

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,†}

¹*Institute for Quantum Computing, Baidu Research, Beijing 100193, China*

(Dated: December 2, 2021)

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.

I. 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], quantum metrology [7, 8], and quantum data compression [9–11]. 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 [12, 13] for a detailed review on VQA.

Quantum entanglement [14], the most nonclassical manifestation of quantum mechanics, has been identified as invaluable resource enabling a tremendous number of tasks ranging from quantum information processing [15, 16], quantum cryptography [17–19], quantum algorithms [20–22], quantum communication [23, 24], to measurement-based quantum computing [25–28]. As so, the ability to manipulate quantum entanglement 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 [14, 29, 30]. For example, entanglement can be detected via entanglement witnesses [31, 32], Bell's inequalities [33], quantum Fisher information [34], realignment criterion [35, 36], range

criterion [37], and majorization criterion [38], to name a few. 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 [39, 40]. 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.

In this paper, we combine VQA and the quasi-probability decomposition technique [41–47] to propose the Variational Entanglement Detection (VED) and Variational Logarithmic Negativity Estimation (VLNE) algorithms, aiming 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. Our main contributions can be summarized as follows:

1. We combine VQA and the quasi-probability decomposition technique to propose the VED framework, con-

*wangkun28@baidu.com

†wangxin73@baidu.com

tributing a feasible solution for detecting entanglement on near-term quantum devices.

2. We combine VQA and the quasi-probability decomposition technique to propose the VLNE algorithm that could estimate the well-known logarithmic negativity [48], which may lead to various applications in near-term quantum information processing.

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

Our paper is structured as follows. In Sec. II, we set the notations and briefly summarize the entanglement criteria based on positive maps. In Sec. III, we present our first main result: the VED framework. In Sec. IV, we elaborate on three prominent positive maps to illustrate how the VED framework is applied. In Sec. V, we present our second main result: the VLNE framework, which estimates the logarithmic negativity entanglement measure. In Secs. VI and VII, experiments on IBM-Q and numerical simulations on Paddle Quantum are conducted on various bipartite quantum states of interests to show the validity of the proposed methods. We finally conclude in Sec. VIII.

II. PRELIMINARIES

A. Notations

In this section, we set the notation and define several quantities that will be used throughout this paper. We will frequently use symbols such as $\mathcal{H}_A$ and $\mathcal{H}_B$ to denote Hilbert spaces associated with quantum systems A and B , respectively. We use d_A to denote the dimension of the system A . The set of linear operators acting on A is denoted by $\mathcal{L}(\mathcal{H}_A)$. We write an operator with subscript to indicate the system that the operator acts on, e.g., X_{AB} , and write $X_A := \text{Tr}_B X_{AB}$. For a linear operator $X \in \mathcal{L}(\mathcal{H}_A)$, we define its modulus $|X| := \sqrt{X^\dagger X}$, where $X^\dagger$ is the adjoint operator of X . The trace norm of X is defined as $\|X\|_1 := \text{Tr}|X|$. We use $X \geq 0$ to indicate that A is positive semidefinite. A quantum map $\mathcal{N}$ that transforms linear operators to linear operators in the system A is abbreviated as $\mathcal{N}_{A \rightarrow A}$. We use the calligraphic letters (e.g., $\mathcal{N}$, $\mathcal{R}$, and $\mathcal{O}$) to represent linear quantum maps and use id_A to represent the identity map on system A . We say $\mathcal{N}_{A \rightarrow A}$ is trace-preserving if $\text{Tr}[\mathcal{N}(X)] = \text{Tr}[X]$ for arbitrary $X \in \mathcal{L}(\mathcal{H}_A)$, is positive if $\mathcal{N}(X) \geq 0$ for arbitrary $X \geq 0$, and is completely positive if $\text{id}_R \otimes \mathcal{N}$ is positive for arbitrary reference system R .

Given a Hermitian operator X in system A , we denote by $\lambda_{\min}(X)$ its minimal eigenvalue. We have the following variational characterization:

$$\lambda_{\min}(X) = \min_{|\psi\rangle} \langle \psi | X | \psi \rangle, \quad (1)$$

where the minimization ranges over the set of pure states on system A .

We use $\mathbb{R}$ to represent the real field. We introduce the sign function $\text{sgn} : \mathbb{R} \rightarrow \{\pm 1, 0\}$ as $\forall x < 0$, $\text{sgn}(x) = -1$;

$\forall x > 0$, $\text{sgn}(x) = 1$; and $\text{sgn}(0) = 0$. All logarithms are in base 2 in this paper.

B. Pauli operators and Pauli channels

The four Pauli operators in the qubit space are defined as

$$I \equiv \sigma_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X \equiv \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (2a)$$

$$Y \equiv \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z \equiv \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2b)$$

They provide a basis for the qubit linear operators, i.e., arbitrary qubit linear operator can be decomposed w.r.t. this basis. For the n -qubit case, one can construct a set of Pauli operators, which we call the Pauli set, as

$$\mathbf{P}_n := \left\{ \bigotimes_{k=1}^n \sigma_{q_k} \mid q_k = 0, 1, 2, 3 \right\} \equiv \{I, X, Y, Z\}^{\otimes n}. \quad (3)$$

The n -qubit Pauli set has size $|\mathbf{P}_n| = 4^n$. Note that $\mathbf{P}_n$ forms a basis for the n -qubit linear operators. One can see from (3) that each Pauli operator $P \in \mathbf{P}_n$ can be represented uniquely by a *quaternary* sequence:

$$\mathbf{q} = q_1 \cdots q_n, \quad \text{with } q_k = 0, 1, 2, 3. \quad (4)$$

As so, we use $P_{\mathbf{q}}$ to represent the n -qubit Pauli operator that is uniquely determined by the sequence $\mathbf{q}$. Since each qubit Pauli operator is unitary, so is each n -qubit Pauli operator, due to the construction. Given a Pauli operator $P \in \mathbf{P}_n$, we denote by $\mathcal{P}(\cdot) := P(\cdot)P^\dagger$ its induced Pauli channel.

C. 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 [49]

$$\rho_{AB} = \sum_x p_x |\psi_x\rangle\langle\psi_x|_A \otimes |\phi_x\rangle\langle\phi_x|_B, \quad (5)$$

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 [50, 51]. There are many separability criteria that have been proposed to determine the separability or entanglement of bipartite quantum states as necessary conditions [14, 29].

