ARTICLE OPEN

Check for updates

Detecting and quantifying entanglement on near-term quantum devices

Kun Wang[1], Zhixin Song[1], Xuanqiang Zhao[1], Zihe Wang[1] and Xin Wang [1]✉

Quantum entanglement is a key resource in quantum technology, and its quantification is a vital task in the current noisy intermediate-scale quantum (NISQ) era. This paper combines hybrid quantum-classical computation and quasi-probability decomposition to propose two variational quantum algorithms, called variational entanglement detection (VED) and variational logarithmic negativity estimation (VLNE), for detecting and quantifying entanglement on near-term quantum devices, respectively. VED makes use of the positive map criterion and works as follows. Firstly, it decomposes a positive map into a combination of quantum operations implementable on near-term quantum devices. It then variationally estimates the minimal eigenvalue of the final state, obtained by executing these implementable operations on the target state and averaging the output states. Deterministic and probabilistic methods are proposed to compute the average. At last, it asserts that the target state is entangled if the optimized minimal eigenvalue is negative. VLNE builds upon a linear decomposition of the transpose map into Pauli terms and the recently proposed trace distance estimation algorithm. It variationally estimates the well-known logarithmic negativity entanglement measure and could be applied to quantify entanglement on near-term quantum devices. Experimental and numerical results on the Bell state, isotropic states, and Breuer states show the validity of the proposed entanglement detection and quantification methods.

INTRODUCTION

It is widely believed that we are now in the noisy intermediate-scale quantum (NISQ) era[1], where quantum computers with 50–100 qubits are available while noise in quantum gates severely limits the quantum circuits that can be executed reliably. It thus becomes important to make the best use of today's NISQ devices to design practical applications. One promising scheme for near-term quantum applications is the variational quantum algorithms (VQA)[2], which have been applied to solve many tasks including Hamiltonian ground and excited states preparation[3,4], quantum state distance estimation[5,6], and quantum data compression[7–9]. These variational quantum algorithms involve evaluating and optimizing loss functions that depend on parameters in parameterized quantum circuits (PQC). They are regarded as well-suited for execution on NISQ devices by combining quantum computers with classical computers. We refer the readers to[10,11] for a detailed review on VQA.

Quantum entanglement[12], the most nonclassical manifestation of quantum mechanics, has been identified as invaluable resource enabling a tremendous number of tasks ranging from quantum information processing[13,14], quantum cryptography[15,16], quantum algorithms[17,18], quantum communication[19], to measurement-based quantum computing[20,21]. As so, the ability to manipulate quantum entanglement[12,22] is the cornerstone to achieve real applications of quantum technologies. A number of theoretical and experimental methods have been proposed in the past 20 years for entanglement detection and quantification[12,23,24]. For example, entanglement can be detected via entanglement witnesses[25,26], Bell's inequalities[27], realignment criterion[28,29], range criterion[30], and majorization criterion[31]. These methods commonly assume that prior information about the target state is known. A direct way to obtain such information is to perform quantum state tomography and reconstruct the density matrix[32,33]. However, tomography becomes unrealistic as the

number of required measurement settings scales exponentially with the size of the system. Briefly speaking, though there are many methods proposed for detecting and quantifying quantum entanglement, they are not specially designed for near-term quantum devices and thus are not directly applicable in most cases, rendering reliable detection and quantification of quantum entanglement on near-term quantum devices a vital challenge. Recently there are a number of works aiming to overcome this challenge[34–37]. The core idea of all these approaches is to perform measurements in randomly sampled bases, leading to ensembles of measurement outcomes whose statistical correlations provide a fingerprint of the system's entanglement.

In this paper, we combine VQA and the quasi-probability decomposition technique[38–44] to propose the variational entanglement detection (VED) and variational logarithmic negativity estimation (VLNE) algorithms, contributing new approaches to detect and quantify quantum entanglement on near-term quantum devices, respectively. VED uses criteria based on positive maps as a bridge and works as follows. Given an unknown target bipartite quantum state, it firstly decomposes the chosen positive map into a linear combination of NISQ implementable quantum operations. Then, it variationally estimates the minimal eigenvalue of the final state, which is obtained by executing these quantum operations on the target state and averaging the output states. Two methods are proposed to compute the average: the first one averages the output states according to the quasi-probability distribution, and the second one estimates the average via the sampling technique and is probabilistic. At last, it asserts that the target state is entangled if the optimized minimal eigenvalue is negative. Following the idea of VED, VLNE variationally computes the well-known log-negativity entanglement measure, building on a linear decomposition of the transpose map into Pauli terms and the recently proposed trace distance estimation algorithm.

[1]Institute for Quantum Computing, Baidu Research, 100193 Beijing, China. ✉email: wangxin73@baidu.com

Experimental and numerical results reveal the validity of the proposed entanglement detection and quantification methods.

RESULTS

Quantum entanglement detection

In this section, we integrate variational quantum algorithms with the quasi-probability decomposition technique[38–44] to propose a bipartite entanglement detection framework specially designed for near-term quantum devices, using positive map criterion as a bridge. For simplicity, we assume A and B are two n-qubit quantum systems throughout this section. However, we remark that the proposed framework can be applied to bipartite systems with different dimensions directly.

Let Δ be a discrete set of quantum operations that are implementable in the near-term quantum devices. For example, one may choose Δ to be the set of implementable operations introduced in[42,43]. Alternatively, one may set Δ to be the set of Pauli channels induced by Pauli operators from the Pauli set [see Supplementary Note 1]. For a positive (but not completely positive) and trace-preserving map $\mathcal{N}_{B \to B}$, we assume that it can be decomposed w.r.t. Δ as

$$\mathcal{N}(\cdot) = \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \mathcal{O}(\cdot), \quad r_{\mathcal{O}} \in \mathbb{R}. \tag{1}$$

Note that such a decomposition always exists if Δ contains a universal basis[42]. The trace-preserving condition imposes $\sum_{\mathcal{O}} r_{\mathcal{O}} = 1$. We emphasize that there must exist negative coefficients $r_{\mathcal{O}}$ since otherwise, $\mathcal{N}$ is completely positive. Given a bipartite quantum state ρ_{AB}, we have

$$\sigma_{AB} := \mathcal{N}_{B \to B}(\rho_{AB}) = \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \mathcal{O}_{B \to B}(\rho_{AB}). \tag{2}$$

To see if ρ_{AB} can be detected by $\mathcal{N}$, i.e., if ρ_{AB} is entangled from $\mathcal{N}$'s perspective, we need to check if the output state σ_{AB} has a negative eigenvalue or not. Denote by $\lambda_{\min}(\sigma_{AB})$ the smallest eigenvalue of σ_{AB}. By the positive map criterion, if ρ_{AB} is separable, then it must hold that $\lambda_{\min}(\sigma_{AB}) \geq 0$. Equivalently, if $\lambda_{\min}(\sigma_{AB}) < 0$, we safely conclude that ρ_{AB} is entangled and it can be detected by the positive map $\mathcal{N}_{B \to B}$. This highlights the importance of computing or estimating $\lambda_{\min}(\sigma_{AB})$ in entanglement detection.

Deterministic detection. As we have argued, σ_{AB} cannot be obtained directly via $\mathcal{N}(\rho)$ since $\mathcal{N}$ does not represent a physically implementable quantum operation. Fortunately, the decomposition in Eq. (2) empowers us an effective way to simulate the role of $\mathcal{N}$ and reconstruct σ_{AB} as an average of a set of output states, obtained using quantum circuits implementable in near-term devices. This decomposition technique, combined with the variational quantum algorithm, enables a general framework that estimates $\lambda_{\min}(\sigma_{AB})$, whose value can witness the entanglement of the input state ρ_{AB}. We call this framework the variational entanglement detection (VED). The core idea is to use the linear decomposition in Eq. (2) of the target state σ_{AB} and the framework goes as follows. First of all, it holds that (see Supplementary Note 1)

$$\lambda_{\min}(\sigma_{AB}) = \min_{|\psi\rangle_{AB}} \langle \psi | \sigma_{AB} | \psi \rangle \tag{3}$$

$$= \min_{|\psi\rangle_{AB}} \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \langle \psi | \mathcal{O}(\rho_{AB}) | \psi \rangle, \tag{4}$$

where the minimization ranges over all pure bipartite quantum states $|\psi\rangle_{AB}$ in AB. We use a variational quantum circuit with parameters $\boldsymbol{\alpha}$ to prepare the test state $|\psi\rangle$. More precisely, we choose a parameterized quantum circuit ansatz that generates a unitary $U(\boldsymbol{\alpha})$ and prepares the test state via $|\psi\rangle = U(\boldsymbol{\alpha})|0\rangle^{\otimes 2n}$. Each inner product $\langle \psi | \mathcal{O}(\rho) | \psi \rangle$ in Eq. (4) can be estimated via the
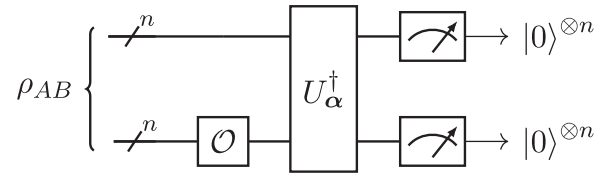
Fig. 1 Circuit estimating the overlap $\langle \psi | \mathcal{O}(\rho_{AB}) | \psi \rangle$ in Equation (7). This simplified quantum circuit estimates the overlap $\langle \psi | \mathcal{O}(\rho_{AB}) | \psi \rangle$ in Eq. (7) for a given implementable operation $\mathcal{O}$, where $|\psi\rangle := U_{\boldsymbol{\alpha}}|0\rangle^{\otimes 2n}$ is the parameterized input state.

