ldots, \Psi_M\}$ each describing an n-qubit state $|\psi_i\rangle$, for $i \in [M]$. Our objective is to use the measurement data obtained from identical copies of a state ρ and learn a model Ψ_i among $i \in [M]$ which achieves the highest overlap $\langle\psi_i|\rho|\psi_i\rangle$. This approach to learning is relevant in applications where we either naturally have a set of M hypotheses (for example, from different theories describing the physics of a quantum system) or where we can obtain such a discrete set by casting a covering net (or carrying out some form of coarse-graining) over a larger and more expressive family of models.

We show in Supplemetary Section J that, assuming the fast mixing condition for the set of models $\{\Psi_1, \cdots, \Psi_M\}$, we can use the shadow overlap to learn a model that achieves a high fidelity with the laboratory state ρ using $\mathcal{O}(\log M)$ copies of ρ. In Supplemetary Section J, we also give a concrete application of this scheme for learning a feedforward neural network representation of a quantum state. We show that the sample complexity of this problem scales as $\tilde{\mathcal{O}}\left(nL^3W^3s^{2L}\right)$ for a network of depth L, width W and spectral norm s that takes n-bit strings as input. In Supplemetary Section J.3, we discuss another application of this learning algorithm in the context of gapped ground states.

Training neural quantum states with shadow overlap

Although hypothesis selection provides a learning scheme with a rigorous sample complexity, the run time of this algorithm scales linearly with the number of models M, rendering it inefficient for many applications where M grows exponentially with the number of qubits n. In practice, though, as shown in Fig. 2 and detailed in Supplemetary Section L.1, we can use the shadow overlap along with the stochastic gradient descent (SGD) to efficiently train and certify an ML model of a quantum state.

To this end, we consider training a neural network representation of an n-qubit state $|\psi\rangle$ using a shadow-based log loss and the data acquired by single-qubit measurements in the shadow overlap protocol. Figure 2 shows an application of this scheme to learning highly entangled phase states

$$|\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} e^{i\phi(x)}|x\rangle \qquad (,6)$$

with random binary phases $\phi(x) \in \{0, \pi\}$ on $n = 120$ qubits. Such quantum states are indistinguishable from states with exponentially large circuit complexity[44,45], and exhibit volume-law scaling of entanglement[46]. The findings reported in Fig. 2 indicate that, beyond a certain