One of the most celebrated criteria for distinguishing separable states from entangled states are 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 $\mathcal{N}_{B \rightarrow B}$ 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 $\mathcal{N}_{B \rightarrow B}(\rho_{AB})$ may be no longer positive; the transpose map is a prominent example [52]. 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 $\mathcal{N}_{B \rightarrow C}$, it holds that $\mathcal{N}_{B \rightarrow C}(\rho_{AB}) \geq 0$ [31].

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 [53] and thus cannot be realized in near-term quantum devices. In the following, we show how to overcome this obstacle and employ the positive map criterion to detect entanglement on NISQ devices.

III. QUANTUM ENTANGLEMENT DETECTION

In this section, we integrate variational quantum algorithms with the quasi-probability decomposition technique [41–47] 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 [45, 46]. Alternatively, one may set Δ to be the set of Pauli channels induced by Pauli operators from the Pauli set (3). For a positive (but not completely positive) and trace-preserving map $\mathcal{N}_{B \rightarrow B}$, we assume that it can be decomposed w.r.t. Δ as

$$\mathcal{N}(\cdot) = \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \mathcal{O}(\cdot), \quad r_{\mathcal{O}} \in \mathbb{R}. \quad (6)$$

Note that such a decomposition always exists if Δ contains a universal basis [45]. 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

$$\sigma_{AB} := \mathcal{N}_{B \rightarrow B}(\rho_{AB}) = \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \mathcal{O}_{B \rightarrow B}(\rho_{AB}). \quad (7)$$

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 $\lambda_{\min}(\sigma_{AB})$ as 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 $\mathcal{N}_{B \rightarrow B}$. This highlights the importance of computing or estimating $\lambda_{\min}(\sigma_{AB})$ in entanglement detection.

A. Deterministic detection

As we have argued, σ_{AB} cannot be obtained directly via $\mathcal{N}(\rho)$ since $\mathcal{N}$ does not represent physically implementable quantum operations. Fortunately, the decomposition (7) 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 (7) of the target state σ_{AB} and the framework goes as follows. First of all, by (1) it holds that

$$\lambda_{\min}(\sigma_{AB}) = \min_{|\psi\rangle_{AB}} \langle \psi | \sigma_{AB} | \psi \rangle \quad (8)$$

$$= \min_{|\psi\rangle_{AB}} \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \langle \psi | \mathcal{O}(\rho_{AB}) | \psi \rangle, \quad (9)$$

where the minimization ranges over all pure bipartite quantum states $|\psi\rangle_{AB}$ in AB . We use a variational quantum circuit with parameters α to prepare the test state $|\psi\rangle$. More precisely, we choose a parametrized quantum circuit ansatz that generates a unitary $U(\alpha)$ and prepare the test state via $|\psi\rangle = U(\alpha)|0\rangle^{\otimes 2n}$. Each inner product $\langle \psi | \mathcal{O}(\rho) | \psi \rangle$ in (9) can be estimated via the canonical Swap Test subroutine [54], as both $U(\alpha)$ 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 efficient estimating procedure which uses $2n$ qubits and avoids the use of expensive SWAP gate. First of all, notice that

$$\langle \psi(\alpha) | \mathcal{O}(\rho_{AB}) | \psi(\alpha) \rangle \quad (10)$$

$$= \text{Tr} [\mathcal{O}(\rho_{AB}) U_{\alpha} |0^{2n}\rangle \langle 0^{2n}| U_{\alpha}^{\dagger}] \quad (11)$$

$$= \langle 0^{2n} | U_{\alpha}^{\dagger} \mathcal{O}(\rho_{AB}) U_{\alpha} | 0^{2n} \rangle, \quad (12)$$

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 (9). With these data in hand, we compute the following loss function

$$L(\alpha) := \sum_{\mathcal{O} \in \Delta} r_{\mathcal{O}} \langle \psi(\alpha) | \mathcal{O}(\rho_{AB}) | \psi(\alpha) \rangle. \quad (13)$$

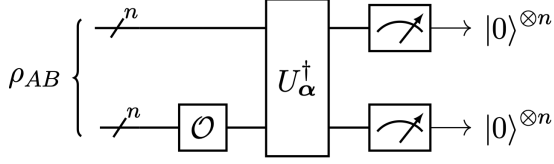


FIG. 1: The simplified quantum circuit that estimates the overlap $\langle \psi | \mathcal{O}(\rho_{AB}) | \psi \rangle$ in (12) for a given implementable operation $\mathcal{O}$, where $|\psi\rangle := U_\alpha |0\rangle^{\otimes 2n}$ is the parameterized input state.

At last, we perform gradient-based optimization methods including SGD [55] and Adam [56] to minimize the loss function $L(\alpha)$ by varying the parameters α , whose value will determine the separability of the input state ρ_{AB} . More precisely, if $L(\alpha)$ 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 $L(\alpha) < -\delta$ 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 $L(\alpha)$ since the condition $L(\alpha) < 0$ 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 $L(\alpha)$ in advance to save the optimization cost. The detailed VED framework is summarized in Algorithm 1. 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 (6) of the positive map $\mathcal{N}$, parameterized quantum circuits $U(\alpha)$ with initial parameters α , and tolerance δ ;
 - 2: Initialize $L(\alpha) = 0$;
 - 3: **for all** $\mathcal{O} \in \Delta$ such that $r_{\mathcal{O}} \neq 0$ **do**
 - 4: Apply U_α to $|0\rangle^{\otimes 2n}$ and obtain test state $|\psi\rangle = U_\alpha |0\rangle^{\otimes 2n}$;
 - 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 $L(\alpha) = L(\alpha) + r_{\mathcal{O}} c_{\mathcal{O}}$, where $r_{\mathcal{O}}$ is given by the decomposition (6);
 - 7: **end for**
 - 8: Perform optimization methods to minimize $L(\alpha)$; terminate the optimization if the error tolerance is satisfied: $L(\alpha) < -\delta$;
 - 9: Output "Entangled" if the optimized $L(\alpha) < -\delta$.
-