canonical Swap Test subroutine[45], as both $U(\boldsymbol{\alpha})$ and $\mathcal{O}$ can be implemented in near-term devices. However, this subroutine costs a total number of $4n + 1$ qubits and requires a $4n$-qubit SWAP gate, which is resource consuming when n becomes large. Here we explore the special structure of the overlap $\langle \psi | \mathcal{O}(\rho) | \psi \rangle$ and propose an qubit efficient estimating procedure which uses $2n$ qubits and avoids the use of expensive SWAP gate. First of all, notice that

$$\langle \psi(\boldsymbol{\alpha}) | \mathcal{O}(\rho_{AB}) | \psi(\boldsymbol{\alpha}) \rangle \tag{5}$$

$$= \langle 0^{2n} | U_{\boldsymbol{\alpha}}^{\dagger} \mathcal{O}(\rho_{AB}) U_{\boldsymbol{\alpha}} | 0^{2n} \rangle \tag{6}$$

$$= \text{Tr} \left[U_{\boldsymbol{\alpha}}^{\dagger} \mathcal{O}(\rho_{AB}) U_{\boldsymbol{\alpha}} | 0^{2n} \rangle \langle 0^{2n} | \right], \tag{7}$$

where the second equality follows from the cyclic property of trace function. Since each $\mathcal{O}$ is implementable on near-term quantum devices, we may use ρ_{AB} as input to the quantum circuit implementing $\mathcal{O}$, and estimate the overlap $\langle \psi | \mathcal{O}(\rho_{AB}) | \psi \rangle$ using the quantum circuit illustrated in Fig. 1. The overlap is obtained by counting the relative frequency of the measurement outcome 0^{2n}. Then, we repeat the estimation procedure $|\Delta|$ times, where $|\Delta|$ is the size of Δ, to obtain the overlaps for different $\mathcal{O}$ in Eq. (4). With these data in hand, we compute the following loss function:

$$L(\boldsymbol{\alpha}) := \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \langle \psi(\boldsymbol{\alpha}) | \mathcal{O}(\rho_{AB}) | \psi(\boldsymbol{\alpha}) \rangle. \tag{8}$$

We remark that this loss function is a global cost since it requires measuring the expectation value of all zero results, which may lead to barren plateaus[46]. We provide detailed discussions and potential solutions in the section "Resource cost and barren plateaus". In particular, it would be interesting to adapt the technique invented in[46] to define a local version for pursuing better scalability and trainability. At last, we perform gradient-based optimization methods including SGD[47] and Adam[48] to minimize the loss function $L(\boldsymbol{\alpha})$ by varying the parameters $\boldsymbol{\alpha}$, whose value will determine the separability of the input state ρ_{AB}. More precisely, if $L(\boldsymbol{\alpha})$ is negative, we conclude that ρ_{AB} is entangled, since by the positive map criterion, separable states cannot yield a negative spectrum.

Taking into account the noise in NISQ quantum devices, we may introduce a tolerance threshold $\delta > 0$ so that $L(\boldsymbol{\alpha}) < -\delta$ implies the input state is entangled. This threshold δ can be set with prior knowledge about the noise characterization on the NISQ devices. What's more, for the purpose of entanglement detection, it is unnecessary to minimize $L(\boldsymbol{\alpha})$ since the condition $L(\boldsymbol{\alpha}) < 0$ is sufficient to assert that the input state is entangled. Based on this observation, we can terminate the optimization procedure that minimizes the loss function $L(\boldsymbol{\alpha})$ in advance to save the optimization cost. It was heuristically observed that the loss function achieves lower values with noise-free training than with noisy training[49–51], where the intuition behind is that the cost landscape is flattened and hence gradient magnitudes are reduced due to the hardware noise[52]. Based on this observation and the fact that separable states have positive eigenvalues in the positive map criteria, we conclude that the optimized loss function for separable states will always be positive, and thus our algorithm will not lead to false-positive results.
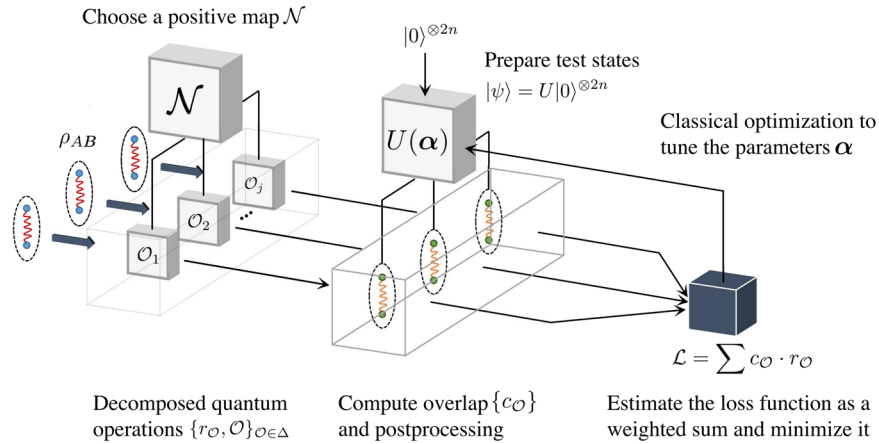
Fig. 2 The VED framework for detecting entanglement on near-term quantum devices. The difference between the deterministic VED in Algorithm 1 and the probabilistic VED in Algorithm 2 lies in how the decomposed quantum operations are sampled.

The detailed VED framework is summarized in Algorithm 1 and illustrated in Fig. 2. We name it the deterministic VED to distinguish it from the probabilistic framework described in the next section.

Algorithm 1. Deterministic VED
1: **Input:** $2n$-qubit quantum state ρ_{AB}, decomposition in Eq. (1) of the positive map $\mathcal{N}$, parameterized quantum circuit $U(\boldsymbol{\alpha})$ with initial parameters $\boldsymbol{\alpha}$, and tolerance δ;
2: Initialize $L(\boldsymbol{\alpha}) = 0$;
3: **for all** $\mathcal{O} \in \Delta$ such that $r_{\mathcal{O}} \neq 0$ **do**
4: Apply $U_{\boldsymbol{\alpha}}$ to $|0\rangle^{\otimes 2n}$ and obtain test state $|\psi\rangle = U_{\boldsymbol{\alpha}}|0\rangle^{\otimes 2n}$;
5: Input ρ_{AB} and compute the overlap $c_{\mathcal{O}} := \langle\psi|\mathcal{O}(\rho_{AB})|\psi\rangle$ using the quantum circuit in Fig. 1;
6: Update the loss function $L(\boldsymbol{\alpha}) = L(\boldsymbol{\alpha}) + r_{\mathcal{O}}c_{\mathcal{O}}$, where $r_{\mathcal{O}}$ is given by the decomposition in Eq. (1);
7: **end for**
8: Perform optimization methods to minimize $L(\boldsymbol{\alpha})$; terminate the optimization if the error tolerance is satisfied: $L(\boldsymbol{\alpha}) < -\delta$;
9: Output "Entangled" if the optimized $L(\boldsymbol{\alpha}) < -\delta$.

Probabilistic detection. In Algorithm 1, we have used a brute-force approach, where we iterate over the set of implementable operations Δ, to estimate the loss function $L(\boldsymbol{\alpha})$. Actually, $L(\boldsymbol{\alpha})$ can be estimated in a probabilistic way using the sampling technique, by virtue of the quasi-probability decomposition in Eq. (1). This method would be beneficial when the number of decomposed operations in Eq. (1) with non-zero coefficients is large while the sampling cost is relatively low. Now we describe the sampling method accurately. First of all, notice that the decomposition in Eq. (1) induces a quasi-probability distribution $\{r_{\mathcal{O}}\}_{\mathcal{O}\in\Delta}$ over Δ. From this quasi-probability distribution, we can construct a probability distribution $\{p_{\mathcal{O}}\}_{\mathcal{O}\in\Delta}$ using the canonical technique, i.e.,

$$p_{\mathcal{O}} := \frac{|r_{\mathcal{O}}|}{\gamma}, \quad \gamma := \sum_{\mathcal{O}\in\Delta} |r_{\mathcal{O}}|. \tag{9}$$

Substituting Eq. (9) into Eq. (8) yields

$$L(\boldsymbol{\alpha}) = \gamma \sum_{\mathcal{O}\in\Delta} \text{sgn}(r_{\mathcal{O}})p_{\mathcal{O}}\langle\psi(\boldsymbol{\alpha})|\mathcal{O}(\rho_{AB})|\psi(\boldsymbol{\alpha})\rangle \tag{10}$$

$$= \mathbb{E}_{\mathcal{O}}[\gamma\text{sgn}(r_{\mathcal{O}})\langle\psi(\boldsymbol{\alpha})|\mathcal{O}(\rho_{AB})|\psi(\boldsymbol{\alpha})\rangle], \tag{11}$$

where $\mathbb{E}(X)$ denotes the expectation of the random variable X, and the expectation in Eq. (11) is evaluated w.r.t. the probability distribution $\{p_{\mathcal{O}}\}_{\mathcal{O}\in\Delta}$. Based on Eq. (11), we propose Algorithm 2,

which can be viewed as a probabilistic version of Algorithm 1. In particular, Algorithm 2 replaces the brute-force approach (Steps 3–7) in Algorithm 1 with the sampling approach, yielding a probabilistic algorithm as summarized in Algorithm 2.

Let's analyze Algorithm 2 in depth. First, we remark that the obtained $L'(\boldsymbol{\alpha})$ in step 10 of Algorithm 2 is an unbiased estimator of true value $L(\boldsymbol{\alpha})$ due to Eq. (11). Second, since $|L^{(m)}| \leq \gamma$, we can apply the Hoeffding's inequality[53] to ensure that $M = 2\gamma^2 \log(2/\varepsilon)/\delta^2$ number of samples would estimate the true value $L(\boldsymbol{\alpha})$ within error δ with success probability no less than $1-\varepsilon$, i.e.,

$$p(|L'(\boldsymbol{\alpha}) - L(\boldsymbol{\alpha})| \leq \delta) \geq 1 - \varepsilon. \tag{12}$$

This confirms the validity of the sampling procedure (steps 4–9) of Algorithm 2. We call γ the sampling cost since it determines M, the number of samples required to achieve the desired precision. At last, we examine the success probability of the algorithm, given the success probability condition in Eq. (12) of the sampling procedure. Assume the optimization procedure repeats K times. The overall success probability of Algorithm 2 is no less than $1-K\varepsilon$, as a direct corollary of Eq. (12) and the union bound. That is to say, if Algorithm 2 outputs "Entangled", ρ_{AB} is entangled with probability larger than $1-K\varepsilon$.