https://doi.org/10.1038/s41567-025-03025-1

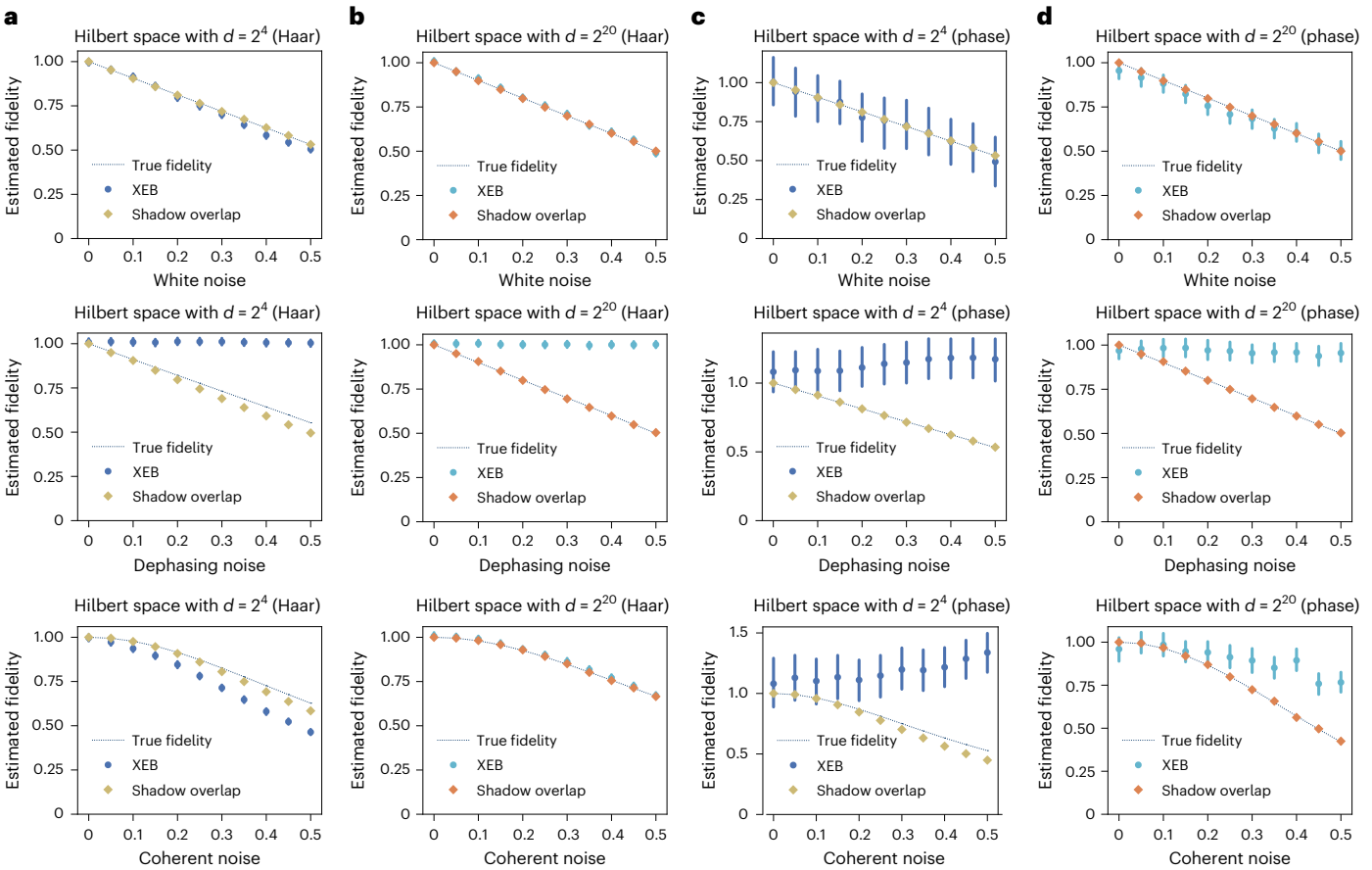

Fig. 3 | Benchmarking with the shadow overlap. a–d, A comparison of the (normalized) shadow overlap, fidelity and XEB in benchmarking noisy quantum states under white noise (that is, global depolarizing noise), dephasing noise and coherent noise (realized as small Gaussian errors in the wavefunction), for a 4-qubit (**a**) and a 20-qubit (**b**) Haar random state, along with a 4-qubit (**c**) and a 20-qubit (**d**) random phase state of the form $U_{\text{phase}} \cdot \otimes_{i=1}^{n} |\psi_i\rangle$, where $|\psi_i\rangle$ are single-qubit states with random real amplitudes and U_{phase} is diagonal with random complex phases. In this numerical experiment, the shadow overlap is normalized as detailed in Supplemetary Section F. Error bars indicate statistical measurement errors. Each independent run uses a sample size of $N = 50$, and the average and variance are computed over 500 independent runs. XEB tracks the fidelity well under white noise but fails under dephasing and coherent noise. In contrast, shadow overlap performs robustly across all regimes with lower variance.

training threshold, the model attains a fidelity very close to 1.00 with the target state. As explained next, this performance can also be certified using the shadow overlap, as an efficient alternative to the fidelity.

Certifying ML models

One drawback of machine learning models for quantum states is that their training usually relies on heuristic algorithms. The absence of performance guarantees highlights the need for certification procedures capable of efficiently verifying the accuracy of the trained models. The result of Theorem 1 can be restated in terms of certifying the overlap between an n-qubit state $|\psi\rangle$ and its trained ML model with a relaxation time τ. This is achieved using single-qubit Pauli measurements performed independently on $\mathcal{O}(\tau^2/\epsilon^2)$ copies of $|\psi\rangle$ along with two queries to the trained ML model per $|\psi\rangle$ copy.

Figure 2 shows a numerical implementation of this certification procedure for a dual-input neural network representation of a 120-qubit random phase state. After training the neural net with 50,000 measurements using shadow-overlap-based stochastic gradient decent, we estimate and compare the shadow overlap of the resulting model with its fidelity. We observe that the predicted shadow overlap closely mirrors the fidelity, serving as an effective proxy.

Estimating sparse observables

The certified ML models of quantum states can be employed to statistically estimate many properties of interest[9,11] if in addition to query access, we assume the models are also equipped with sampling access: the ability to sample from the measurement distribution corresponding to $|\psi(x)|^2 := |\langle x|\psi\rangle|^2$. The sampling access can be obtained in various ways as follows: Markov chain sampling via running the random walk defined in equation (5) or the Metropolis–Hastings algorithm, measuring the certified state ρ in the computational basis, or using autoregressive methods to obtain direct sampling access[47].

In Supplemetary Section K, we show how to apply a verified ML model of a quantum state with query and sampling access to estimate the expectation value of any sparse observable G, such as the energy of a local Hamiltonian, or highly non-local properties, such as Rényi entanglement entropies, up to an error ϵ with a number of samples that scales as $T = \mathcal{O}\left(\langle\psi| G^2 |\psi\rangle / \epsilon^2\right)$. When no certified ML model is available, estimating certain non-linear observables, such as the subsystem purity $\text{Tr}(\rho_A^2)$, requires a number of samples that are exponential in the size of the subsystem A; for example, see ref. 48 for an exponential lower bound that applies to any single-copy measurements, and ref. 12 for an upper bound via the classical shadow formalism. In Supplemetary Section K, the same task can be conducted using a verified ML model with a sample complexity $\mathcal{O}\left(1/\epsilon^2\right)$, independent of the system size.

In Fig. 2, we demonstrate this feature with a numerical experiment on the trained neural network representation of the random phase state in equation (6). The purity $\text{Tr}(\rho_A^2)$ of the phase state is estimated for subsystems of size $|A| \in \{1, \ldots, 120\}$, confirming that the state of the

https://doi.org/10.1038/s41567-025-03025-1

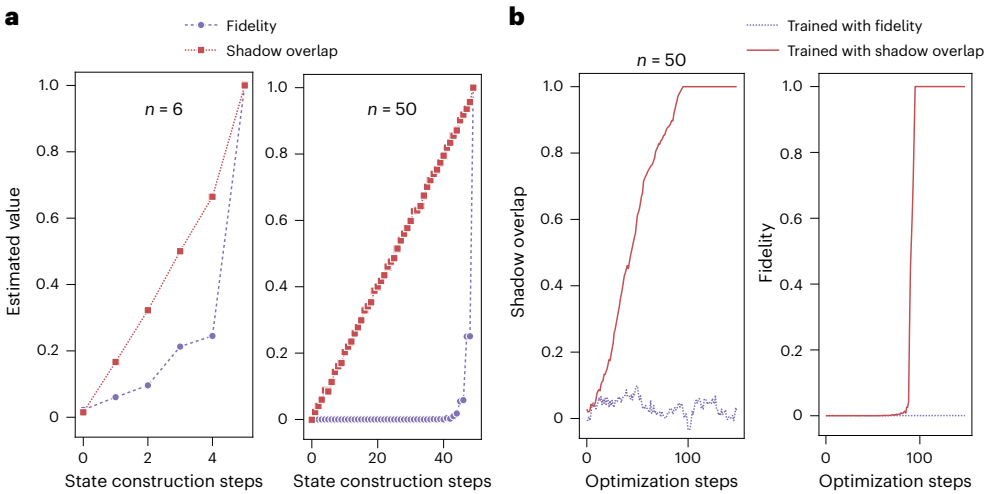

Fig. 4 | Optimizing quantum circuits for state preparation. Training a low-depth quantum circuit consisting of Hadamard, controlled-Z and T gates to prepare a target state $|\psi\rangle$ given as a matrix product state. The target state corresponds to the output of a one-dimensional IQP circuit[57] infused with random T gates. The training is first performed by variationally optimizing for maximum shadow overlap. We then assess this method's performance against fidelity-based training. Changes in both fidelity and shadow overlap are monitored across optimization steps. **a**, We see that, akin to the linear decrease in Hamming distance between two binary strings as suitable bits are flipped, under shadow-overlap-based training the deviation of the shadow overlap from 1 decreases linearly as suitable gates are added to the circuit. This contrasts with the fidelity, which fails to exhibit a steady, gradual increase as the number of state construction steps increases. **b**, Because the optimization landscape of fidelity has a barren plateau, training with fidelity fails to find a high-fidelity state-preparation circuit. In contrast, training with shadow overlap successfully finds a high-fidelity circuit. The improvement in optimization landscape is comparable to that achieved by local fidelity[53–56], but with a key advantage, namely, while local fidelity requires either deep quantum circuits or exponentially many single-qubit measurements for general target states $|\psi\rangle$, shadow overlap needs only polynomially many single-qubit measurements.

subsystems is close to the maximally mixed state as guaranteed by the rigorous proofs.