B. 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 $L(\alpha)$. Actually, $L(\alpha)$ can be estimated in a probabilistic way using the sampling technique, by virtue of the quasi-probability decomposition (6). This new method would be beneficial when the number of decomposed operations in (6) with non-zero coefficients is large while the sampling cost is relatively low. Now we describe the sam-

pling method accurately. First of all, notice that the decomposition (6) 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.,

$$p_{\mathcal{O}} := \frac{|r_{\mathcal{O}}|}{\gamma}, \quad \gamma := \sum_{\mathcal{O} \in \Delta} |r_{\mathcal{O}}|. \quad (14)$$

Substituting (14) into (13) yields

$$L(\alpha) = \gamma \sum_{\mathcal{O} \in \Delta} \text{sgn}(r_{\mathcal{O}}) p_{\mathcal{O}} \langle \psi(\alpha) | \mathcal{O}(\rho_{AB}) | \psi(\alpha) \rangle \quad (15)$$

$$= \mathbb{E}_{\mathcal{O}} [\gamma \text{sgn}(r_{\mathcal{O}}) \langle \psi(\alpha) | \mathcal{O}(\rho_{AB}) | \psi(\alpha) \rangle], \quad (16)$$

where $\mathbb{E}(X)$ denotes the expectation of the random variable X , and the expectation in (16) is evaluated w.r.t. the probability distribution $\{p_{\mathcal{O}}\}_{\mathcal{O} \in \Delta}$. Based on (16), 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.

Algorithm 2 Probabilistic VED

- 1: Input: $2n$ -qubit quantum state ρ_{AB} , decomposition (6) of the positive map $\mathcal{N}$, parameterized quantum circuits $U(\alpha)$ with initial parameters α , error tolerance δ , and fail probability ε ;
 - 2: Initialize $L'(\alpha) = 0$;
 - 3: Compute γ defined in (14) and set $M = 2\gamma^2 \log(2/\varepsilon)/\delta^2$;
 - 4: **for all** $m = 1, \dots, M$ **do**
 - 5: Apply U_α to $|0\rangle^{\otimes 2n}$ and obtain test state $|\psi\rangle = U_\alpha |0\rangle^{\otimes 2n}$;
 - 6: Sample a quantum operation $\mathcal{O}^{(m)}$ from Δ according to the probability distribution $\{p_{\mathcal{O}}\}_{\mathcal{O} \in \Delta}$ in (14); Let $r^{(m)}$ be the coefficient of $\mathcal{O}^{(m)}$ in (6);
 - 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 $L^{(m)} = \gamma \text{sgn}(r^{(m)}) c^{(m)}$;
 - 9: **end for**
 - 10: Compute the loss function $L'(\alpha) = \frac{1}{M} \sum_{m=1}^M L^{(m)}$;
 - 11: Perform optimization methods to minimize $L'(\alpha)$; terminate the optimization if the error tolerance is satisfied: $L'(\alpha) < -\delta$;
 - 12: Output "Entangled" if the optimized $L'(\alpha) < -\delta$.
-

Let's analyze Algorithm 2 in depth. First, we remark that the obtained $L'(\alpha)$ in step 11 of Algorithm 2 is an unbiased estimator of true value $L(\alpha)$ due to (16). Second, since $|L^{(m)}| \leq \gamma$, we can apply the Hoeffding inequality [57] to ensure that $M = 2\gamma^2 \log(2/\varepsilon)/\delta^2$ number of samples would estimate the true value $L(\alpha)$ within error δ with success probability no less than $1 - \varepsilon$, i.e.,

$$p(|L'(\alpha) - L(\alpha)| \leq \delta) \geq 1 - \varepsilon. \quad (17)$$

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 (17) 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 (17)

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 (6) 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 sacrifices precision for efficiency in entanglement detection.

IV. PROMINENT POSITIVE MAPS

In Sec. III 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 [52], the reduction map [58], and the enhanced reduction map [59, 60]—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 (3), i.e.,

$$\Delta := \{P \mid \mathcal{P}(\cdot) = P(\cdot)P^\dagger, P \in \mathbf{P}_n\}. \quad (18)$$

For each of the three positive maps under consideration, we firstly decompose it w.r.t. Δ as (6) 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 PPT 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 output "Entangled." We also show by example how VED works in qutrit systems in Appendix A, utilizing the Choi map [61, 62].

A. PPT criterion

A necessary condition for entanglement detection is the positive partial transpose (PPT) criterion [52], which we briefly review as follows. Let ρ_{AB} be a bipartite quantum

state. We can express it as

$$\rho_{AB} = \sum_{ijkl} \alpha_{ijkl} |i\rangle\langle j|_A \otimes |k\rangle\langle l|_B, \quad (19)$$

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

$$\rho_{AB}^{T_B} := (\text{id}_A \otimes T_B)(\rho_{AB}) \quad (20)$$

$$= \sum_{ijkl} \alpha_{ijkl} |i\rangle\langle j| \otimes (|k\rangle\langle l|)^T \quad (21)$$

$$= \sum_{ijkl} \alpha_{ijkl} |i\rangle\langle j| \otimes |l\rangle\langle k|, \quad (22)$$

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 [31, 63, 64].

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 (18):

$$T(\rho) = \frac{\rho + X\rho X - Y\rho Y + Z\rho Z}{2}, \quad (23)$$

where X, Y, Z are the Pauli matrices defined in (2). The validity of this decomposition can be checked by direct calculation. Substituting (23) into (22), we obtain

$$\rho_{AB}^{T_B} := (\text{id}_A \otimes T_B)(\rho_{AB}) \quad (24)$$

$$= \frac{1}{2}(\rho + X_B \rho X_B - Y_B \rho Y_B + Z_B \rho Z_B), \quad (25)$$

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 $Y \rho Y$ and $Z \rho Z$. Adapting the decomposition (25) 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 $B \equiv B_1 B_2 \cdots B_n$ be a composite system with n qubits, i.e., B_i represents the i -th qubit system. A key observation is that the transpose operation satisfies the tensor product property: transposing the composite system B is equivalent to transposing the local qubit systems B_i individually. More precisely,

$$T_B = \bigotimes_{i=1}^n T_{B_i}, \quad (26)$$

where T_{B_i} is the transpose operation on the i -th qubit. Eqs. (26) and (23) together give T_B a linear combination into Pauli channels of 4^n terms in total. Using this decomposition, we may apply VED (Algorithm 1 or Algorithm 2) to accomplish the multi-qubit PPT criterion deterministically or probabilistically.