To summarize, we have proposed two variational entanglement detection methods. Algorithm 1 is deterministic in the sense that whenever it outputs "Entangled", one can safely assert that ρ_{AB} is entangled. On the other hand, Algorithm 2 is probabilistic in the sense that even if it outputs "Entangled", one can only declare that ρ_{AB} is entangled with certain success probability. Nevertheless, when the number of decomposed operations in Eq. (1) with non-zero coefficients is large while the simulation cost γ is relatively low, the latter method may be beneficial. In this case, one can reduce the number of iterations via sampling and thus save computational resources. Algorithm 2 scarifies precision for efficiency in entanglement detection.

Algorithm 2. Probabilistic VED
1: Input: $2n$-qubit quantum state ρ_{AB}, decomposition in Eq. (1) of the positive map $\mathcal{N}$, parameterized quantum circuit $U(\boldsymbol{\alpha})$ with initial parameters $\boldsymbol{\alpha}$, error tolerance δ, and fail probability ε.
2: Initialize $L'(\boldsymbol{\alpha}) = 0$;
3: Compute γ defined in Eq. (9) and set $M = 2\gamma^2 \log(2/\varepsilon)/\delta^2$;
4: **for all** $m = 1, \cdots, M$ **do**
5: Apply $U_{\boldsymbol{\alpha}}$ to $|0\rangle^{\otimes 2n}$ and obtain test state $|\psi\rangle = U_{\boldsymbol{\alpha}}|0\rangle^{\otimes 2n}$;
6: Sample a quantum operation $\mathcal{O}^{(m)}$ from Δ according to the probability distribution $\{p_{\mathcal{O}}\}_{\mathcal{O}\in\Delta}$ in Eq. (9); Let $r^{(m)}$ be the coefficient of $\mathcal{O}^{(m)}$ in Eq. (1);

7: Input ρ_{AB} and compute the overlap $c^{(m)} := \langle\psi|\mathcal{O}^{(m)}(\rho)|\psi\rangle$
   using the quantum circuit in Fig. 1;
8: Compute $L^{(m)} = \gamma\mathrm{sgn}(r^{(m)})c^{(m)}$;
9: **end for**
10: Compute the loss function $L'(\boldsymbol{a}) = \frac{1}{M}\sum_{m=1}^{M} L^{(m)}$;
11: Perform optimization methods to minimize $L'(\boldsymbol{a})$; terminate
    the optimization if the error tolerance is satisfied: $L'(\boldsymbol{a}) < -\delta$;
12: Output "Entangled" if the optimized $L'(\boldsymbol{a}) < -\delta$.

Prominent positive maps

In section "Quantum entanglement detection" we have outlined the general deterministic and probabilistic VED frameworks for detecting entanglement via positive map criterion. In this section, we elaborate on three prominent positive maps—the transpose map[54], the reduction map[55], and the enhanced reduction map[56,57]—to illustrate how the deterministic VED framework works. We choose the set of NISQ implementable quantum operations Δ to be the set of Pauli channels induced by Pauli operators from the Pauli set $\boldsymbol{P}_n$, i.e.,

$$\Delta := \left\{\mathcal{P} \mid \mathcal{P}(\cdot) = P(\cdot)P^\dagger, P \in \boldsymbol{P}_n\right\}. \tag{13}$$

For each of the three positive maps under consideration, we firstly decompose it w.r.t. Δ as in Eq. (1) and then adopt the variational framework summarized in Algorithm 1 to fulfill entanglement detection. However, we remind that not all positive maps can be decomposed w.r.t. the set of Pauli channels.

Here are remarks for the three criteria under consideration. First, the reduction criterion is strictly weaker than both the transpose criterion and the enhanced reduction criterion, in the sense that the states that can be detected by the first criterion can also be detected by the latter two criteria. Second, there is no inclusion relation between the transpose criterion and the enhanced reduction criterion. That is, there are states that can be detected by one but not by the other. As so, given an unknown state, one may execute VED twice. One adopts the PPT criterion, and the other adopts the enhanced reduction criterion. The state is necessarily entangled if at least one of these two VEDs outputs "Entangled". We also show by example how VED works in qutrit systems in Supplementary Note 2, utilizing the Choi map[58,59].

PPT criterion. A necessary condition for entanglement detection is the positive partial transpose (PPT) criterion[54], which we briefly review as follows. Let ρ_{AB} be a bipartite quantum state. We can express it as

$$\rho_{AB} = \sum_{ijkl} a_{ijkl}|i\rangle\langle j|_A \otimes |k\rangle\langle l|_B, \tag{14}$$

where $\{|i\rangle\}_i$ and $\{|k\rangle\}_k$ are the computational bases of A and B, respectively. Its partial transpose with respect to system B is defined as

$$\rho_{AB}^{T_B} := (\mathrm{id}_A \otimes T_B)(\rho_{AB}) \tag{15}$$

$$= \sum_{ijkl} a_{ijkl}|i\rangle\langle j| \otimes (|k\rangle\langle l|)^\mathsf{T} \tag{16}$$

$$= \sum_{ijkl} a_{ijkl}|i\rangle\langle j| \otimes |l\rangle\langle k|, \tag{17}$$

where T_B denotes the transpose map on system B. The PPT criterion says that if ρ_{AB} is separable, then $\rho_{AB}^{T_B} \geq 0$. Conversely, the negative spectrum witnesses entanglement of ρ_{AB}. What's more, the PPT criterion is not only necessary but also sufficient for separability of the $2 \otimes 2$ and $2 \otimes 3$ cases[25,60,61].

We begin with the two-qubit bipartite quantum state case. Notice that the qubit transpose map admits the following

decomposition w.r.t. Δ specialized in Eq. (13):

$$T(\rho) = \frac{\rho + X\rho X - Y\rho Y + Z\rho Z}{2}, \tag{18}$$

where X, Y, Z are the Pauli matrices. The validity of this decomposition can be checked by direct calculation. Substituting Eq. (18) into Eq. (17), we obtain

$$\rho_{AB}^{T_B} := (\mathrm{id}_A \otimes T_B)(\rho_{AB}) \tag{19}$$

$$= \frac{1}{2}(\rho + X_B\rho X_B - Y_B\rho Y_B + Z_B\rho Z_B), \tag{20}$$

where the quantum operation $X_B\rho_{AB}X_B$ should be understood as $(I_A \otimes X_B)\rho_{AB}(I_A \otimes X_B)$, and similarly for $Y_B\rho Y_B$ and $Z_B\rho Z_B$. Adapting the decomposition in Eq. (20) into Algorithm 1, we successfully apply the proposed VED to accomplish the PPT criterion in the qubit case.

Now we show the above detection method can be generalized to the multi-qubit bipartite quantum state case. Let $\boldsymbol{B} \equiv B_1 B_2 \cdots B_n$ be a composite system with n qubits, i.e., B_i represents the ith qubit system. A key observation is that the transpose operation satisfies the tensor product property: transposing the composite system $\boldsymbol{B}$ is equivalent to transposing the local qubit systems B_i individually. More precisely,

$$T_{\boldsymbol{B}} = \bigotimes_{i=1}^{n} T_{B_i}, \tag{21}$$

where T_{B_i} is the transpose operation on the ith qubit. Equations (21) and (18) together give $T_{\boldsymbol{B}}$ as a linear combination of 4^n Pauli channels in total. Using this decomposition, we may apply VED (Algorithm 1 or Algorithm 2) to accomplish the multi-qubit PPT criterion deterministically or probabilistically.

Reduction criterion. In this section, we first review the reduction criterion[55] and then propose a variational algorithm implementing this criterion within the VED framework. Let

$$\mathcal{R}_{B\to B}(X_B) := \mathrm{Tr}[X_B]I_B - X_B, \tag{22}$$

which is known as the reduction map. The reduction criterion says that if a bipartite quantum state ρ_{AB} is separable, then it must hold that

$$\sigma_{AB} := (\mathrm{id}_A \otimes \mathcal{R}_{B\to B})(\rho_{AB}) \geq 0. \tag{23}$$

Equivalently, if σ_{AB} has negative eigenvalues, then ρ_{AB} is entangled. It is based on this observation that our variational algorithm works.

To apply the framework in the section "Quantum entanglement detection", we have to first decompose $\mathcal{R}_{B\to B}$ into a linear combination of Pauli channels. Indeed, we can do so since

$$\mathcal{R}_{B\to B}(\rho_B) := \mathrm{Tr}[\rho_B]I_B - \rho_B \tag{24}$$

$$= \frac{1}{2^n}\sum_{P\in\boldsymbol{P}_n}\mathcal{P}(\rho_B)I_B - \rho_B \tag{25}$$

$$= \frac{1-2^n}{2^n}\rho_B + \frac{1}{2^n}\sum_{P\neq I^{\otimes n}}\mathcal{P}(\rho_B), \tag{26}$$

where the second equality follows from the twirling property of Pauli channels[62], Exercise 4.7.3. Using this decomposition, we can call Algorithm 1 or Algorithm 2 to accomplish the reduction criterion. Specially, in the qubit case where $n = 1$, the reduction map is of the form

$$\mathcal{R}_{B\to B}(\rho) = \frac{-\rho + X\rho X + Y\rho Y + Z\rho Z}{2}. \tag{27}$$

As one might see, deterministic VED using the reduction criterion is not efficient in the many qubits case since it has to compute exponentially many overlaps: for a $2n$-qubit bipartite

quantum state, one has to compute about 4^n overlaps. Notice that the probabilistic VED using the reduction criterion is slightly better than the deterministic one since the sampling cost satisfies $\gamma \approx 2^n$. The quasi-probability sampling method does not show notable advantages over the deterministic method for the positive maps used in our work. However, the probabilistic VED provides a different perspective to reduce the cost using the quasi-probability sampling technique and it may be useful for entanglement detection via other positive maps. In the section "VED based on reduction criterion without decomposition", we also propose another entanglement detection method (cf. Algorithm 3) with better efficiency in measurement cost by exploring the simple structure of the reduction map in Eq. (22) and showcase the efficiency of Algorithm 3 compared to the regular VED using typical entangled states.