Benchmarking quantum devices

Certifying the fidelity between a state ρ prepared using a quantum device and a known quantum state $|\psi\rangle$ offers a rigorous approach for benchmarking quantum machines. However, the exponential resources and the high level of control needed for estimating fidelity limit the applicability of this approach in practice. To address these challenges, a number of studies have proposed and deployed other statistical quantities that act as a form of proxy for the fidelity in the situations often encountered practically[38,39,49,50]. One such proxy is the XEB, a prominent metric employed in the evaluation of quantum supremacy experiments with local random quantum circuits[38,49]. Given an n-qubit target state $|\psi\rangle = \sum_{x\in\{0,1\}^n} \sqrt{\pi(x)}e^{i\phi(x)}|x\rangle$ and the laboratory state ρ, their XEB score is defined by

$$\mathrm{XEB} = \frac{2^n E_{x\sim\langle x|\rho|x\rangle}\pi(x) - 1}{2^n E_{x\sim\pi(x)}\pi(x) - 1}, \tag{7}$$

where $\mathrm{E}_{x\text{-}\pi(x)}$ and $E_{x\sim\langle x|\rho|x\rangle}$ denote the expectation when x is sampled according to measurement distribution $\pi(x)$ and $\langle x|\rho|x\rangle$, respectively. The XEB score is designed to ensure that XEB matches the fidelity when the laboratory state $\rho = (1-p)|\psi\rangle\langle\psi| + p\frac{I}{2^n}$ equals the target state subject to white noise with parameter p; see refs. 38,49,51 and Supplemetary Section A. From equation (7), we can also observe that XEB entirely ignores the phase information $e^{i\phi(x)}$ in the target state $|\psi\rangle$. Consequently, it is not a reliable proxy for the fidelity under various physically relevant noise models, such as dephasing or coherent noise.

In Fig. 3, we compare the performance of shadow overlap, fidelity and XEB. We observe that, when the laboratory state ρ is subject to white noise (first row), XEB performs well for Haar random states but exhibits much larger variance for random phase states. However, under dephasing noise (second row), XEB incorrectly reports near perfect fidelity, even when the actual fidelity is far from 1. In the presence of coherent noise (third row), XEB can also deviate substantially from fidelity. In contrast, the shadow overlap tracks the fidelity quite well across different noise regimes and system sizes for white noise, dephasing noise and coherent noise. Refer to Supplemetary Sections A, F and L.2 for a more detailed discussion.

Optimizing quantum circuits for state preparation

Many variational quantum algorithms employ the fidelity between two quantum states as their cost function. However, these fidelity-based cost functions suffer from exponentially vanishing gradients, a phenomenon known as barren plateaus[52–55], and require a high sample complexity for statistical estimation. Shadow overlap $\mathrm{E}[\omega]$ offers an alternative to the fidelity $\langle\psi|\rho|\psi\rangle$ in such algorithms. Beyond demanding a substantially lower sample complexity, the shadow overlap may provide an improved optimization landscape with non-vanishing gradients. Specifically, when the target state exhibits no global correlations and its probability amplitudes are well distributed across the Boolean hypercube $\{0,1\}^n$, the shadow overlap behaves similarly to the Hamming distance. This correspondence becomes exact for bit strings in the X-basis, where the Hamming distance and shadow overlap are identical.

Consider the illustrative case of $|\psi\rangle = |+\rangle^{\otimes n}$. As detailed in the Methods, shadow overlap is the expectation value $\mathrm{Tr}(L\rho)$ of the observable $L = \frac{1}{n}\sum_{i=1}^{n}|+\rangle\langle+|_i \otimes \mathbb{1}_{\setminus i}$. This observable is a sum of local terms, offering favourable optimization features compared with the non-local observable $|+\rangle\langle+|^{\otimes n}$ used in fidelity estimation. A related approach in quantum machine learning[53–56], known as local fidelity, employs the observable $\frac{1}{n}\sum_{i=1}^{n} U(|0\rangle\langle 0|_i \otimes \mathbb{1}_{\setminus i})U^\dagger$ for a target state $|\psi\rangle = U|0^n\rangle$. Both local fidelity and shadow overlap offer improved optimization landscapes. However, local fidelity is experimentally challenging. While it can be efficiently estimated using classical shadows[12] for low-entanglement states $|\psi\rangle$, the general case is highly inefficient. For $|\psi\rangle = U|0^n\rangle$ with deep circuit U, local fidelity is a sum of highly non-local terms: $U(|0\rangle\langle 0|_i \otimes \mathbb{1}_{\setminus i})U^\dagger$. Consequently, measuring local fidelity requires either implementing a deep circuit $U^\dagger$ or performing an exponential number of single-qubit measurements (see Supplemetary Section A for further details). In contrast, shadow overlap maintains

high measurement efficiency, requiring only a polynomial number of single-qubit measurements.

In a numerical experiment presented in Fig. 4 and discussed in Supplemetary Section L.3, we investigate this feature of the shadow overlap in the context of training quantum circuits to optimally prepare a target state, corresponding to the output of a one-dimensional IQP circuit[57] infused with random T gates. When employing $n = 50$ qubits, fidelity-based training encounters barren plateaus, whereas shadow overlap-based training successfully prepares the target state with a fidelity very close to 1.