B. Reduction criterion

In this section, we first review the reduction criterion [58] and then propose a variational algorithm implementing this criterion within the VED framework described in III.

$$\mathcal{R}_{B \rightarrow B}(X_B) := \text{Tr}[X_B]I_B - X_B, \quad (27)$$

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

$$\sigma_{AB} := (\text{id}_A \otimes \mathcal{R}_{B \rightarrow B})(\rho_{AB}) \geq 0. \quad (28)$$

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 III, we have first to decompose $\mathcal{R}_{B \rightarrow B}$ into a linear combination of Pauli channels. Indeed, we can do so since

$$\mathcal{R}_{B \rightarrow B}(\rho_B) := \text{Tr}[\rho_B]I_B - \rho_B \quad (29)$$

$$= \frac{1}{2^n} \sum_{\mathbf{q} \in \{0,1,2,3\}^{\otimes n}} \mathcal{P}_{\mathbf{q}}(\rho_B)I_B - \rho_B \quad (30)$$

$$= \frac{1-2^n}{2^n} \rho_B + \frac{1}{2^n} \sum_{\mathbf{q} \neq \mathbf{0}} \mathcal{P}_{\mathbf{q}}(\rho_B), \quad (31)$$

where $\mathbf{0} \equiv (0, \dots, 0)$ of size n , and the second equality follows from the twirling property of Pauli channels [65, 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

$$\mathcal{R}_{B \rightarrow B}(\rho) = \frac{-\rho + X\rho X + Y\rho Y + Z\rho Z}{2}. \quad (32)$$

As one might see, deterministic VED using the reduction criterion is not efficient in the multi-qubit case since it has to compute exponentially many numbers of overlaps: for a $2n$ -qubit bipartite quantum state, one has to compute 4^n overlaps. In Section IV E, we consider another version with better efficiency by exploring the simple structure of the reduction map (27).

On the other hand, probabilistic VED using the reduction criterion is also not efficient due to the sampling cost $\gamma \approx 2^n$ for $\mathcal{R}$. This observation leads to a simple way to improve the efficiency of the probabilistic VED. For example, we introduce the trace-preserving reduction map, defined via

$$\hat{\mathcal{R}}_{B \rightarrow B}(\rho) := \frac{1}{2^n - 1} \mathcal{R}_{B \rightarrow B}(\rho_B). \quad (33)$$

One can check that as a positive map, $\hat{\mathcal{R}}$ has the same entanglement detection range as the original reduction map $\mathcal{R}$. On the contrary, the sampling cost of $\hat{\mathcal{R}}$ is $\gamma = 1 + 1/2^n$, which is exponentially smaller than that of $\mathcal{R}$. This implies that probabilistic VED using the trace-preserving reduction

map $\hat{\mathcal{R}}$ is very efficient in terms of the number of samples consumed. However, the efficiency is achieved at the cost of the high precision required in estimating the minimal eigenvalue of $\hat{\mathcal{R}}(\rho_{AB})$, which decreases exponentially in n in general. We point out this is a trade off between efficiency and precision. Similar arguments can be applied to the transpose map and the enhanced reduction map discussed in the next section.

C. Enhanced reduction criterion

In this section, we consider an enhanced version of the reduction map [59, 60] 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 Sec. III to detect entanglement.

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

$$U_a = \text{antidiag}(1, -1, 1, -1, \dots, 1, -1), \quad (34)$$

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

$$U_a = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} = X \otimes iY. \quad (35)$$

Indeed, one can check that the n -qubit U_a can be decomposed w.r.t. the Pauli set as $U_a = X \otimes \dots \otimes X \otimes iY$, where there are $n - 1$ X operators in the tensor product. Based on U_a , we define the following map [59]

$$\mathcal{K}_{B \rightarrow B}(\rho_B) := \mathcal{R}_{B \rightarrow B}(\rho_B) - U_a T_B(\rho) U_a^\dagger, \quad (36)$$

where $\mathcal{R}_{B \rightarrow B}$ is the reduction map defined in (27) and T_B is the transpose map defined in (22). This map has been shown to be positive but not completely positive [59]. 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 (31) of $\mathcal{R}$ and the Pauli decomposition (26) of T_B into (36) 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 on this decomposition, we can use VED (Algorithm 1 or Algorithm 2) to accomplish the enhanced reduction criterion.

D. Discussion on the measurement cost

Note that another way to detect and quantify entanglement of a state ρ_{AB} is to obtain its density matrix via quantum state tomography [66]. Full density matrix reconstruction of an unknown $(n_A + n_B)$ -qubit state in the worst-case costs exponential copies of the state [67, 68], e.g., $\tilde{\Omega}(4^{n_A + n_B})$ measurement

results are necessary to reconstruct a matrix close to ρ in the sense of trace distance [68]. 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 gradient-based optimization loops. Besides the measurement cost, methods based on state tomography need vast memory to store and process the density matrix on a classical computer, which are resource-demanding as well. The VED framework, on the other hand, doesn't require such classical memory and post-processing.

E. VED based on reduction criterion without decomposition

In Sec. IV B 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 $\mathcal{R}_{B \rightarrow B}(\rho_{AB})$ is not semidefinite positive. Using (1), this means that

$$\min_{|\psi_{AB}\rangle} \langle \psi | \mathcal{R}_{B \rightarrow B}(\rho_{AB}) | \psi \rangle \quad (37)$$

$$= \min_{|\psi_{AB}\rangle} \langle \psi | (I_A \otimes \rho_B - \rho_{AB}) | \psi \rangle \quad (38)$$

$$= \min_{|\psi_{AB}\rangle} \{ \text{Tr}[\psi_B \rho_B] - \text{Tr}[\psi_{AB} \rho_{AB}] \} < 0, \quad (39)$$

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 Eq. (39), 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 new 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 circuits $U(\alpha)$ with initial parameters α , and tolerance δ ;
 - 2: Apply $U(\alpha)$ to $|00\rangle_{AB}$ on system AB and obtain the test state $|\psi\rangle_{AB} = U(\alpha)|00\rangle_{AB}$;
 - 3: Compute the overlap between state ψ_B and ρ_B on subsystem B using the Swap Test and obtain $c_1 = \text{Tr}[\psi_B \rho_B]$;
 - 4: Apply $U(\alpha)$ to $|00\rangle_{AB}$ on system AB and obtain the test state $|\psi\rangle_{AB} = U(\alpha)|00\rangle_{AB}$;
 - 5: Compute the overlap between state ψ_{AB} and ρ_{AB} using the Swap Test and obtain $c_2 = \text{Tr}[\psi_{AB} \rho_{AB}]$;
 - 6: Compute the loss function $L(\alpha) = c_1 - c_2$;
 - 7: Perform optimization methods to minimize $L(\alpha)$; terminate the optimization if the error tolerance is satisfied: $L(\alpha) < -\delta$.
 - 8: Output "Entangled" if the optimized $L(\alpha) < -\delta$.
-

We remark that this idea can also be adopted to improve the efficiency of VED using the enhanced reduction criterion.