Enhanced reduction criterion. In this section, we consider an enhanced version of the reduction map[56,57] for bipartite quantum states. This enhanced criterion is based on an elementary positive map which operates on state spaces with even dimension. It is known that the enhanced reduction criterion detects many bound entangled states (states that satisfy the PPT criterion). As before, we first review this enhanced reduction criterion and show how to combine it with the VED framework proposed in the section "Quantum entanglement detection" to detect entanglement.

Define the following anti-symmetric unitary in an *n*-qubit Hilbert space:

$$U_a = \text{antidiag}(1, -1, 1, -1, \dots, 1, -1),\qquad(28)$$

where antidiag means anti-diagonal. For example, when $n = 2$, the corresponding anti-symmetric unitary has the form

$$U_a = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} = X \otimes iY.\qquad(29)$$

Indeed, one can check that the *n*-qubit U_a can be decomposed w. r.t. the Pauli set as $U_a = X \otimes \cdots \otimes X \otimes iY$, where there are $n-1$ X operators in the tensor product. Based on U_a, we define the following map[56]

$$\mathcal{K}_{B \to B}(\rho_B) := \mathcal{R}_{B \to B}(\rho_B) - U_a T_B(\rho) U_a^\dagger,\qquad(30)$$

where $\mathcal{R}_{B \to B}$ is the reduction map defined in Eq. (22) and T_B is the transpose map defined in Eq. (17). This map has been shown to be positive but not completely positive[56]. What's more, this map improves the reduction criterion and can detect bound entangled states that cannot be detected by the PPT criterion. Substituting the Pauli decomposition in Eq. (26) of $\mathcal{R}$ and the Pauli decomposition in Eq. (21) of T_B into Eq. (30) and regrouping the Pauli terms, we obtain a Pauli decomposition of $\mathcal{K}$, where there is a total number of 4^n Pauli terms. Using this decomposition, we can use VED (Algorithm 1 or Algorithm 2) to accomplish the enhanced reduction criterion.

Resource cost and barren plateaus
Note that another way to detect and quantify entanglement of a state ρ_{AB} is to obtain its density matrix via quantum state tomography[63]. Full density matrix reconstruction of an unknown $(n_A + n_B)$-qubit state in the worst-case costs exponential copies of the state[64,65], e.g., $\widetilde{\Omega}(4^{n_A+n_B})$ measurement results are necessary to reconstruct a matrix close to ρ in terms of trace distance[65]. Using the learned density matrix, we can either numerically apply a positive map on it or compute the fidelity between ρ_{AB} and any entangled target states. However, such methods are resource-demanding compared to the VED framework.

To use the decomposed maps for entanglement detection on the state ρ_{AB}, we can apply them either to subsystem A or to subsystem B, which requires $\text{poly}(D)4^{\min\{n_A,n_B\}}$ measurement results with circuit depth D in optimization loops. Under the assumption that the PQC could be well trained, our method is considerably better than the state tomography method. The barren plateaus phenomena might weaken the advantage of our method over the state tomography in terms of the measurement cost. However, for large-scale quantum systems, our method could work in the proof of concept while computing the minimum eigenvalue in the tomography method is extremely difficult. We further elaborate the barren plateaus problem from different aspects. In particular, methods based on the state tomography need vast memory to store and process the density matrix on a classical computer, which are unbearably resource-demanding as the scale of the number of qubits increases. The VED framework, on the other hand, does not require such classical memory and post-processing.

As our current methods use a global cost function and choose a hardware-efficient ansatz, it is possible that VED exhibits a barren plateau, resulting an exponentially suppressed gradient with respect to the problem dimension[66]. Under this circumstance, there exists a vast flat area on the loss/optimization landscape. This phenomenon is known as the barren plateau (BP) and is independent of the optimizer utilized, meaning that a gradient-free optimizer would not help in mitigating this phenomenon[67]. Furthermore, noise and entanglement could also induce BP[52,68]. In order to mitigate BP, one can adopt the following strategies: variable structure ansatzes[69–71], layerwise learning[72], meta-learning[73], and parameter initialization and parameter correlation strategies[74,75]. Enormous evidences show that a local cost function could help mitigate BP by extending the trainable circuit depth to a shallow level $\mathcal{O}(\log(n))$[46]. What's more, we can suppress the hardware noise using various error mitigation techniques (see, e.g., refs. [41,42,76–79]), further improving the optimization results. By exploiting the above methods and continuous progresses in the areas of barren plateau and error mitigation, we believe that the effect and practicability of our VED framework could be further improved.

VED based on reduction criterion without decomposition
In the section "Prominent positive maps" we have shown how VED uses the reduction criterion to detect entanglement; it works by decomposing the reduction map $\mathcal{R}$ into a linear combination of Pauli channels and then variationally estimate the minimal eigenvalue of the averaged output state.

Here we propose another variational entanglement detection algorithm for the reduction criterion, motivated by the simple structure of the reduction map. The intuition behind this protocol is as follows. We know that ρ_{AB} is entangled if $\mathcal{R}_{B \to B}(\rho_{AB})$ is not semidefinite positive. Using the variational characterization of a Hermitian operator's minimum eigenvalue [see Supplementary Note 1], this means that

$$\min_{|\psi_{AB}\rangle} \langle \psi | \mathcal{R}_{B \to B}(\rho_{AB}) | \psi \rangle\qquad(31)$$

$$= \min_{|\psi_{AB}\rangle} \langle \psi | (I_A \otimes \rho_B - \rho_{AB}) | \psi \rangle\qquad(32)$$

$$= \min_{|\psi_{AB}\rangle} \{ \text{Tr}[\psi_B \rho_B] - \text{Tr}[\psi_{AB}\rho_{AB}] \} < 0,\qquad(33)$$

where the minimization ranges over all pure bipartite quantum states $|\psi_{AB}\rangle$ in system AB, $\psi_{AB} \equiv |\psi\rangle\langle\psi|_{AB}$ and $\psi_B := \text{Tr}_A \psi_{AB}$. From Equation (33), one can see that it suffices to compute the difference of two overlaps and then variationally estimate the minimal eigenvalue. The crucial point is that the number of overlaps is independent on the dimension of the *n*-qubit system B. This detection method could save a large amount of computing

resources when n becomes large. The improved VED based on the reduction criterion is summarized in Algorithm 3.

### Algorithm 3. Improved VED based on reduction criterion

1: Input: $2n$-qubit quantum state ρ_{AB}, parameterized quantum circuit $U(\boldsymbol{a})$ with initial parameters $\boldsymbol{a}$, and tolerance δ;
2: Apply $U(\boldsymbol{a})$ to $|00\rangle_{AB}$ on system AB and obtain the test state $|\psi\rangle_{AB} = U(\boldsymbol{a})|00\rangle_{AB}$;
3: Compute the overlap between state ψ_B and ρ_B on subsystem B using the Swap Test and obtain $c_1 = \mathrm{Tr}[\psi_B \rho_B]$;
4: Apply $U(\boldsymbol{a})$ to $|00\rangle_{AB}$ on system AB and obtain the test state $|\psi\rangle_{AB} = U(\boldsymbol{a})|00\rangle_{AB}$;
5: Compute the overlap between state ψ_{AB} and ρ_{AB} using the Swap Test and obtain $c_2 = \mathrm{Tr}[\psi_{AB} \rho_{AB}]$;
6: Compute the loss function $L(\boldsymbol{a}) = c_1 - c_2$;
7: Perform optimization methods to minimize $L(\boldsymbol{a})$; terminate the optimization if the error tolerance is satisfied: $L(\boldsymbol{a}) < -\delta$;
8: Output "Entangled" if the optimized $L(\boldsymbol{a}) < -\delta$.

To showcase the advantage of the improved VED, we use the reduction map's simple structure exploited by this algorithm to estimate $\langle 0|\mathcal{R}_{B \to B}(\rho_{AB})|0\rangle$ for $\rho_{AB} = |W\rangle\langle W|$ being the four-qubit generalized W state, where

$$|W\rangle = \frac{1}{2}(|1000\rangle + |0100\rangle + |0010\rangle + |0001\rangle), \qquad (34)$$

and both local systems have two qubits. The simulation results are summarized as box plots and compared to the estimation obtained by the Pauli channel decomposition of the reduction map in Fig. 3. When taking the same number of measurement shots, the method used by Algorithm 3 gives estimation results more concentrated in a range close to the ideal value than the Pauli channel decomposition method adopted by the standard VED. Thus, the improved VED can achieve the desired accuracy using fewer measurement shots, which means that fewer copies of the input quantum state are required.