Outlook

Our certification protocol based on the shadow overlap demonstrates that almost all quantum states can be certified using polynomially many independent single-qubit Pauli measurements. Further extending the reach of our certification protocol raises many interesting open questions.

Quantum states with fast relaxation times

What families of quantum states provably admit a poly(n) relaxation time with respect to the Markov chain (5) introduced in our analysis? We show that Haar random quantum states exhibits a relaxation time bounded by $\tau \leq O(n^2)$. Can we show that states prepared with (random) quantum circuits of arbitrary depth satisfy a relaxation time $\tau \leq \text{poly}(n)$?

Certifying any states with few single-qubit measurements

While our protocol successfully certifies almost all states, the ultimate goal would be universal certification of any state using only poly(n) single-qubit measurements. A recent follow-up work[58] demonstrates that adaptive single-qubit measurements with mid-circuit classical feedforward can indeed certify arbitrary target states. However, such adaptive protocols require sophisticated experimental control and are equivalent to universal quantum computation[59]. Furthermore, their result assumes a more powerful form of algorithmic access to the target state than in our setting, namely, the ability to compute amplitudes in any product basis rather than just the computational basis. This stronger assumption limits the applicability of their protocol to some tasks discussed in 'Applications' section. This raises a refined open question, namely: can we achieve universal state certification using only poly(n) non-adaptive single-qubit measurements without mid-circuit classical feedforward?

Mixed states

Can a similar certification protocol be developed when the target state belongs to a certain family of mixed quantum states? In particular, can almost all approximately low-rank mixed states be certified with few single-qubit measurements?

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-025-03025-1.

References

1. Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
2. Verstraete, F. & Cirac, J. I. Matrix product states represent ground states faithfully. *Phys. Rev. B* **73**, 094423 (2006).
3. Shi, Y. -Y., Duan, L. -M. & Vidal, G. Classical simulation of quantum many-body systems with a tree tensor network. *Phys. Rev. A* **74**, 022320 (2006).
4. Pfau, D., Spencer, J. S., Matthews, A. G. D. G. & Foulkes, W. M. C. Ab initio solution of the many-electron schrödinger equation with deep neural networks. *Phys. Rev. Res.* **2**, 033429 (2020).
5. Hibat-Allah, M., Ganahl, M., Hayward, L. E., Melko, R. G. & Carrasquilla, J. Recurrent neural network wave functions. *Phys. Rev. Res.* **2**, 023358 (2020).
6. Carrasquilla, J., Torlai, G., Melko, R. G. & Aolita, L. Reconstructing quantum states with generative models. *Nat. Mach. Intell.* **1**, 155–161 (2019).
7. Melko, R. G., Carleo, G., Carrasquilla, J. & Cirac, J. I. Restricted Boltzmann machines in quantum physics. *Nat. Phys.* **15**, 887–892 (2019).
8. Torlai, G. et al. Neural-network quantum state tomography. *Nat. Phys.* **14**, 447–450 (2018).
9. Iouchtchenko, D., Gonthier, J. érômeF., Perdomo-Ortiz, A. & Melko, R. G. Neural network enhanced measurement efficiency for molecular groundstates. *Mach. Learn. Sci. Technol.* **4**, 015016 (2023).
10. Zhao, H., Carleo, G. & Vicentini, F. Empirical sample complexity of neural network mixed state reconstruction. *Quantum* **8**, 1358 (2024).
11. Torlai, G., Mazzola, G., Carleo, G. & Mezzacapo, A Precise measurement of quantum observables with neural-network estimators. *Phys. Rev. Res.* **2**, 022060 (2020).
12. Huang, H. -Y., Kueng, R. & Preskill, J Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
13. Montanaro, A. & de Wolf, R. A survey of quantum property testing. Preprint at https://arxiv.org/abs/1310.2035 (2013).
14. O'Donnell, R. & Wright, J. Efficient quantum tomography. In *Proc. 48th Annual ACM SIGACT Symposium on Theory of Computing* 899–912 (Association for Computing Machinery, 2016).
15. O'Donnell, R. & Wright, J. Efficient quantum tomography ii. In *Proc. 49th Annual ACM SIGACT Symposium on Theory of Computing* 962–974 (Association for Computing Machinery, 2017).
16. Haah, J., Harrow, A. W., Ji, Z., Wu, X. & Yu, N. Sample-optimal tomography of quantum states. *IEEE Trans. Inf. Theory* **63**, 5628–5641 (2017).
17. Buadescu, C., O'Donnell, R. & Wright, J. Quantum state certification. In *STOC'19—Proc. 51st Annual ACM SIGACT Symposium on Theory of Computing* 503–514 (Association for Computing Machinery, 2019).
18. Renes, J. M., Blume-Kohout, R., Scott, A. J. & Caves, C. M. Symmetric informationally complete quantum measurements. *J. Math. Phys.* **45**, 2171–2180 (2004).
19. Scott, A. J. & Grassl, M. Symmetric informationally complete positive-operator-valued measures: a new computer study. *J. Math. Phys.* https://doi.org/10.1063/1.3374022 (2010).
20. Scott, A. J. Tight informationally complete quantum measurements. *J. Phys. A* **39**, 13507 (2006).
21. Scott, A.J. SICs: extending the list of solutions. Preprint at https://arxiv.org/abs/1703.03993 (2017).
22. Kueng, R., Rauhut, H. & Terstiege, U. Low rank matrix recovery from rank one measurements. *Appl. Comput. Harmon. Anal.* **42**, 88–116 (2017).
23. Guta, M., Kahn, J., Kueng, R. & Tropp, J. A. Fast state tomography with optimal error bounds. *J. Phys. A* **53**, 204001 (2020).
24. Brandão, F. G. S. L., Kueng, R & França, D. S. Fast and robust quantum state tomography from few basis measurements. In *16th Conference on the Theory of Quantum Computation, Communication and Cryptography* 7:1–7:13 (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2021).
25. Flammia, S. T. & Liu, Y. -K. Direct fidelity estimation from few pauli measurements. *Phys. Rev. Lett.* **106**, 230501 (2011).
26. da Silva, M. P., Landon-Cardinal, O. & Poulin, D. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.* **107**, 210404 (2011).