V. QUANTUM ENTANGLEMENT QUANTIFICATION

One of the most well-known entanglement measure is the logarithmic negativity [69, 70], which has various applications in quantum information theory. For a bipartite state ρ_{AB} , its logarithmic negativity is defined as

$$E_N(\rho_{AB}) := \log \|\rho_{AB}^{T_B}\|_1. \quad (40)$$

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

$$\|\rho_{AB}^{T_B}\|_1 = 2 \max_U \text{Tr} |0\rangle\langle 0|_R Q_R - \text{Tr} \rho_{AB}^{T_B} \quad (41)$$

$$= 2 \max_U \text{Tr} |0\rangle\langle 0|_R Q_R - 1, \quad (42)$$

where $Q_R = \text{Tr}_{AB} Q_{ABR}$, $Q_{ABR} = U(\rho_{AB}^{T_B} \otimes |0\rangle\langle 0|_R)U^\dagger$, 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 (25) and (26), compute the overlaps in (42) 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 (26). 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.

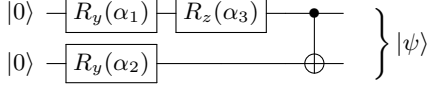


FIG. 2: Parameterized two-qubit quantum circuit $U(\alpha)$ used for preparing the test state $\psi_{AB}(\alpha)$ on the ibmq-santiago hardware. The parameters α are randomly initialized as $(\alpha_1, \alpha_2, \alpha_3) = (3.2292, 4.8579, 5.4691)$.

Algorithm 4 Variational Logarithmic Negativity Estimation

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

$$\rho_{AB} \otimes |0\rangle\langle 0|_R, \quad (43)$$

$$(I_A \otimes X_B)\rho_{AB}(I_A \otimes X_B) \otimes |0\rangle\langle 0|_R, \quad (44)$$

$$(I_A \otimes Y_B)\rho_{AB}(I_A \otimes Y_B) \otimes |0\rangle\langle 0|_R, \quad (45)$$

$$(I_A \otimes Z_B)\rho_{AB}(I_A \otimes Z_B) \otimes |0\rangle\langle 0|_R, \quad (46)$$

and obtain the states $\sigma^{(0)}, \sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}$, respectively.

- 3: Obtain $o_j = \text{Tr}[\sigma_R^{(j)} |0\rangle\langle 0|_R]$ 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 $\mathcal{L}_1(\alpha)$;
 - 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 [71–73] based on the sandwiched Rényi relative entropy [74, 75] of order $1/2$, making use of the recently proposed variational quantum algorithm estimating the fidelity between two quantum states [6].

VI. EXPERIMENTS IN 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 $\mathcal{R}_{B \rightarrow B}$ defined in Eq. (27). After implementing the decomposed reduction map by 4 Pauli terms as Eq. (32), we use a parametrized quantum circuit $U(\alpha)$ to prepare 4 identical test states $\psi_{AB}(\alpha) = U(\alpha)|00\rangle\langle 00|_{AB}U^\dagger(\alpha)$ and compute the loss function defined in Eq. (13). The PQC used is depicted in Fig. 2 with three randomly initialized parameters $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. During the optimization procedure, we apply the gradient descent algorithm [76] to guide the learning process where the analytical gradient is calculated via the following parameter-shift rule [77]:

$$\frac{\partial \mathcal{L}(\alpha)}{\partial \alpha_j} := \frac{1}{2} \left[\mathcal{L}\left(\alpha_j + \frac{\pi}{2}\right) - \mathcal{L}\left(\alpha_j - \frac{\pi}{2}\right) \right]. \quad (47)$$

Due to the finite sampling restriction for measurements, the optimization procedure essentially falls into the regime of Stochastic Gradient Descent (SGD) [55]. 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 [12] 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 to 5 rounds are sufficient) to obtain the detection result. As mentioned in Ref. [78], 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. As a comparison, we conduct numerical simulations on Quantum Leaf platform and receive a similar result by repeating the simulation on qiskit-Aer simulator. We summarize the experimental and numerical results in Fig. 3.

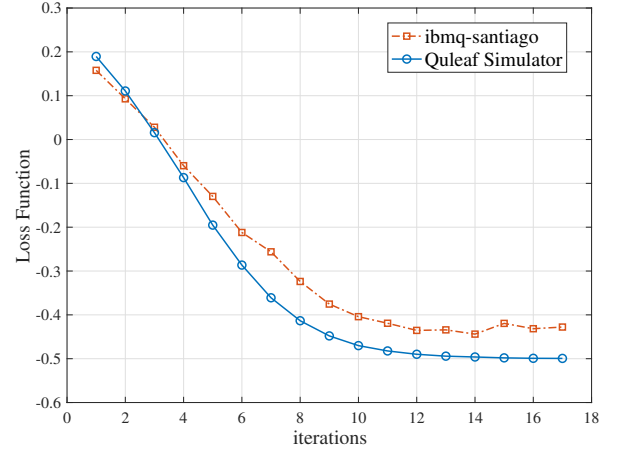


FIG. 3: Estimated $\lambda_{\min}$ by VED using the reduction criterion on the Bell state $|\Phi\rangle$. The red curve records the results from ibmq-santiago with shots = 8192 for each circuit evaluation. The blue curve records the simulation results on Quantum Leaf platform using the Baidu Quleaf simulator [79]. Learning rate in the gradient descent algorithm is set to be LR = 0.5.

VII. NUMERICAL SIMULATIONS

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 [80] toolkit on the PaddlePaddle Deep Learning Platform [81, 82].

A. Entanglement detection

For the entanglement detection purpose, we adopt the circuit ansatz shown in Fig. 4 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.

