

Variational quantum circuit learning of entanglement purification in multiple degrees of freedomHao Zhang,¹ Xusheng Xu,^{2,*} Chen Zhang,¹ Man-Hong Yung,³ Tao Huang,^{1,4,†} and Yunjie Liu^{1,4}¹*Purple Mountain Laboratories, Nanjing 211111, China*²*Department of Physics, State Key Laboratory of Low-Dimensional Quantum Physics, Tsinghua University, Beijing 100084, China*³*Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*⁴*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*

(Received 21 September 2022; revised 20 June 2023; accepted 20 September 2023; published 16 October 2023)

Entanglement purification is a crucial technique for promising the effective entanglement channel in noisy large-scale quantum networks, yet complicated in designing protocols in multi-degree of freedom (DoF). To execute the above tasks easily and effectively, developing a learning framework for designing the entanglement purification with multi-DoF is a promising way and still an open research question. Inspired by variational quantum circuit (VQC) with remarkable advantage in learning optimal quantum operations with near-term quantum devices, in this paper we propose an effective VQC framework for the entanglement purification in multi-DoF and exploit it to learn the optimal purification protocols of the objective function which are based on postselection. By properly introducing additional circuit lines for representing each of the ancillary DoFs of all the particles, e.g., space and time, the parametrized quantum circuit can effectively simulate scalable entanglement purification. To verify our framework, the well-known protocols in linear optics are learned well with alternative operations in low-depth quantum circuit. Moreover, we simulate the multipair cases with multi-DoF to show the scalability and discover one-round protocols. Our work provides an effective way for exploring the entanglement purification protocols in multi-DoF and multipair with near-term quantum devices.

DOI: [10.1103/PhysRevA.108.042611](https://doi.org/10.1103/PhysRevA.108.042611)**I. INTRODUCTION**

Quantum entanglement which shows the nonlocal correlation between two or more objects is an intriguing phenomenon in quantum mechanics and has no classical counterpart [1]. One usually uses quantum entanglement as a crucial resource for building quantum channel in quantum networks [2–6]. However, in practice, entanglement is so fragile in a noisy environment that it is hard to be used directly as an effective quantum channel. The reason is that under the influence of noise, a pure maximally entangled state becomes a mixed one. To overcome this problem, a technique called entanglement purification is proposed to improve the fidelity of the damaged entangled state [7–28]. The first entanglement purification protocol utilizes another copy of mixed entangled state in Werner form as an auxiliary “target” state and executes bilateral controlled-NOT (CNOT) operations and parity check to acquire the information of “source” pair [7]. Subsequently, the protocol is developed without requirement of Werner form and has higher efficiency in recursive procedure [8]. The above protocols are based on CNOT gates between two entangled pairs. It is hard to accomplish this operation in experiment, especially for photons. Therefore, in optical systems, the feasible way is to bring in ancillary photonic degrees of freedom (DoF), such as space and time. The first photonic entanglement purification protocol which

makes use of ancillary DoF is based on selecting the spatial modes of entangled pairs [10] and subsequently performed in experiment [19]. Besides, one can design entanglement purification protocols with only one pair of photons with multi-DoF hyperentanglement [11–16, 23–26], e.g., Simon-Pan [11] and Hu-Huang-Sheng-Zhou *et al.* (HHSZ+) [23] protocols. Up to now, many interesting purification protocols have been proposed for various cases but lack of a learning framework. As the number of entangled pair and DoF increases, the design of entanglement purification becomes more complicated and challenging.

In recent years, machine learning has been considered for processing quantum information [29–32]. Some basic protocols in quantum communication, such as quantum teleportation [33], entanglement purification, and quantum repeaters [34], are discovered by classical agents [35]. Quantum gate operations play a key role in quantum information processing. Therefore, compared with classical machine learning, an approach directly optimizing quantum gates called variational quantum circuit (VQC) has its inherent advantage for handling quantum information tasks [36–38]. The VQC learning has been widely applied in various areas, such as quantum computing and quantum chemistry, for its remarkable advantage in training with near-term quantum devices. This inspires that VQC learning may be an effective way to design the protocols of entanglement purification in multi-DoF. Recently, entanglement purification has been performed as a simple instance in local operation and classical communication framework based on parametrized quantum circuits [39]. However, directly simulating operations on two

*thuxxs@163.com

†htao@bupt.edu.cn

pairs is limited in some cases, such as the difficult CNOT operation of photons. Therefore, developing an effective VQC learning framework for entanglement purification including multi-DoF has practical significance but there is still no research.