27. Aolita, L., Gogolin, C., Kliesch, M. & Eisert, J. Reliable quantum certification of photonic state preparations. *Nat. Commun.* **6**, 8498 (2015).
28. Elben, A. et al. The randomized measurement toolbox. *Nat. Rev. Phys.* **5**, 9–24 (2023).
29. Pallister, S., Linden, N. & Montanaro, A. Optimal verification of entangled states with local measurements. *Phys. Rev. Lett.* **120**, 170502 (2018).
30. Tóth, G. éza & Gühne, O. Detecting genuine multipartite entanglement with two local measurements. *Phys. Rev. Lett.* **94**, 060501 (2005).
31. Gluza, M., Kliesch, M., Eisert, J. & Aolita, L. Fidelity witnesses for fermionic quantum simulations. *Phys Rev. Lett.* **120**, 190501 (2018).
32. Takeuchi, Y. & Morimae, T. Verification of many-qubit states. *Phys. Rev. X* **8**, 021060 (2018).
33. Huang, H. Y. et al. Learning shallow quantum circuits. In *Proc. 56th Annual ACM SIGACT Symposium on Theory of Computing* 1343–1351 (Association for Computing Machinery, 2024).
34. Zhu, H. & Hayashi, M. Optimal verification and fidelity estimation of maximally entangled states. *Phys. Rev. A* **99**, 052346 (2019).
35. Wang, K. & Hayashi, M. Optimal verification of two-qubit pure states. *Phys. Rev. A* **100**, 032315 (2019).
36. Morimae, T., Takeuchi, Y. & Hayashi, M. Verification of hypergraph states. *Phys. Rev. A* **96**, 062321 (2017).
37. Zhu, H. & Hayashi, M. Efficient verification of hypergraph states. *Phys. Rev. Appl.* **12**, 054047 (2019).
38. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
39. Choi, J. et al. Preparing random states and benchmarking with many-body quantum chaos. *Nature* **613**, 468–473 (2023).
40. Cotler, J. S. et al. Emergent quantum state designs from individual many-body wave functions. *PRX Quantum* **4**, 010311 (2023).
41. Dalzell, A. M., Hunter-Jones, N. & Brandão, F. G. S. L. Random quantum circuits transform local noise into global white noise. *Commun. Math. Phys.* **405**, 78 (2024).
42. Sharir, O., Shashua, A. & Carleo, G. Neural tensor contractions and the expressive power of deep neural quantum states. *Phys. Rev. B* **106**, 205136 (2022).
43. Wu, D., Rossi, R., Vicentini, F. & Carleo, G. From tensor-network quantum states to tensorial recurrent neural networks. *Phys. Rev. Res.* **5**, L032001 (2023).
44. Ji, Z., Liu, Y. -K., & Song, F. Pseudorandom quantum states. In *Advances in Cryptology–CRYPTO 2018: 38th Annual International Cryptology Conference, Proceedings, Part III* 126–152 (Springer, 2018).
45. Brakerski, Z. & Shmueli, O. (pseudo) random quantum states with binary phase. In *Theory of Cryptography Conference* 229–250 (Springer, 2019).
46. Aaronson, S. et al. Quantum pseudoentanglement. In *15th Innovations in Theoretical Computer Science Conference* 21 (Univ California, Berkley, 2024).
47. Sharir, O., Levine, Y., Wies, N., Carleo, G. & Shashua, A. Deep autoregressive models for the efficient variational simulation of many-body quantum systems. *Phys. Rev. Lett.* **124**, 020503 (2020).
48. Chen, S., Cotler, J., Huang, H. -Y., & Li, J. Exponential separations between learning with and without quantum memory. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 574–585. IEEE, 2022.
49. Boixo, S. et al. Characterizing quantum supremacy in near-term devices. *Nat. Phys.* **14**, 595–600 (2018).
50. Mark, D. K., Choi, J., Shaw, A. L., Endres, M. & Choi, S. Benchmarking quantum simulators using ergodic quantum dynamics. *Phys. Rev. Lett.* **131**, 110601 (2023).
51. Cross entropy benchmarking theory. *Google Quantum AI* https://quantumai.google/cirq/noise/qcvv/xeb_theory (2024).
52. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **9**, 1–6 (2018).
53. Cerezo, M., Sone, A., Volkoff, T., Cincio, L. & Coles, P. J. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nat. Commun.* **12**, 1791 (2021).
54. Caro, M. C. et al. Out-of-distribution generalization for learning quantum dynamics. *Nat. Commun.* **14**, 3751 (2023).
55. Jerbi, S. et al. The power and limitations of learning quantum dynamics incoherently. Preprint at https://arxiv.org/2303.12834 (2023).
56. Khatri, S. et al. Quantum-assisted quantum compiling. *Quantum* **3**, 140 (2019).
57. Bremner, M. J., Jozsa, R. & Shepherd, D. J. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proc. R. Soc. A* **467**, 459–472 (2011).
58. Gupta, M., He, W. & O'Donnell, R. Few single-qubit measurements suffice to certify any quantum state. Preprint at https://arxiv.org/abs/2506.11355 (2025).
59. Briegel, H. J., Browne, D. E., Dür, W., Raussendorf, R. & Van den Nest, M. Measurement-based quantum computation. *Nat. Phys.* **5**, 19–26 (2009).

Methods

Our certification protocol achieves high efficiency by exploiting a deep connection between quantum state fidelity and Markov chain mixing times. We present the complete framework underlying our results, showing how the shadow overlap naturally serves as a proxy for fidelity and how its performance is linked to the mixing properties of an associated Markov chain, leading to the proof of Theorem 1.