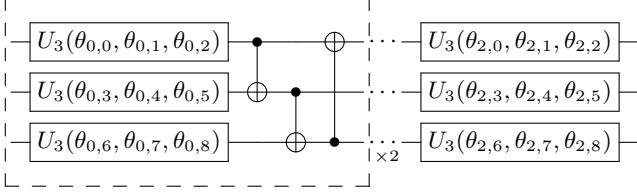


FIG. 4: Three qubit parameterized ansatz $U(\alpha)$ used for VED. The quantum circuit within the dotted block is repeated twice.

1. Four-qubit isotropic states

The four-qubit isotropic state family is defined as [83, Eq. (32)]

$$\rho_{AB}^{\text{iso}}(p) := p\Phi_{AB} + (1-p)\frac{I_{AB}}{16}, \quad (48)$$

where $p \in [0, 1]$ is a parameter, Φ_{AB} is the four-qubit maximally entangled state, and I_{AB} is the identity operator in AB in which $A \equiv A_1A_2$ 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. Intuitively, the isotropic state is a convex combination of the maximally entangled state Φ_{AB} and the maximally mixed state $I_{AB}/16$. It has been shown that $\rho_{AB}^{\text{iso}}(p)$ is separable (w.r.t. the $A:B$ cut) if and only if $p \leq 1/5$ [83].

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 Sec. IV, 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. 5. 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 $L(\alpha)$, 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.

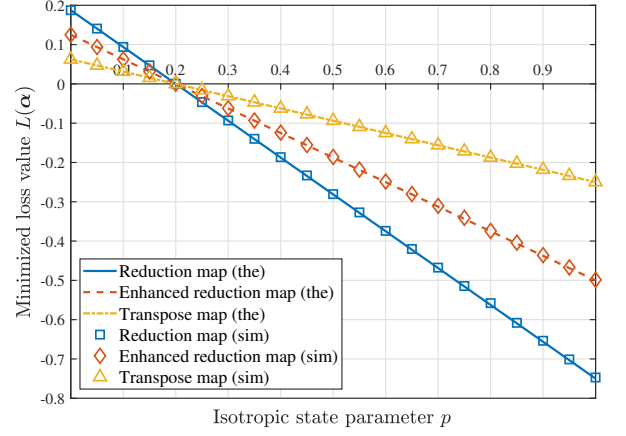


FIG. 5: Numerical results on the four-qubit isotropic states defined in (48). Each line depicts the smallest eigenvalue of every isotropic state with parameter $p \in [0, 1]$ under the corresponding map. This line of the smallest eigenvalues is a lower bound of the loss function $L(\alpha)$. Each marker depicts the minimized loss value obtained by simulations (sim) of Algorithm 1 on a chosen isotropic state, aligning with the theoretical line.

2. Four-qubit Breuer states

As we have mentioned in Sec. IV C, 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 [59, Eq. (7)]

$$\rho_{AB}^{\text{Breuer}}(\lambda) := \begin{pmatrix} \frac{1-\lambda}{3} & 0 & 0 & 0 \\ 0 & \frac{1+2\lambda}{6} & \frac{1-4\lambda}{6} & 0 \\ 0 & \frac{1-4\lambda}{6} & \frac{1+2\lambda}{6} & 0 \\ 0 & 0 & 0 & \frac{1-\lambda}{3} \end{pmatrix}, \quad (49)$$

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 $\rho_{AB}^{\text{Breuer}}$ is separable (w.r.t. the $A:B$ cut) if and only if $\lambda = 0$ and can be detected by the enhanced reduction criterion [59]. On the other hand, it has positive partial transpose if and only if $\lambda \leq 1/6$ [59], 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. 6 by markers, which again aligns 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.

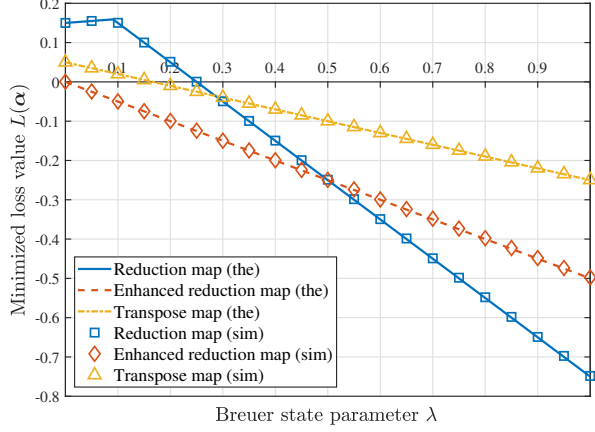


FIG. 6: Numerical results on the four-qubit Breuer states defined in (49). Each line depicts the smallest eigenvalue of every Breuer state with parameter $p \in [0, 1]$ under the corresponding map. This line of the smallest eigenvalues is a lower bound of the loss function $L(\alpha)$. Each marker depicts the minimized loss value obtained by simulations (sim) of Algorithm 1 on a chosen Breuer state, aligning with the theoretical line.

B. Logarithmic negativity estimation

For simulations of variational entanglement quantification with logarithmic negativity, we adopt the hardware efficient ansatz used for trace distance estimation in [6] where the circuit depth is 4. The simulations are carried out on two-qubit isotropic states, which is defined as

$$\rho_{AB}^{\text{iso}}(p) := p\Phi_{AB} + (1-p)\frac{I_{AB}}{4}, \quad (50)$$

where Φ_{AB} is the two-qubit maximally entangled state. As shown in Fig. 7, 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. 7, agree with the precisely calculated values given by the blue line.

VIII. CONCLUSION AND OUTLOOK

In this work, we combined two novel 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

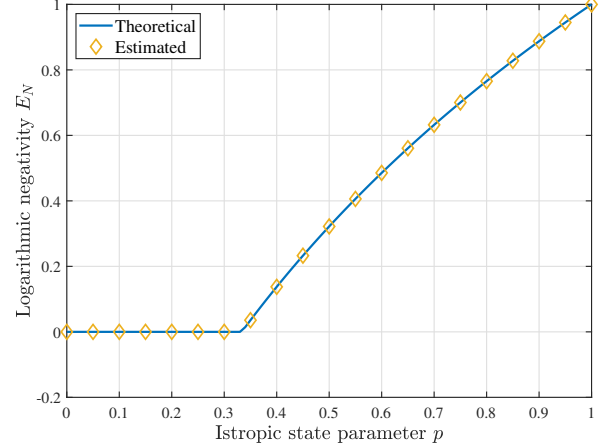


FIG. 7: Numerical results on the two-qubit isotropic states. The blue line represents the precisely calculated logarithmic negativity of isotropic states with parameter $p \in [0, 1]$. The yellow markers depict the estimated logarithmic negativity by simulations of Algorithm 4 on selected isotropic states.