In this paper, we propose an effective VQC learning framework for the entanglement purification in multi-DoF and exploit it to learn the optimal purification protocols of the objective function which is based on postselection. Additional quantum circuit lines are introduced in VQC to represent the high-dimensional DoFs of the particle. In our VQC learning framework, the parametrized ansatz part plays the central role in learning the local quantum operations of entanglement purification. The classical communications used for exchanging the information between two users are included in a measurement part. As examples, the well-known linear optical entanglement purification protocols, including Pan-Simon-Brukner-Zeilinger (PSBZ) [10], HHSZ+ [23], Simon-Pan [11], etc., are learned well and the different operations of entanglement purification are discovered automatically. To verify the scalability of our framework, the cases of multipair with multi-DoF are also performed well and the results indicate that multipair entangled states can be purified by a one-round way that is different from the conventional recursive process. Moreover, entanglement purification is also simulated with different noisy channels. Our framework provides an alternative way to understand and design the entanglement purification in multi-DoF by variational quantum learning and has extensive applications for the other areas of quantum information, such as quantum networks.

The article is organized as follows: In Sec. II, we introduce our VQC learning framework for the entanglement purification in multi-DoF. In Sec. III, the well-known protocols of the entanglement purification with multi-DoF are learned. In Sec. IV, the cases of multipair with multi-DoF are verified. In Sec. V, the noisy channels are considered in simulations. Section VI covers discussion and summary.

II. VQC FRAMEWORK FOR THE ENTANGLEMENT PURIFICATION IN MULTI-DOF

A pure entangled state labeled with ρ_{pure} will become a mixed one ρ_{mixed} when it is distributed over a noisy quantum channel. However, entanglement purification is an inverse process which can improve the fidelity of a mixed state. Those two processes can be described by the formula

$$\rho_{\text{pure}} \xrightarrow{\text{noise}} \rho_{\text{mixed}} \xrightarrow{\text{EP}} \rho_{\text{purified}}. \quad (1)$$

Here, EP is entanglement purification, and ρ_{purified} is the density matrix of a purified mixed state and can equal to ρ_{pure} after a perfect entanglement purification protocol. In long-distance quantum channel, an entangled state is usually shared by two nonlocal quantum nodes shown in the upper part of Fig. 1(a). Because of nonlocality, the two users are only allowed to execute entanglement purification using local operations and classical communications. The local operations are a series of local quantum gates, and the classical communications are used for exchanging the measurement results in two nodes. The variational quantum algorithm

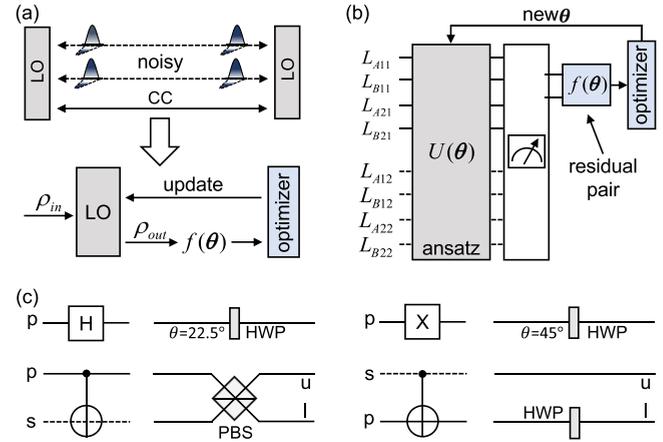


FIG. 1. The VQC learning framework of the entanglement purification with multi-DoF. (a) The simple schematic diagrams of the entanglement purification in optical systems and VQC. LO, local operations; CC, classical communications. (b) An instance for the VQC of the entanglement purification with polarization and spatial DoFs. (c) The physical implementations of some basic quantum gates in linear optics. The meanings of symbols are p (s), polarization (spatial) DoF; u (l), upper (lower) spatial DoF; H, Hadamard gate; X, Pauli-X gate; HWP, half-wave plate; PBS, polarizing beam splitter.

shown in lower part of Fig. 1(a) is a hybrid quantum-classical simulator using classical optimizer to optimize parametrized quantum circuits. Here, we propose a general VQC framework shown in Fig. 1(b) to simulate the entanglement purification with multi-DoF. Its main architecture includes ansatz, measurement and classical optimizer. In conventional entanglement purification protocols, local operations usually have two parts arranged before and after measurement, we assume the ansatz in our VQC framework expresses only the part before measurement for simplicity. The ansatz is the main concern in learning entanglement purification and prepared with parametrized quantum circuits based on practical conditions. The gate operations in the ansatz between circuit lines of Alice and Bob are forbidden because of only allowing local quantum operations. By optimizing the parametrized quantum circuits in the ansatz, the process of local operations in entanglement purification can be simulated effectively. Each circuit line represents one DoF of entangled particle. We assume that Alice and Bob share n pairs of entangled photons denoted by sets $A = \{A_1, A_2, \dots, A_n\}$ and $B = \{B_1, B_2, \dots, B_n\}$, respectively. A_i and B_i are entangled pairs. The photons with m DoFs in the sets A and B are also described as sets $A_i = \{D_{Ai1}, D_{Ai2}, \dots, D_{Aim}\}$ and $B_i = \{D_{Bi1}, D_{Bi2}, \dots, D_{Bim}\}$, and the elements D_{Aij} and D_{Bij} stand for the different DoFs of photons, such as polarization, space, and time. Preparing the quantum circuits in the VQC for entanglement purification, the circuit lines arranged for Alice and Bob's photons are represented with sets $C_A = \{C_{A1}, C_{A2}, \dots, C_{An}\}$ and $C_B = \{C_{B1}, C_{B2}, \dots, C_{Bn}\}$, respectively. All the elements in the sets C_{Ai} and C_{Bi} contain the circuit lines for the DoFs of photons and expressed by $C_{Ai} = \{L_{Ai1}, L_{Ai2}, \dots, L_{Aim}\}$ and $C_{Bi} = \{L_{Bi1}, L_{Bi2}, \dots, L_{Bim}\}$. The construction of the circuit lines in VQC is actually a map from entangled photons to lines, i.e., $f: D_{Aij}(D_{Bij}) \mapsto L_{Aij}(L_{Bij})$. The dimension of each line