Certification protocol

We analyse the performance of a generalized version of the shadow overlap protocol introduced in the main text. In this version detailed below, we choose m qubits uniformly at random and measure each in a randomized basis. This section includes the proof of Theorem 1 of the main text, corresponding to the $m = 1$ case, as well as the equivalent statement for the general level-m protocol. Given an n-qubit target state $|\psi\rangle$ specified by a model Ψ, and an arbitrary laboratory state ρ, the level-m protocol proceeds as follows:

(1) Among the total n qubits of ρ, choose a uniformly random subset of size at most m qubits. Denote these qubits by $k = \{k_1, \ldots, k_r\}$ where $r \leq m$ is the size of the subset.
(2) Perform single-qubit Z-basis measurements on all but qubits $k_1, \ldots, k_r$ of ρ. Denote the measurement outcomes collectively by $z_k \in \{0,1\}^{n-r}$.
(3) For each qubit $k_1, \ldots, k_r$, choose an X, Y or Z-basis measurement uniformly at random and measure that qubit of ρ. Denote the post-measurement state of the qubits $k_1, \ldots, k_r$ by $|s_1\rangle, \ldots, |s_r\rangle$, respectively. Compute the classical shadow

$$\boldsymbol{\sigma} = (3|\boldsymbol{s_1}\rangle\langle\boldsymbol{s_1}| - \mathbb{1}) \otimes (3|\boldsymbol{s_2}\rangle\langle\boldsymbol{s_2}| - \mathbb{1}) \otimes \cdots \otimes (3|\boldsymbol{s_r}\rangle\langle\boldsymbol{s_r}| - \mathbb{1}). \tag{8}$$

(4) Query the model Ψ for all choices of r-bit strings ℓ_1 and ℓ_2 that differ exactly in r bits (that is, $\ell_1, \ell_2 \in \{0,1\}^r$ and $\mathrm{dist}(\ell_1, \ell_2) = r$) to obtain the normalized states

$$\left|\Psi_{z_k}^{\ell_1,\ell_2}\right\rangle := \frac{\Psi\left(z_k^{(\ell_1)}\right)|\ell_1\rangle + \Psi\left(z_k^{(\ell_2)}\right)|\ell_2\rangle}{\sqrt{\left|\Psi\left(z_k^{(\ell_1)}\right)\right|^2 + \left|\Psi\left(z_k^{(\ell_2)}\right)\right|^2}}. \tag{9}$$

Here the n-bit string $\boldsymbol{z_k^{(\ell)}}$ matches $\ell \in \{0,1\}^r$ on bits $k_1, \ldots, k_r$ and equals $z_k \in \{0,1\}^{n-r}$ on the remaining $n-r$ bits.
(5) Compute the overlap

$$\omega := \mathrm{Tr}(L_{z_k}\sigma) \text{ with } L_{z_k} := \sum_{\substack{\ell_1,\ell_2 \in \{0,1\}^r \\ \mathrm{dist}(\ell_1,\ell_2)=r}} \left|\Psi_{z_k}^{\ell_1,\ell_2}\right\rangle\left\langle\Psi_{z_k}^{\ell_1,\ell_2}\right|. \tag{10}$$

(6) Repeat steps 1–5 for T times to obtain overlaps $\omega_1, \ldots, \omega_T$. Report the estimated shadow overlap given by $\hat{\boldsymbol{\omega}} := \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\omega_t}$.
(7) If the estimated shadow overlap $\hat{\boldsymbol{\omega}} \geq 1 - \frac{3\epsilon}{4\tau}$, the output is certified. Otherwise, the output failed.

Certification efficiency and mixing time

Fix a level m for the certification protocol. The measurement distribution $\pi(x) = |\langle x|\psi\rangle|^2$ is a distribution on a graph $G = (V, E)$ where the vertices $V = \{0,1\}^n$ are n-bit strings, and an edge $e = (x, y)$ exists between vertices x and y when they differ in $k \in \{1, \ldots, m\}$ bits. Let $\mathcal{S} := \{x : \pi(x) > 0\}$ denote the support of $\pi(x)$. Let $N = \sum_{k=1}^{m}\binom{n}{k}$ be number of neighbours of each vertex. We define a random walk or a Markov chain on this graph. The transition matrix $P = \sum_{x\in\{0,1\}^n} P(x,y)|x\rangle\langle y|$ of this walk is defined by

$$P(x,y) = \begin{cases} \frac{1}{N}\frac{\pi(y)}{\pi(x)+\pi(y)} & (x,y) \in E, \\ \frac{1}{N}\sum_{x':(x',x)\in E}\frac{\pi(x)}{\pi(x)+\pi(x')} & x = y, \\ 0 & \text{otherwise}. \end{cases} \tag{11}$$

In our application, it is more convenient to consider a normalized version of the transition matrix P given by $S^{\frac{1}{2}}PS^{-\frac{1}{2}}$ where $S^{-\frac{1}{2}} = \sum_{x\in\mathcal{S}}\frac{1}{\sqrt{\pi(x)}}|x\rangle\langle x|$. We claim that the shadow overlap estimated in level m of our protocol equals the expected value of an observable directly linked to the normalized transition matrix.

Proposition 3. Suppose the level-m certification protocol is performed on copies of the state ρ and a model Ψ of the quantum state $|\psi\rangle = \sum_{x\in\{0,1\}^n}\sqrt{\pi(x)}e^{i\phi(x)}|x\rangle$. Define the 'phase matrix' by $F = \sum_{x\in\{0,1\}^n}e^{i\phi(x)}|x\rangle\langle x|$ and let L be the Hermitian operator given by

$$L = FS^{\frac{1}{2}}PS^{-\frac{1}{2}}F^{\dagger}. \tag{12}$$

We have $L|\psi\rangle = |\psi\rangle$ and $\mathrm{Tr}(L\rho) = \mathbf{E}[\omega]$, where $\mathrm{E}[\omega]$ denotes the expected output of the certification protocol.