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 Sec. V we showed by case how to variationally compute the log-negativity entanglement measure. It would be meaningful to design novel quantum algorithms to estimate other distance-based entanglement measures [72, 84–87].

Acknowledgements

We thank Runyao Duan for helpful suggestions. XW would like to thank Youle Wang for useful discussions.

-
- [1] J. Preskill, *Quantum* **2**, 79 (2018).
- [2] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, *New Journal of Physics* **18**, 023023 (2016).
- [3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, *Nature communications* **5**, 4213 (2014).
- [4] K. M. Nakanishi, K. Mitarai, and K. Fujii, *Physical Review Research* **1**, 033062 (2019).
- [5] M. Cerezo, A. Poremba, L. Cincio, and P. J. Coles, *Quantum* **4**, 248 (2020).
- [6] R. Chen, Z. Song, X. Zhao, and X. Wang, *arXiv:2012.05768*, **1** (2020), *arXiv:2012.05768*.
- [7] J. L. Beckey, M. Cerezo, A. Sone, and P. J. Coles, *arXiv:2010.10488*, **1** (2020), *arXiv:2010.10488*.
- [8] B. Koczor, S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin, *arXiv:1908.08904*, **1** (2019), *arXiv:1908.08904*.
- [9] J. Romero, J. P. Olson, and A. Aspuru-Guzik, *Quantum Science and Technology* **2**, 045001 (2017), *arXiv:1612.02806*.
- [10] K. H. Wan, O. Dahlsten, H. Kristjánsson, R. Gardner, and M. S. Kim, *npj Quantum Information* **3**, 36 (2017), *arXiv:1612.01045*.
- [11] C. Cao and X. Wang, *arXiv preprint arXiv:2012.08331* (2020), *arXiv:2012.08331*.
- [12] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, *arXiv:2011.01382*, **1** (2020), *arXiv:2011.01382*.
- [13] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, *arXiv:2012.09265*, **1** (2020), *arXiv:2012.09265*.
- [14] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Reviews of Modern Physics* **81**, 865 (2009).
- [15] C. H. Bennett and S. J. Wiesner, *Physical Review Letters* **69**, 2881 (1992).
- [16] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Physical review letters* **70**, 1895 (1993).
- [17] J. S. Bell, *Physics Physique Fizika* **1**, 195 (1964).
- [18] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Physical Review Letters* **24**, 549 (1970).
- [19] A. K. Ekert, *Physical Review Letters* **67**, 661 (1991).
- [20] P. W. Shor, in *Proceedings 35th annual symposium on foundations of computer science* (Ieee, 1994) pp. 124–134.
- [21] R. Jozsa, *arXiv preprint quant-ph/9707034* (1997).
- [22] A. M. Childs and W. Van Dam, *Reviews of Modern Physics* **82**, 1 (2010).
- [23] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, *Physical Review Letters* **83**, 3081 (1999).
- [24] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, *IEEE Transactions on Information Theory* **48**, 2637 (2002).
- [25] R. Raussendorf and H. J. Briegel, *Physical Review Letters* **86**, 5188 (2001).
- [26] R. Raussendorf, D. E. Browne, and H. J. Briegel, *Physical review A* **68**, 022312 (2003).
- [27] D. Gross and J. Eisert, *Physical review letters* **98**, 220503 (2007).
- [28] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest, *Nature Physics* **5**, 19 (2009).
- [29] O. Gühne and G. Tóth, *Physics Reports* **474**, 1 (2009).
- [30] N. Friis, G. Vitagliano, M. Malik, and M. Huber, *Nature Reviews Physics* **1**, 72 (2019), *arXiv:1906.10929*.
- [31] M. Horodecki, P. Horodecki, and R. Horodecki, *Physics Letters A* **223**, 1 (1996).
- [32] O. Gühne and N. Lütkenhaus, *Physical review letters* **96**, 170502 (2006).
- [33] J. S. Bell and J. S. Bell, *Speakable and unspeakable in quantum mechanics: Collected papers on quantum philosophy* (Cambridge university press, 2004).
- [34] L. Pezzé and A. Smerzi, *Physical review letters* **102**, 100401 (2009).
- [35] K. Chen and L.-A. Wu, *arXiv preprint quant-ph/0205017* (2002).
- [36] O. Rudolph, *Quantum Information Processing* **4**, 219 (2005).
- [37] P. Horodecki, *Physics Letters A* **232**, 333 (1997).
- [38] M. A. Nielsen and J. Kempe, *Physical Review Letters* **86**, 5184 (2001).
- [39] G. M. D'Ariano, M. G. Paris, and M. F. Sacchi, *Advances in Imaging and Electron Physics* **128**, 206 (2003).
- [40] M. Steffen, M. Ansmann, R. C. Bialczak, N. Katz, E. Lucero, R. McDermott, M. Neeley, E. M. Weig, A. N. Cleland, and J. M. Martinis, *Science* **313**, 1423 (2006).
- [41] F. Buscemi, M. Dall'Arno, M. Ozawa, and V. Vedral, *arXiv:1312.4240*, **1** (2013), *arXiv:1312.4240*.
- [42] F. Buscemi, M. Dall'Arno, M. Ozawa, and V. Vedral, *International Journal of Quantum Information* **12**, 1 (2014), *arXiv:1409.2237*.
- [43] H. Pashayan, J. J. Wallman, and S. D. Bartlett, *Physical Review Letters* **115**, 070501 (2015).
- [44] K. Temme, S. Bravyi, and J. M. Gambetta, *Physical Review Letters* **119**, 180509 (2017).
- [45] S. Endo, S. C. Benjamin, and Y. Li, *Physical Review X* **8**, 031027 (2018).
- [46] R. Takagi, *arXiv preprint arXiv:2006.12509* (2020).
- [47] J. Jiang, K. Wang, and X. Wang, (2020), *arXiv:2012.10959v1*, *2012.10959*.
- [48] M. B. Plenio, *Physical review letters* **95**, 090503 (2005).
- [49] R. F. Werner, *Physical Review A* **40**, 4277 (1989).
- [50] L. Gurvits, in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing* (ACM, 2003) pp. 10–19.
- [51] S. Gharibian, *arXiv arXiv:0810.4507* (2008).
- [52] A. Peres, *Physical Review Letters* **77**, 1413 (1996).
- [53] K. Kraus, *States, effects, and Operations* (Springer-Verlag, Berlin, 1983).
- [54] H. Buhrman, R. Cleve, J. Watrous, and R. De Wolf, *Physical Review Letters* **87**, 167902 (2001).
- [55] J. Kiefer, J. Wolfowitz, *et al.*, *The Annals of Mathematical Statistics* **23**, 462 (1952).
- [56] D. P. Kingma and J. Ba, *arXiv preprint arXiv:1412.6980* (2014).
- [57] W. Hoeffding, in *The Collected Works of Wassily Hoeffding* (Springer, 1994) pp. 409–426.
- [58] N. J. Cerf, C. Adami, and R. M. Gingrich, *Physical Review A* **60**, 898 (1999).
- [59] H.-P. Breuer, *Physical Review Letters* **97**, 080501 (2006).
- [60] W. Hall, *Journal of Physics A: Mathematical and General* **39**, 14119 (2006).
- [61] M.-D. Choi, *Linear Algebra and its Applications* **12**, 95 (1975).
- [62] M.-D. Choi, *Journal of Operator Theory*, 271 (1980).
- [63] E. Stormer, *Acta Mathematica* **110**, 233 (1963).
- [64] S. L. Woronowicz, *Reports on Mathematical Physics* **10**, 165 (1976).
- [65] M. M. Wilde, *Quantum Information Theory 2nd Edition* (Cambridge University Press, 2016).
- [66] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2011).