We remark that this idea can also be adopted to improve the efficiency of VED using the enhanced reduction criterion.
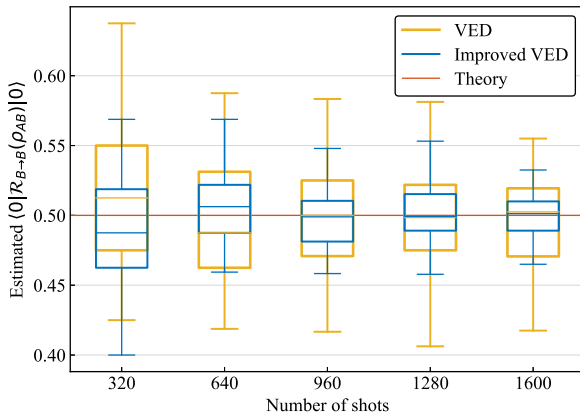
Fig. 3 Comparison between VED and improved VED on estimate accuracy. The overlap $\langle 0|\mathcal{R}_{B \to B}(\rho_{AB})|0\rangle$ is estimated using the reduction map's Pauli decomposition (Equation (26) used in VED) and its simple structure (Eq. (33) in the improved VED), where ρ_{AB} is the four-qubit generalized W state. The methods are simulated under five different numbers of measurement shots, and for each given number of shots, they are, respectively, repeated 50 times. The distributions of both methods' estimations are summarized in box plots, where each box extends from the first quartile to the third quartile of the corresponding data, with a line at the median. The whiskers extend to the most extreme data points within 1.5 times the interquartile range. The red horizontal line gives the theoretical value for reference. The improved VED (cf. Algorithm 3) achieves a better performance.

Quantum entanglement quantification

One of the most well-known entanglement measure is the logarithmic negativity[80,81], which has various applications in quantum information theory. For a bipartite state ρ_{AB}, its logarithmic negativity is defined as

$$E_N(\rho_{AB}) := \log \|\rho_{AB}^{T_B}\|_1. \qquad (35)$$

Based on the recently developed near-term quantum algorithm for trace distance estimation[6] and the fact that E_N is defined via the transpose map T_B, we introduce a variational quantum algorithm to estimate E_N using an ancillary qubit system R. According to [6,Corollary 3], it holds that

$$\|\rho_{AB}^{T_B}\|_1 = 2 \max_U \mathrm{Tr}|0\rangle\langle 0|_R Q_R - \mathrm{Tr}\rho_{AB}^{T_B} \qquad (36)$$

$$= 2 \max_U \mathrm{Tr}|0\rangle\langle 0|_R Q_R - 1, \qquad (37)$$

where $Q_R = \mathrm{Tr}_{AB} Q_{ABR}$, $Q_{ABR} = U(\rho_{AB}^{T_B} \otimes |0\rangle\langle 0|_R)U^\dagger$, and the maximization ranges over all unitaries on the composite system ABR. Note that the second equality follows from the fact that T_B is trace-preserving. Following the idea of VED, we may decompose the transpose map T_B appeared in the operator Q_{ABR} (correspondingly, Q_R) into a linear combination of Pauli terms via Eqs. (20) and (21), compute the overlaps in Eq. (37) one by one, and then variationally estimate the maximal value. For illustrative purposes, we give Algorithm 4, the variational logarithmic negativity estimation (VLNE), as an example of estimating the logarithmic negativity of a two-qubit quantum state ρ_{AB}. However, we emphasize that method outlined in Algorithm 4 can be easily generalized to quantify multi-qubit bipartite entanglement, as the transpose operation satisfies the preferable tensor product property in Eq. (21). What's more, Algorithm 4 can be modified to use the sampling technique to estimate the average state, following the idea illustrated in Algorithm 2.

Algorithm 4. Variational logarithmic negativity estimation

1: Input: a 2-qubit quantum state ρ_{AB} and parameterized circuit $U_{ABR}(\boldsymbol{a})$ with initial parameters $\boldsymbol{a}$;
2: Apply $U_{ABR}(\boldsymbol{a})$, respectively, to

$$\rho_{AB} \otimes |0\rangle\langle 0|_R, \qquad (38)$$

$$(I_A \otimes X_B)\rho_{AB}(I_A \otimes X_B) \otimes |0\rangle\langle 0|_R, \qquad (39)$$

$$(I_A \otimes Y_B)\rho_{AB}(I_A \otimes Y_B) \otimes |0\rangle\langle 0|_R, \qquad (40)$$

$$(I_A \otimes Z_B)\rho_{AB}(I_A \otimes Z_B) \otimes |0\rangle\langle 0|_R, \qquad (41)$$

and obtain the states $\sigma^{(0)}$, $\sigma^{(1)}$, $\sigma^{(2)}$, $\sigma^{(3)}$, respectively.
3: Obtain $o_j = \mathrm{Tr}[\sigma_R^{(j)}|0\rangle\langle 0|_R]$ for $j = 0, 1, 2, 3$ by measurements on system R.
4: Compute the loss function $\mathcal{L}_1 := -(o_0 + o_1 - o_2 + o_3)/2$.
5: Perform optimization methods to minimize $\mathcal{L}_1(\boldsymbol{a})$;
6: Compute $\beta = 2|\mathcal{L}_1| - 1$ as the estimated trace norm of $\rho_{AB}^{T_B}$;
7: Output $\log \beta$ as the estimated logarithmic negativity.

One may also evaluate the entanglement measures[82-84] based on the sandwiched Rényi relative entropy[85,86] of order 1/2, making use of the recently proposed variational quantum algorithm estimating the fidelity between two quantum states[6].

Experiments on IBMQ

In this section, we discuss how to apply the VED framework to detect the two-qubit maximally entangled state $|\Phi\rangle := (|00\rangle + |11\rangle)/\sqrt{2}$ on IBM-Q superconducting quantum hardware accessible to the public. The specific quantum device used is ibmq-santiago (5 qubits) with a quantum volume of 32. The positive map adopted here for detection purpose is the qubit reduction map $\mathcal{R}_{B \to B}$ defined in Eq. (22). After implementing the
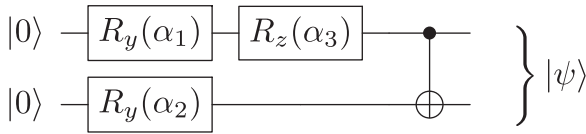
Fig. 4 Circuit for experiments on IBMQ. This parameterized two-qubit quantum circuit $U(\boldsymbol{\alpha})$ is used for preparing the test state $\psi_{AB}(\boldsymbol{\alpha})$ on the ibmq-santiago hardware. The parameters $\boldsymbol{\alpha}$ are randomly initialized as $(\alpha_1, \alpha_2, \alpha_3) = (3.2292, 4.8579, 5.4691)$.

decomposed reduction map by 4 Pauli terms as Eq. (27), we use a parameterized quantum circuit $U(\boldsymbol{\alpha})$ to prepare four identical test states $\psi_{AB}(\boldsymbol{\alpha}) = U(\boldsymbol{\alpha})|00\rangle\langle00|_{AB}U^\dagger(\boldsymbol{\alpha})$ and compute the loss function defined in Eq. (8). The PQC used is depicted in Fig. 4 with three randomly initialized parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$. During the optimization procedure, we apply the gradient descent algorithm to guide the learning process where the analytical gradient is calculated via the following parameter-shift rule[87]:

$$\frac{\partial \mathcal{L}(\boldsymbol{\alpha})}{\partial a_j} := \frac{1}{2}\left[\mathcal{L}\left(a_j + \frac{\pi}{2}\right) - \mathcal{L}\left(a_j - \frac{\pi}{2}\right)\right]. \quad (42)$$

Due to the finite sampling restriction for measurements, the optimization procedure essentially falls into the regime of Stochastic Gradient Descent (SGD)[47]. The optimized loss values converges to $\mathcal{L}_{min} \approx -0.43$. The gap between the experiment data and simulation result $\lambda_{min} = -0.5$ is due to various hardware noises on the ibmq-santiago processor. One can further adopt error mitigation methods[10] to improve the result. This result proves the validity of our VED framework. Note that if we adopt the termination setup in Algorithm 1, it will require much fewer optimization iterations (4–5 rounds are sufficient) to obtain the detection result. As mentioned in ref. [88], the communication bottleneck between the IBM-Q hardware and classical optimizer blocks us from efficiently conducting experiments without any specified reservation. This leads to a 9-minutes waiting time on average for each circuit evaluation from the IBM-Q cloud service. We also implement the probabilistic detection Algorithm 2 for comparison. In the experiment, we set the backend to the Aer simulator which imitates the behavior of ibmq-santiago and choose $\varepsilon = 0.01$ and $\delta = 0.1$. The experimental results reveal that the probabilistic VED achieves the same detection precision level compared to the deterministic VED. As comparison, we conduct numerical simulations on the Baidu Quantum Leaf platform[89] and obtain similar results. We summarize the experimental and numerical results in Fig. 5.