Proof. The entries of the observable L for any $x \in \mathcal{S}$ are given by

$$\langle x|L|y\rangle = \begin{cases} \frac{1}{N}\frac{\sqrt{\pi(x)\pi(y)}}{\pi(x)+\pi(y)}e^{i(\phi(x)-\phi(y))} & (x,y) \in E, \\ \frac{1}{N}\sum_{x':(x',x)\in E}\frac{\pi(x)}{\pi(x)+\pi(x')} & x = y, \\ 0 & \text{otherwise}. \end{cases} \tag{13}$$

For $x \in \mathcal{S}$, we have

$$\begin{aligned} \langle x|L|\psi\rangle &= \langle x|L|x\rangle\langle x|\psi\rangle + \sum_{y\neq x}\langle x|L|y\rangle\langle y|\psi\rangle \\ &= \frac{1}{N}\sum_{y:(y,x)\in E}\frac{\pi(x)}{\pi(x)+\pi(y)}\sqrt{\pi(x)}e^{i\phi(x)} \\ &\quad + \frac{1}{N}\sum_{y:(y,x)\in E}\frac{\sqrt{\pi(x)\pi(y)}}{\pi(x)+\pi(y)}e^{i(\phi(x)-\phi(y))}\sqrt{\pi(y)}e^{i\phi(y)} \\ &= \frac{1}{N}\sum_{y:(y,x)\in E}\left(\frac{\pi(x)}{\pi(x)+\pi(y)} + \frac{\pi(y)}{\pi(x)+\pi(y)}\right)\sqrt{\pi(x)}e^{i\phi(x)} \\ &= \sqrt{\pi(x)}e^{i\phi(x)}. \end{aligned} \tag{14}$$

This shows that $L|\psi\rangle = |\psi\rangle$. Next, we prove that $\mathrm{Tr}(L\rho) = E[\omega]$. Consider subsets of qubits with size $r \leq m$. There are $N = \sum_{k=1}^{m}\binom{n}{k}$ choices for the location of these qubits. For any r, we enumerate the chosen qubits by $k_1, \ldots, k_r$ and collectively denote them by $k = \{k_1, \ldots, k_r\}$. For a fixed k, the set $\{z_k \in \{0,1\}^{n-r}\}$ denotes all the possible bit strings on the remaining $n-r$ bits. Direct inspection reveals that the observable L corresponding to the model Ψ can be expressed as

$$L = \frac{1}{N}\sum_{r\in[m]}\sum_{k=\{k_1,\ldots,k_r\}}\sum_{z_k\in\{0,1\}^{n-r}}|z_k\rangle\langle z_k| \otimes L_{z_k}, \tag{15}$$

where L_{z_k} is an operator acting on r qubits $\{k_1, \ldots, k_r\}$ and is given by

$$\begin{aligned} L_{z_k} &:= \sum_{\substack{\ell_1,\ell_2 \in \{0,1\}^r \\ \mathrm{dist}(\ell_1,\ell_2)=r}}\left|\Psi_{z_k}^{\ell_1,\ell_2}\right\rangle\left\langle\Psi_{z_k}^{\ell_1,\ell_2}\right| \text{ with } \left|\Psi_{z_k}^{\ell_1,\ell_2}\right\rangle \\ &:= \frac{\Psi\left(z_k^{(\ell_1)}\right)|\ell_1\rangle + \Psi\left(z_k^{(\ell_2)}\right)|\ell_2\rangle}{\sqrt{\left|\Psi\left(z_k^{(\ell_1)}\right)\right|^2 + \left|\Psi\left(z_k^{(\ell_2)}\right)\right|^2}}. \end{aligned} \tag{16}$$

In this expression, the binary string $z_k^{(\ell)}$ equals $\ell \in \{0,1\}^r$ on bits $k_1, \ldots, k_r$ and equals $z_k \in \{0,1\}^{n-r}$ on the remaining $n-r$ bits. Let σ denote the classical shadow obtained after performing randomized Pauli measurements on the post-measurement state of qubits r. That is, if the r single Pauli measurements return states $|s_1\rangle, \ldots, |s_r\rangle$, we set $\sigma = (3|s_1\rangle\langle s_1| - \mathbb{1}) \otimes (3|s_2\rangle\langle s_2| - \mathbb{1}) \otimes \cdots \otimes (3|s_r\rangle\langle s_r| - \mathbb{1})$. It follows from the discussion in Supplemetary Section B that $E_{\text{shadows}}[\sigma] = \frac{\langle z_k|\rho|z_k\rangle}{\text{Tr}(\langle z_k|\rho|z_k\rangle)}$, where the expectation is over Pauli measurements on qubits k. Using this, we can expand $\text{Tr}[L\rho]$ as follows:

$$\begin{aligned}
\text{Tr}[L\rho] &= \frac{1}{N} \sum_{r \in [m]} \sum_{z_k \in \{0,1\}^{n-r}} \text{Tr}(\langle z_k|\rho|z_k\rangle) \text{Tr}\left(L_{z_k} \frac{\langle z_k|\rho|z_k\rangle}{\text{Tr}(\langle z_k|\rho|z_k\rangle)}\right) \\
&\qquad k = \{k_1, \ldots, k_r\} \\
&= E_{k,z_k} \text{Tr}\left(L_{z_k} \frac{\langle z_k|\rho|z_k\rangle}{\text{Tr}(\langle z_k|\rho|z_k\rangle)}\right) \\
&= E_{k,z_k} \text{Tr}\left(L_{z_k} E_{\text{shadows}}[\sigma]\right) \\
&= E_{k,z_k} E_{\text{shadows}}[\text{Tr}\left(L_{z_k}\sigma\right)] = E[\omega].
\end{aligned}$$

In the last expression, the expectation is with respect to the location of the Pauli Z measurements, their outcomes, as well as the randomized measurements on the remaining qubits.

Shadow overlap as a proxy for fidelity

When we first average over the classical shadows, the shadow overlap $E[\omega]$ is equal to the average overlap between the postselected state on ρ and the postselected state on the target state $|\psi\rangle$. Hence $0 \le E[\omega] = \text{Tr}(L\rho) \le 1$ for any state ρ. This implies that $0 \preccurlyeq L \preccurlyeq I$.

Theorem 4. *Let $\lambda_1 = 1 - \frac{1}{\tau}$ be the second largest eigenvalue of the transition matrix P defined with respect to the measurement distribution $\pi(x)$ of the state $|\psi\rangle$. The shadow overlap satisfies*

$$\text{if} \quad E[\omega] \ge 1 - \epsilon \quad \text{then we have} \quad \langle\psi|\rho|\psi\rangle \ge 1 - \tau\epsilon; \tag{17}$$

$$\text{if} \quad \langle\psi|\rho|\psi\rangle \ge 1 - \epsilon \quad \text{then we have} \quad E[\omega] \ge 1 - \epsilon. \tag{18}$$

This implies that we can check if fidelity is close to one by checking if shadow overlap is close to one.