- [67] R. O'Donnell and J. Wright, in *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing* (2016) pp. 899–912.
- [68] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, *IEEE Transactions on Information Theory* **63**, 5628 (2017).
- [69] G. Vidal and R. F. Werner, *Physical Review A* **65**, 032314 (2002).
- [70] M. B. Plenio, *Physical Review Letters* **95**, 090503 (2005).
- [71] L. Chen, M. Aulbach, and M. Hajdušek, *Physical Review A* **89**, 042305 (2014).
- [72] H. Zhu, M. Hayashi, and L. Chen, *Journal of Physics A: Mathematical and Theoretical* **50**, 475303 (2017).
- [73] K. Wang, X. Wang, and M. M. Wilde, arXiv preprint arXiv:1911.07433 (2019).
- [74] M. Müller-Lennert, F. Dupuis, O. Szechr, S. Fehr, and M. Tomamichel, *Journal of Mathematical Physics* **54**, 122203 (2013).
- [75] M. M. Wilde, A. Winter, and D. Yang, *Communications in Mathematical Physics* **331**, 593 (2014).
- [76] C. Lemairechal, *Doc Math Extra* **251**, 254 (2012).
- [77] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, *Physical Review A* **98**, 032309 (2018).
- [78] P. J. Coles, S. Eidenbenz, S. Pakin, A. Adedoyin, J. Ambrosiano, P. Anisimov, W. Casper, G. Chennupati, C. Coffrin, H. Djidjev, *et al.*, arXiv, arXiv (2018).
- [79] “Baidu Quleaf Simulator,” (2020).
- [80] “Paddle Quantum,” (2020).
- [81] “<https://github.com/paddlepaddle/paddle>,” .
- [82] Y. Ma, D. Yu, T. Wu, and H. Wang, *Frontiers of Data and Computing* **1**, 105 (2019).
- [83] M. Horodecki and P. Horodecki, *Physical Review A* **59**, 4206 (1999).
- [84] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, *Physical Review Letters* **78**, 2275 (1997).
- [85] V. Vedral and M. B. Plenio, *Physical Review A* **57**, 1619 (1998).
- [86] E. M. Rains, *IEEE Transactions on Information Theory* **47**, 2921 (2001).
- [87] N. Datta, *IEEE Transactions on Information Theory* **55**, 2816 (2009).

Appendix A: Qutrit entanglement detection via the Choi map

In this Appendix, we show by example how VED can be adapted to detect the entanglement of an unknown two-qutrit state via the celebrated Choi map [61, 62]. To be specific, the Choi map is defined as follows

$$\Phi_C(\rho) := \begin{bmatrix} \rho_{1,1} + \rho_{2,2} & -\rho_{1,2} & -\rho_{1,3} \\ -\rho_{2,1} & \rho_{2,2} + \rho_{3,3} & -\rho_{2,3} \\ -\rho_{3,1} & -\rho_{3,2} & \rho_{3,3} + \rho_{1,1} \end{bmatrix}, \quad (\text{A1})$$

where $\rho_{i,j}$ is the element of ρ in the i -th row and j -th column. This map is the first known example of a positive map that is indecomposable (i.e., it cannot be decomposed into a sum of a completely positive map and a completely copositive map). The Choi map in particular can be used to detect entanglement of some PPT states (states that can not be detected by the transpose map).

For the qutrit system, the generalized Pauli matrices are de-

fined through the unitary boost and shift operators

$$X = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{bmatrix}, \quad (\text{A2})$$

where $\omega := e^{2\pi i/3}$ is the 3-th root of unity. The Choi map can be decomposed w.r.t. X and Z as follows:

$$\begin{aligned} \Phi_C(\rho) = & \frac{1}{3}[X^2 Z \rho (X^2 Z)^\dagger + X^2 Z^2 \rho (X^2 Z^2)^\dagger + X^2 \rho (X^2)^\dagger] \\ & + \frac{2}{3}[Z \rho Z^\dagger + Z^2 \rho (Z^2)^\dagger] - \frac{1}{3}\rho. \end{aligned} \quad (\text{A3})$$

Based on (A3), one could adopt the deterministic VED in Algorithm 1 or the probabilistic VED in Algorithm 2 to detect entanglement on qutrit systems by considering

$$\min_{\psi(\alpha)} \langle \psi(\alpha) | \Phi_C(\rho_{AB}) | \psi(\alpha) \rangle, \quad (\text{A4})$$

where the test state $\psi(\alpha)$ can be generated by parameterized quantum circuits.