Numerical simulations for entanglement detection
In this section, we carry out numerical simulations that apply VED to detect a variety of bipartite quantum states of interest to investigate the performance of VED and its motivated entanglement quantification algorithm. All simulations, including optimization loops, are conducted using the Paddle Quantum[90] toolkit on the PaddlePaddle Deep Learning Platform[91].
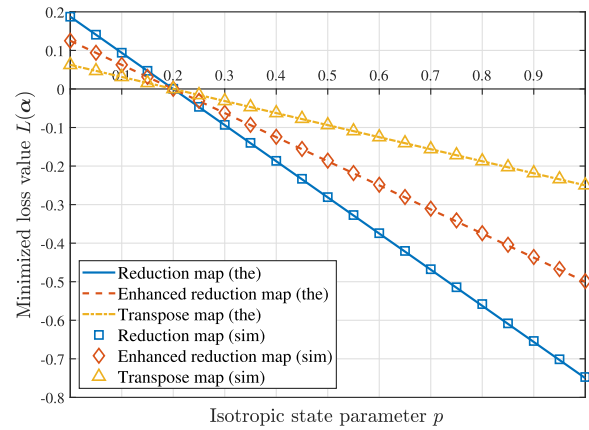
Isotropic states. The n-qubit isotropic state family (n is even and each local system has $n/2$-qubits) is defined as[92,Eq. (32)]

$$\rho_{AB}^{iso}(p) := p\Phi_{AB} + (1-p)\frac{I_{AB}}{2^n}, \quad (43)$$

where $p \in [0, 1]$ is a parameter, Φ_{AB} is the n-qubit maximally entangled state, and I_{AB} is the identity operator in $\boldsymbol{AB}$ in which $\boldsymbol{A} \equiv A_1 \cdots A_{n/2}$ and $\boldsymbol{B} \equiv B_1 \cdots B_{n/2}$. The qubit systems $\{A_i\}_i$ are at Alice's hand, while the qubit systems $\{B_i\}_i$ are at Bob's hand. Intuitively, the isotropic state is a convex combination of the maximally entangled state Φ_{AB} and the maximally mixed state $I_{AB}/2^n$. It has been shown that $\rho_{AB}^{iso}(p)$ is separable (w.r.t. the $\boldsymbol{A}:\boldsymbol{B}$ cut) if and only if $p \le 1/(2^{n/2} + 1)$[92].
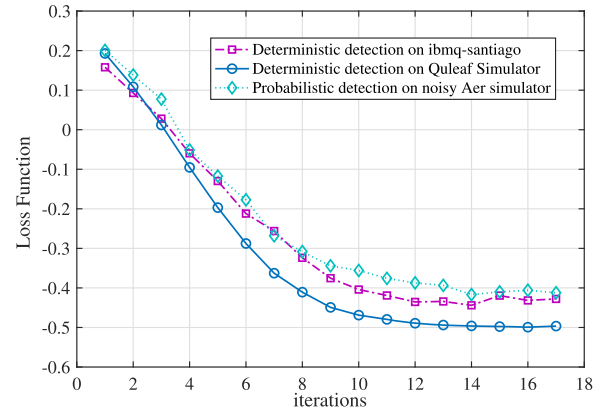
Fig. 5 Estimated minimum eigenvalue λ_{min} by VED using the reduction criterion on the Bell state $|\Phi\rangle$. The red square curve records the results of Algorithm 1 on ibmq-santiago with shots = 8192 for each circuit evaluation. The cyan diamond curve records the results of Algorithm 2 by choosing $\delta = 0.1$ and $\varepsilon = 0.01$. The blue circle curve records the numerical results on the Baidu Quantum Leaf platform[89]. Learning rate in the gradient descent algorithm is set to be LR = 0.5.

Fig. 6 Numerical results on the four-qubit isotropic states defined in Eq. (43). Each line depicts the smallest eigenvalue of every isotropic state with parameter $p \in [0, 1]$ under the corresponding map. This line of the smallest eigenvalues is a lower bound of the loss function $L(\boldsymbol{\alpha})$. Each marker depicts the minimized loss value obtained by simulations (sim) of Algorithm 1 on a chosen isotropic state, aligning with the theoretical (the) line.

We numerically carry out Algorithm 1 together with the three prominent positive maps—the PPT criterion, the reduction criterion, and the enhanced reduction criterion—introduced in the section "Prominent positive maps", using four-qubit isotropic states as inputs. The minimized loss values of these three maps obtained by our simulations on the isotropic states are represented by different markers in Fig. 6. As can be seen from this figure, VED can successfully identify the range of p for which the corresponding isotropic state can be detected by each positive map. The markers representing results from simulations fall on the lines that give the minimums of the loss function $L(\boldsymbol{\alpha})$, verifying the validity and viability of our VED framework. Note that for detecting entanglement in four-qubit isotropic states, all three maps are both necessary and sufficient. However, this phenomenon is not universal for all four-qubit states, as we shall see in the experiment using Breuer states.

Table 1. Estimated λ_{min} by VED using the reduction criterion on isotropic states with up to 10 qubits and the noise parameter $p = 0.7$.

	Number of qubits				
	$n = 2$	$n = 4$	$n = 6$	$n = 8$	$n = 10$
VED	−0.2750	−0.4687	−0.5793	−0.6378	−0.6663
Theory	−0.2750	−0.4688	−0.5797	−0.6387	−0.6690

Each value shown in the VED row is the average of optimized losses from 10 independent runs of 100 training iterations. When estimating each overlap, we use a circuit as shown in Fig. 9 of depth $n/2$, with randomly initialized parameters and number of shots 8192. The Theory row gives the theoretical λ_{min} for comparison.

To explore the barren plateau phenomenon that might be possible in our proposed VED algorithms, we carry out extensive numerical simulations on isotropic states by ranging the total number of qubits from 2 to 10 and fixing the noise parameter p to 0.7. In the following Table 1, we compare the average loss achieved by VED using the reduction criterion to the theoretical minimum, assuming that the number of shots does not scale with the number of qubits when estimating each overlap induced by the decomposition. As the number of qubits increases, there exhibits no large discrepancy between the estimated value by VED and the theoretical value. The scaling behavior of the optimized loss suggests that the VED algorithm is resilient to the barren plateau phenomenon when detecting bipartite states with a moderate number of qubits.

Breuer states. As we have mentioned in the section "Prominent positive maps", there are states that can be detected by the enhanced reduction criterion yet cannot be detected by the PPT criterion. In this section, we use the proposed VED framework to numerically consolidate this statement. The four-qubit Breuer state family is defined as[56], Eq. (7).

$$\rho_{AB}^{Breuer}(\lambda) := \begin{pmatrix} \frac{1-\lambda}{3} & 0 & 0 & 0 \\ 0 & \frac{1+2\lambda}{6} & \frac{1-4\lambda}{6} & 0 \\ 0 & \frac{1-4\lambda}{6} & \frac{1+2\lambda}{6} & 0 \\ 0 & 0 & 0 & \frac{1-\lambda}{3} \end{pmatrix}, \quad (44)$$

where $\lambda \in [0, 1]$ is a parameter, $A \equiv A_1A_2$, and $B \equiv B_1B_2$. The qubit systems A_1 and A_2 are at Alice's hand while the qubit systems B_1 and B_2 are at Bob's hand. It has been shown that ρ^{Breuer} is separable (w.r.t. the $A:B$ cut) if and only if $\lambda = 0$ and can be detected by the enhanced reduction criterion[56]. On the other hand, it has positive partial transpose if and only if $\lambda \leq 1/6$[56], witnessing the power of the enhanced reduction criterion.

Following the same line of the case of the isotropic state, we carry out Algorithm 1 on the three criteria using four-qubit Breuer states as inputs. The minimized loss values obtained by our simulations on selected Breuer states are represented in Fig. 7 by markers, which again align with the theoretical lines. From the numeric results, we can see that while the enhanced reduction criterion is still necessary and sufficient for entanglement detection in the four-qubit Breuer states, neither the reduction criterion nor the PPT criterion can detect all entangled states in the Breuer state family, attesting the advantage of the enhanced reduction criterion in this case.

Numerical simulations for logarithmic negativity estimation
For simulations of variational entanglement quantification with logarithmic negativity, we adopt the hardware efficient ansatz used for trace distance estimation in ref. [6] where the circuit depth
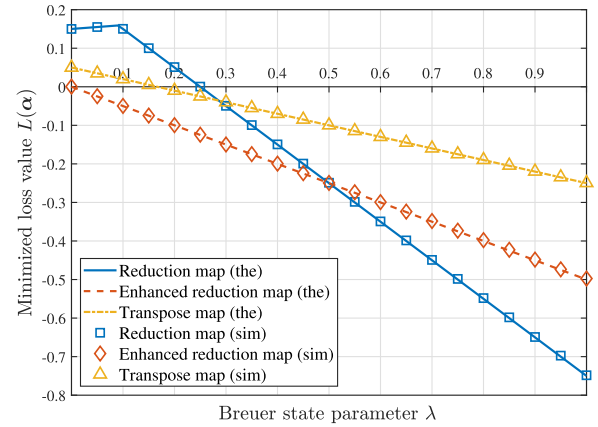
Fig. 7 Numerical results on the four-qubit Breuer states defined in Eq. (44). Each line depicts the smallest eigenvalue of every Breuer state with parameter $p \in [0, 1]$ under the corresponding map. This line of the smallest eigenvalues is a lower bound of the loss function $L(\alpha)$. Each marker depicts the minimized loss value obtained by simulations (sim) of Algorithm 1 on a chosen Breuer state, aligning with the theoretical (the) line.
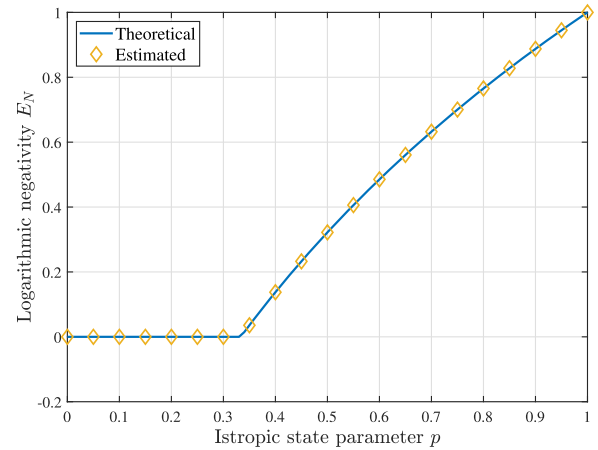
Fig. 8 Numerical results on the two-qubit isotropic states. The blue line represents the precisely calculated logarithmic negativity of isotropic states with parameter $p \in [0, 1]$. The yellow markers depict the estimated logarithmic negativities by simulations of Algorithm 4 on selected isotropic states.

is 4. The simulations are carried out on two-qubit isotropic states, which is defined as

$$\rho_{AB}^{iso}(p) := p\Phi_{AB} + (1 - p)\frac{I_{AB}}{4}, \quad (45)$$

where Φ_{AB} is the two-qubit maximally entangled state. As shown in Fig. 8, the logarithmic negativity of a two-qubit isotropic state is positive if and only if its parameter $p > 1/3$, which matches the range of p where the corresponding isotropic states are entangled. The estimated logarithmic negativities by our method, which are represented by markers in Fig. 8, agree with the precisely calculated values given by the blue line.