Proof. We first study the spectrum of the observable L. From the previous theorem and the fact that $0 \preccurlyeq L \preccurlyeq I$, the eigenvalues of L are given by $1 = \lambda_0 \ge \lambda_1 \ge \lambda_2 \ge \cdots \ge 0$. The two operators P and L are related by a similarity transformation. Hence, they have the same set of eigenvalues. Let $|\lambda_i\rangle$ denote the eigenstate of observable L corresponding to the eigenvalue λ_i. We claim that the top eigenstate $|\lambda_0\rangle$ of the operator L is the quantum state $|\psi\rangle$. This can be seen by the direct calculation or by noting that the measurement distribution $\pi(x)$ is the unique stationary distribution of P. Hence, we have $P|\pi\rangle = |\pi\rangle$, where $|\pi\rangle = \sum_{x\in\{0,1\}^n} \sqrt{\pi(x)}|x\rangle$. From this and the fact that $|\pi\rangle = F^\dagger|\psi\rangle$, we have $L|\psi\rangle = |\psi\rangle$, as claimed.

Now we prove the implication stated in equation (17). From Proposition 3, we know that $E[\omega] = \text{Tr}[L\rho]$. Assuming $E[\omega] \ge 1 - \epsilon$, we have

$$\begin{aligned}
1 - \epsilon \le E[\omega] &= \text{Tr}[L\rho] \\
&= \langle\psi|\rho|\psi\rangle + \sum_{i\ge 1} \lambda_i \langle\lambda_i|\rho|\lambda_i\rangle \qquad \text{using } |\lambda_0\rangle = |\psi\rangle \\
&\le \langle\psi|\rho|\psi\rangle + \lambda_1 \sum_{i\ge 1} \langle\lambda_i|\rho|\lambda_i\rangle \qquad \text{definition of } \lambda_1 \\
&\le \langle\psi|\rho|\psi\rangle + \lambda_1(1 - \langle\psi|\rho|\psi\rangle) \qquad \text{since } \text{Tr}(\rho) = 1.
\end{aligned}$$

By rearranging the two sides of the inequality, we arrive at the bound $\langle\psi|\rho|\psi\rangle \ge 1 - \frac{\epsilon}{1-\lambda_1} = 1 - \tau\epsilon$. We next prove the implication stated in (18).

$$\begin{aligned}
E[\omega] &= \text{Tr}[L\rho] \\
&= \langle\psi|\rho|\psi\rangle + \textstyle\sum_{i\ge 1} \lambda_i \langle\lambda_i|\rho|\lambda_i\rangle \text{ using } |\lambda_0\rangle = |\psi\rangle \\
&\ge \langle\psi|\rho|\psi\rangle \ge 1 - \epsilon.
\end{aligned}$$

This concludes the proof of this theorem.

Sample complexity

As we show in Supplemetary Section C, using equations (18) and (17), we can certify the overlap between the target state $|\psi\rangle$ and an unknown state ρ using $T = 2^{2m+4} \cdot \frac{\tau^2}{\epsilon^2} \cdot \log(\frac{2}{\delta})$ copies of the state ρ with probability at least $1 - \delta$. Moreover, we can improve the dependency of the sample complexity on τ/ϵ for the level $m = 1$ protocol by replacing randomized Pauli measurements with measurement in the basis $\{|\Psi_{k,z}\rangle\langle\Psi_{k,z}|, \mathbb{1} - |\Psi_{k,z}\rangle\langle\Psi_{k,z}|\}$. In this setting, we can show that $T = \mathcal{O}(\frac{\tau}{\epsilon} \cdot \log(\frac{1}{\delta}))$ samples of the unknown state ρ are sufficient for the certification protocol to succeed with probability at least $1 - \delta$.

Data availability

The data used to generate all figures are openly available via Zenodo[60]. In the Zenodo repository, the data associated with the three main numerical experiments can be found in text files in `neural-quantum-state/`, pickle files in `benchmarking-vs-XEB/` and pickle files in `benchmarking-vs-shadow/`.

Code availability

The code for conducting the numerical experiments and generating the data in this work is available via Zenodo at https://zenodo.org/records/15873712 (ref. 60). The current development version is available via GitHub at https://github.com/hsinyuan-huang/certify-quantum-states.

References

60. Huang, H. -Y., Preskill, J. & Soleimanifar, M. Certifying almost all quantum states with few single-qubit measurements. *Zenodo* https://doi.org/10.5281/zenodo.15873712 (2025).

Acknowledgements

The authors thank A. Anshu, R. Babbush, M. Broughton, D. Gosset, R. Kothari and J.R. McClean for their valuable input and inspiring discussions. H.-Y.H. is supported by a Google PhD fellowship and a MediaTek Research Young Scholarship. H.-Y.H. acknowledges the visiting associate position at the Massachusetts Institute of Technology. J.P. acknowledges support from the US Department of Energy Office of Science, Office of Advanced Scientific Computing Research (DE-NA0003525, DE-SC0020290), the US Department of Energy, Office of Science, National Quantum Information Science Research Centres, Quantum Systems Accelerator and the National Science Foundation (PHY-1733907). M.S. was supported by AWS Quantum Postdoctoral Scholarship and funding from the National Science Foundation NSF CAREER award CCF-2048204. The Institute for Quantum Information and Matter is an NSF Physics Frontiers Centre.

Author contributions

All authors contributed equally and are listed alphabetically by last name. H.-Y.H., J.P. and M.S. conceived the project. H.-Y.H. and M.S. developed the core mathematical theory and proofs. H.-Y.H., J.P. and M.S. conceived the applications. M.S. extended the mathematical theory to support these applications. H.-Y.H. designed and conducted the numerical experiments illustrating

the applications. M.S. drafted the initial manuscript. All authors contributed to writing, reviewing and finalizing 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/s41567-025-03025-1.

Correspondence and requests for materials should be addressed to Hsin-Yuan Huang, John Preskill or Mehdi Soleimanifar.

Peer review information *Nature Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.