DISCUSSION
In this paper, we combined two techniques that find crucial applications in the NISQ quantum devices, the variational quantum algorithms and the quasi-probability decomposition method, to propose the variational entanglement detection (VED) and variational logarithmic negativity estimation (VLNE) frameworks, contributing feasible solutions to detect and quantify

entanglement on near-term devices. VED is built upon the positive map criterion and works as follows. Firstly, it decomposes a chosen positive map into a linear combination of NISQ implementable quantum operations. Then, it variationally estimates the minimal eigenvalue of the output state of some positive map acting on the target bipartite state. Two methods are proposed to generate the output state: the first one averaged the output states according to the quasi-probability distribution; the second one estimated the average via the sampling technique. At last, it asserts that the target state is entangled if the optimized minimal eigenvalue is negative, guaranteed by the positive map criterion. We elaborated three well-known positive maps to illustrate how the VED framework is applied. Following the idea of VED, VLNE variationally computes the log-negativity entanglement measure, relying on a linear decomposition of the transpose map into Pauli terms and the recently proposed trace distance estimation algorithm. Experimental and numerical results on various bipartite states of interest have validated the proposed entanglement detection and quantification methods.

We expect that the VED framework can be upgraded to detect more entangled states. A crucial step towards this aim is to explore what kind of positive maps can be decomposed into a linear combination of Pauli channels. In the section "Quantum entanglement quantification" we showed by case how to variationally compute the log-negativity entanglement measure. It would be meaningful to design quantum algorithms to estimate other distance-based entanglement measures (see, e.g., refs. [93–95]).

METHODS

Entanglement detection via positive maps

Let ρ_{AB} be a bipartite quantum state in the composite system AB. By definition ρ_{AB} is separable if it can be decomposed into a convex combination of tensor products of states describing local systems as[96]

$$\rho_{AB} = \sum_x p_x |\psi_x\rangle \langle \psi_x|_A \otimes |\phi_x\rangle \langle \phi_x|_B, \quad (46)$$

where $p_x \geq 0$, $\sum_x p_x = 1$, and $\{|\psi_x\rangle\}_x$ and $\{|\phi_x\rangle\}_x$ are two sets of pure states in systems A and B, respectively. Otherwise, ρ_{AB} is entangled. Given the definition, it is natural to ask whether a given unknown bipartite quantum state is separable or entangled, known as the separability problem. This problem has been shown to be NP-hard[97,98]. There are many separability criteria that have been proposed to determine the separability or entanglement of bipartite quantum states as necessary conditions[12,23].

One of the most celebrated criteria for distinguishing separable states from entangled states is the positive map criterion. The core of the positive map criterion is that one subjects a subsystem of ρ_{AB} to a positive (but not completely positive) map $\mathcal{N}_{B \to B}$ that preserves the positivity of inputs. If ρ_{AB} is a product state, i.e., it is of the form $\rho_A \otimes \rho_B$, the resulting operator $\rho_A \otimes \mathcal{N}(\rho_B)$ is still positive. Consequently, due to the linearity, an arbitrary separable state is mapped into some positive operator by this map. On the other hand, if ρ_{AB} is entangled, the output operator $\mathcal{N}_{B \to B}(\rho_{AB})$ may be no longer positive; the transpose map is a prominent example[54]. That is to say, the negative spectrum of the output operator indicates entanglement of the input state. Mathematically, the positive map criterion states that a bipartite quantum state ρ_{AB} is separable if and only if for arbitrary system C and arbitrary positive (but not completely positive) map $\mathcal{N}_{B \to C}$, it holds that $\mathcal{N}_{B \to C}(\rho_{AB}) \geq 0$[25].

Despite its proven efficiency in entanglement detection, the positive map criterion is not directly applicable in practice, especially on recent NISQ devices. This is an immediate consequence of the fact that generically positive but not completely positive maps do not represent physically implementable quantum operations[99] and thus cannot be realized in near-term quantum devices. In this work, we showed how to overcome this obstacle and employ the positive map criterion to detect entanglement on NISQ devices.

Ansatz design

For the entanglement detection purpose, we adopt the circuit ansatz shown in Fig. 9 to prepare the test state $|\psi\rangle_{AB}$. It consists of parameterized single-qubit gates $U_3(\theta, \phi, \varphi) = R_z(\phi) R_y(\theta) R_z(\varphi)$ and circular layers of CNOT gates. Note that this ansatz can be easily generalized to multi-qubit case.
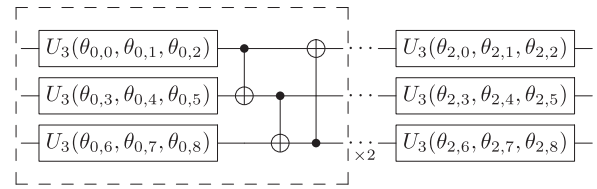
Fig. 9 Three qubit parameterized ansatz $U(\alpha)$ used for VED. The quantum circuit within the dotted block is repeated twice.

DATA AVAILABILITY

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

CODE AVAILABILITY

Code that supports the findings of this study is available from the corresponding authors upon reasonable request.

REFERENCES

1. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
2. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 023023 (2016).
3. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
4. Nakanishi, K. M., Mitarai, K. & Fujii, K. Subspace-search variational quantum eigensolver for excited states. *Phys. Rev. Res.* **1**, 033062 (2019).
5. Cerezo, M., Poremba, A., Cincio, L. & Coles, P. J. Variational quantum fidelity estimation. *Quantum* **4**, 248 (2020).
6. Chen, R., Song, Z., Zhao, X. & Wang, X. Variational quantum algorithms for trace distance and fidelity estimation. *Quantum Sci. Technol.* **7**, 015019 (2021).
7. Romero, J., Olson, J. P. & Aspuru-Guzik, A. Quantum autoencoders for efficient compression of quantum data. *Quantum Sci. Technol.* **2**, 045001 (2017).
8. Wan, K. H., Dahlsten, O., Kristjánsson, H., Gardner, R. & Kim, M. S. Quantum generalisation of feedforward neural networks. *npj Quantum Inf.* **3**, 36 (2017).
9. Cao, C. & Wang, X. Noise-assisted quantum autoencoder. *Phys. Rev. Appl.* **15**, 054012 (2021).
10. Endo, S., Cai, Z., Benjamin, S. C. & Yuan, X. Hybrid quantum-classical algorithms and quantum error mitigation. *J. Phys. Soc. Jpn.* **90**, 032001 (2021).
11. Cerezo, M. et al. Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 1–29 (2020).
12. Horodecki, R., Horodecki, P., Horodecki, M. & Horodecki, K. Quantum entanglement. *Rev. Mod. Phys.* **81**, 865 (2009).
13. Bennett, C. H. & Wiesner, S. J. Communication via one-and two-particle operators on Einstein–Podolsky–Rosen states. *Phys. Rev. Lett.* **69**, 2881 (1992).
14. Bennett, C. H. et al. Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Phys. Rev. Lett.* **70**, 1895 (1993).
15. Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed experiment to test local hidden variable theories. *Phys. Rev. Lett.* **24**, 549 (1970).
16. Ekert, A. K. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.* **67**, 661 (1991).
17. Shor, P. W. Algorithms for quantum computation: discrete logarithms and factoring. In *Proc. 35th Annual Symposium on Foundations of Computer Science*, 124–134 (IEEE, 1994).
18. Jozsa, R. Entanglement and quantum computation. Preprint at https://arxiv.org/abs/quant-ph/9707034 (1997).
19. Bennett, C. H., Shor, P. W., Smolin, J. A. & Thapliyal, A. V. Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem. *IEEE Trans. Inf. Theory* **48**, 2637–2655 (2002).
20. Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Phys. Rev. Lett.* **86**, 5188 (2001).
21. Raussendorf, R., Browne, D. E. & Briegel, H. J. Measurement-based quantum computation on cluster states. *Phys. Rev. A* **68**, 022312 (2003).
22. Zhao, X., Zhao, B., Wang, Z., Song, Z. & Wang, X. LOCCNet: a machine learning framework for distributed quantum information processing. *npj Quantum Inf.* **7**, 159 (2021).

23. Gühne, O. & Tóth, G. Entanglement detection. *Phys. Rep.* **474**, 1–75 (2009).
24. Friis, N., Vitagliano, G., Malik, M. & Huber, M. Entanglement certification from theory to experiment. *Nat. Rev. Phys.* **1**, 72–87 (2019).
25. Horodecki, M., Horodecki, P. & Horodecki, R. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A* **223**, 1–8 (1996).
26. Gühne, O. & Lütkenhaus, N. Nonlinear entanglement witnesses. *Phys. Rev. Lett.* **96**, 170502 (2006).
27. Bell, J. S. & Bell, J. S. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy* (Cambridge University Press, 2004).
28. Chen, K. & Wu, L. -A. A matrix realignment method for recognizing entanglement. *Quantum Inf. Comput.* **3**, 193–202 (2003).
29. Rudolph, O. Further results on the cross norm criterion for separability. *Quantum Inf. Process.* **4**, 219–239 (2005).
30. Horodecki, P. Separability criterion and inseparable mixed states with positive partial transposition. *Phys. Lett. A* **232**, 333–339 (1997).
31. Nielsen, M. A. & Kempe, J. Separable states are more disordered globally than locally. *Phys. Rev. Lett.* **86**, 5184 (2001).
32. D'Ariano, G. M., Paris, M. G. & Sacchi, M. F. Quantum tomography. *Adv. Imaging Electron Phys.* **128**, 206–309 (2003).
33. Steffen, M. et al. Measurement of the entanglement of two superconducting qubits via state tomography. *Science* **313**, 1423–1425 (2006).
34. Audenaert, K. M. R. & Plenio, M. B. When are correlations quantum?—Verification and quantification of entanglement by simple measurements. *New J. Phys.* **8**, 266–266 (2006).
35. Zhou, Y., Zeng, P. & Liu, Z. Single-copies estimation of entanglement negativity. *Phys. Rev. Lett.* **125**, 1–39 (2020). 2004.11360.
36. Elben, A. et al. Mixed-state entanglement from local randomized measurements. *Phys. Rev. Lett.* **125**, 200501 (2020).
37. Ketterer, A., Imai, S., Wyderka, N. & Gühne, O. Statistically significant tests of multiparticle quantum correlations based on randomized measurements. Preprint at https://arxiv.org/abs/2012.12176 (2020).
38. Buscemi, F., Dall'Arno, M., Ozawa, M. & Vedral, V. Direct observation of any two-point quantum correlation function. Preprint at https://arxiv.org/abs/1312.4240 (2013).
39. Pashayan, H., Wallman, J. J. & Bartlett, S. D. Estimating outcome probabilities of quantum circuits using quasiprobabilities. *Phys. Rev. Lett.* **115**, 070501 (2015).
40. Wang, X., Wilde, M. M. & Su, Y. Efficiently computable bounds for magic state distillation. *Phys. Rev. Lett.* **124**, 090505 (2020).
41. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
42. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).
43. Takagi, R. Optimal resource cost for error mitigation. *Phys. Rev. Res.* **3**, 033178 (2020).
44. Jiang, J., Wang, K. & Wang, X. Physical implementability of linear maps and its application in error mitigation. *Quantum* **5**, 600 (2021).
45. Buhrman, H., Cleve, R., Watrous, J. & De Wolf, R. Quantum fingerprinting. *Phys. Rev. Lett.* **87**, 167902 (2001).
46. Cerezo, M., Sone, A., Volkoff, T., Cincio, L. & Coles, P. J. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nat. Commun.* **12**, 1–12 (2021).
47. Kiefer, J. & Wolfowitz, J. et al. Stochastic estimation of the maximum of a regression function. *Ann. Math. Stat.* **23**, 462–466 (1952).
48. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at https://arxiv.org/abs/1412.6980 (2014).
49. Kübler, J. M., Arrasmith, A., Cincio, L. & Coles, P. J. An adaptive optimizer for measurement-frugal variational algorithms. *Quantum* **4**, 263 (2020).
50. Arrasmith, A., Cincio, L., Somma, R. D. & Coles, P. J. Operator sampling for shot-frugal optimization in variational algorithms. *Quantum* **4**, 263 (2020).
51. Fontana, E., Cerezo, M., Arrasmith, A., Rungger, I. & Coles, P. J. Optimizing parametrized quantum circuits via noise-induced breaking of symmetries. Preprint at https://arxiv.org/abs/2011.08763 (2020).
52. Wang, S. et al. Noise-induced barren plateaus in variational quantum algorithms. *Nat. Commun.* **12**, 6961 (2021).
53. Hoeffding, W. Probability inequalities for sums of bounded random variables. In Fisher, N. I., Sen, P. K. (eds) *The Collected Works of Wassily Hoeffding*, 409–426 (Springer, 1994).
54. Peres, A. Separability criterion for density matrices. *Phys. Rev. Lett.* **77**, 1413 (1996).
55. Cerf, N. J., Adami, C. & Gingrich, R. M. Reduction criterion for separability. *Phys. Rev. A* **60**, 898–909 (1999).
56. Breuer, H.-P. Optimal entanglement criterion for mixed quantum states. *Phys. Rev. Lett.* **97**, 080501 (2006).
57. Hall, W. A new criterion for indecomposability of positive maps. *J. Phys. A: Math. Gen.* **39**, 14119 (2006).
58. Choi, M.-D. Positive semidefinite biquadratic forms. *Linear Algebra Appl.* **12**, 95–100 (1975).
59. Choi, M.-D. Some assorted inequalities for positive linear maps on c*-algebras. *J. Oper. Theory* **4**, 271–285 (1980).
60. Stormer, E. Positive linear maps of operator algebras. *Acta Math.* **110**, 233–278 (1963).
61. Woronowicz, S. L. Positive maps of low dimensional matrix algebras. *Rep. Math. Phys.* **10**, 165–183 (1976).
62. Wilde, M. M. *Quantum Information Theory* 2nd edn (Cambridge University Press, 2016).
63. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2011).
64. O'Donnell, R. & Wright, J. Efficient quantum tomography. In *Proc. 48th Annual ACM Symposium on Theory of Computing* 899–912 (ACM, 2016).
65. 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).
66. 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).
67. Arrasmith, A., Cerezo, M., Czarnik, P., Cincio, L. & Coles, P. J. Effect of barren plateaus on gradient-free optimization. *Quantum* **5**, 558 (2020).
68. Marrero, C. O., Kieferová, M. & Wiebe, N. Entanglement-induced barren plateaus. *PRX Quantum* **2**, 040316 (2021).
69. Grimsley, H. R., Economou, S. E., Barnes, E. & Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* **10**, 1–9 (2019).
70. Rattew, A. G., Hu, S., Pistoia, M., Chen, R. & Wood, S. A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver. Preprint at https://arxiv.org/abs/1910.09694 (2019).
71. Bilkis, M., Cerezo, M., Verdon, G., Coles, P. J. & Cincio, L. A semi-agnostic ansatz with variable structure for quantum machine learning. Preprint at https://arxiv.org/abs/2103.06712 (2021).
72. Skolik, A., McClean, J. R., Mohseni, M., van der Smagt, P. & Leib, M. Layerwise learning for quantum neural networks. *Quantum Mach. Intell.* **3**, 1–11 (2021).
73. Verdon, G. et al. Learning to learn with quantum neural networks via classical neural networks. Preprint at https://arxiv.org/abs/1907.05415 (2019).
74. Grant, E., Wossnig, L., Ostaszewski, M. & Benedetti, M. An initialization strategy for addressing barren plateaus in parametrized quantum circuits. *Quantum* **3**, 214 (2019).
75. Volkoff, T. & Coles, P. J. Large gradients via correlation in random parameterized quantum circuits. *Quantum Sci. Technol.* **6**, 025008 (2021).
76. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
77. Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495 (2019).
78. Zheng, M., Li, A., Terlaky, T. & Yang, X. A bayesian approach for characterizing and mitigating gate and measurement errors. Preprint at https://arxiv.org/abs/2010.09188 (2020).
79. Wang, K., Chen, Y.-A. & Wang, X. Measurement error mitigation via truncated Neumann series. Preprint at https://arxiv.org/abs/2103.13856 (2021).
80. Vidal, G. & Werner, R. F. Computable measure of entanglement. *Phys. Rev. A* **65**, 032314 (2002).
81. Plenio, M. B. Logarithmic negativity: a full entanglement monotone that is not convex. *Phys. Rev. Lett.* **95**, 090503 (2005).
82. Chen, L., Aulbach, M. & Hajdušek, M. Comparison of different definitions of the geometric measure of entanglement. *Phys. Rev. A* **89**, 042305 (2014).
83. Zhu, H., Hayashi, M. & Chen, L. Coherence and entanglement measures based on rényi relative entropies. *J. Phys. A: Math. Theor.* **50**, 475303 (2017).
84. Wang, K., Wang, X. & Wilde, M. M. Quantifying the unextendibility of entanglement. Preprint at https://arxiv.org/abs/1911.07433 (2019).
85. Müller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S. & Tomamichel, M. On quantum Rényi entropies: a new generalization and some properties. *J. Math. Phys.* **54**, 122203 (2013).
86. Wilde, M. M., Winter, A. & Yang, D. Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched Rényi relative entropy. *Commun. Math. Phys.* **331**, 593–622 (2014).
87. Mitarai, K., Negoro, M., Kitagawa, M. & Fujii, K. Quantum circuit learning. *Phys. Rev. A* **98**, 032309 (2018).
88. Abhijith, J. et al. Quantum algorithm implementations for beginners. Preprint at https://arxiv.org/abs/1804.03719 (2018).
89. Baidu Quleaf Simulator. https://github.com/baidu/QCompute (2020).
90. Paddle Quantum. https://github.com/paddlepaddle/Quantum (2020).
91. Ma, Y., Yu, D., Wu, T. & Wang, H. PaddlePaddle: an open-source deep learning platform from industrial practice. *Front. Data Comput.* **1**, 105–115 (2019).

92. Horodecki, M. & Horodecki, P. Reduction criterion of separability and limits for a class of distillation protocols. *Phys. Rev. A* **59**, 4206 (1999).

93. Wang, X. & Wilde, M. M. Cost of quantum entanglement simplified. *Phys. Rev. Lett.* **125**, 040502 (2020).

94. Rains, E. M. A semidefinite program for distillable entanglement. *IEEE Trans. Inf. Theory* **47**, 2921–2933 (2001).

95. Wang, X. & Duan, R. Improved semidefinite programming upper bound on distillable entanglement. *Phys. Rev. A* **94**, 050301 (2016).

96. Werner, R. F. Quantum states with Einstein–Podolsky–Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **40**, 4277 (1989).

97. Gurvits, L. Classical deterministic complexity of Edmonds' problem and quantum entanglement. In *Proc. 35th Annual ACM Symposium on Theory of Computing*, San Diego, CA, USA 10–19 (ACM Press, New York, NY, USA, 2003).

98. Gharibian, S. Strong NP-hardness of the quantum separability problem. *Quantum Inf. Comput.* **10**, 343–360 (2010).

99. Kraus, K. *States, Effects, and Operations*. (Springer-Verlag, Berlin, 1983).

AUTHOR CONTRIBUTIONS

X.W. formulated the initial idea and the framework; K.W. and X.W. developed the theory; Z.S. and X.Z. developed the entanglement quantification method; Z.S., X.Z., and Z.W. performed the experiments and delivered the analysis. All co-authors contributed to the preparation of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-022-00556-w.

Correspondence and requests for materials should be addressed to Xin Wang.

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

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