L_{Aij} , L_{Bij} is decided by D_{Aij} , D_{Bij} and has the relation of $\dim(L_{Aij}) = \dim(D_{Aij})$, $\dim(L_{Bij}) = \dim(D_{Bij})$. An instance of photons in which two entangled pairs with two DoFs shared by Alice and Bob is given in Fig. 1(b), the horizontal solid and dashed lines represent photonic polarization and spatial DoFs, respectively. The circuit lines $C_{A1} = \{L_{A11}, L_{A12}\}$, $C_{A2} = \{L_{A21}, L_{A22}\}$ belong to Alice and $C_{B1} = \{L_{B11}, L_{B12}\}$, $C_{B2} = \{L_{B21}, L_{B22}\}$ belong to Bob. The C_{A1} , C_{B1} are nonlocal entangled pair, and the C_{A2} , C_{B2} are another entangled one. The polarization has horizontal (H) and vertical (V) directions, and the spatial DoF is assumed with only upper and lower paths. Therefore, the dimension of DoF is $\dim(L_{Aij}) = \dim(L_{Bij}) = 2$. In entanglement purification, the goal is to obtain the output state with higher fidelity. The target state $|\psi_{\text{target}}\rangle$ usually is a pure state and the learning objective function is chosen with fidelity $f_{\text{out}} = \langle \psi_{\text{target}} | \rho_{\text{out}} | \psi_{\text{target}} \rangle$ whose direction of optimization is maximization. Here, ρ_{out} is the density operator of residual entangled pairs after entanglement purification, i.e., depending on $U(\theta)\rho_{\text{in}}U^\dagger(\theta)$ and postselection. Therefore, when the measurement outcome is chosen, the goal is to train the parameter vector θ . Figure 1(c) shows the instances of the physical implementations of several basic quantum gates on polarization and spatial DoFs in linear optics. The half-wave plate can act as the Hadamard or Pauli-X gate of photonic polarization DoF with different input angles. The polarizing beam splitter (PBS) which reflects V and transmits H polarization of photon realizes a CNOT gate on photonic polarization (source) and spatial (target) DoFs. A reversed CNOT between polarization (target) and spatial (source) DoFs can be realized by adding a Pauli-X gate (a half-wave plate) in the lower path.

III. LEARNING THE ENTANGLEMENT PURIFICATION WITH MULTI-DOF IN LINEAR OPTICS

A. PSBZ protocol

Four Bell states are considered in entanglement purification as follows:

$$\begin{aligned}
 |\Phi^\pm\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \\
 |\Psi^\pm\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle).
 \end{aligned} \quad (2)$$

To learn entanglement purification by our VQC framework, we first study the PSBZ protocol proposed by Pan *et al.* [10] for linear optics. In the protocol, photonic spatial DoF is introduced to overcoming the problem of difficult CNOT operations between photons. The ideal case is that Alice and Bob share Bell pairs $|\Phi_{ab}^+\rangle = \frac{1}{\sqrt{2}}(|0_a0_b\rangle \pm |1_a1_b\rangle)$ in polarization DoF from ideal source. Here, the states 0 and 1 in $|\Phi_{ab}^+\rangle$ represent for the V and H polarization of photon, respectively. The photons labeled with “a” and “b” belong to Alice and Bob, respectively. A mixed state considered with only bit-flip error before entanglement purification is given by

$$\rho_{\text{in}}^{ab} = f_{\text{in}}|\Phi_{ab}^+\rangle\langle\Phi_{ab}^+| + (1 - f_{\text{in}})|\Psi_{ab}^+\rangle\langle\Psi_{ab}^+|. \quad (3)$$

The two copies of this mixed state $\rho_{\text{in}}^{a1b1} \otimes \rho_{\text{in}}^{a2b2}$ have four components, including $|\Phi_{a1b1}^+\rangle|\Phi_{a2b2}^+\rangle$ with fidelity f_{in}^2 , $|\Phi_{a1b1}^+\rangle|\Psi_{a2b2}^+\rangle$ with fidelity $f_{\text{in}}(1 - f_{\text{in}})$, $|\Psi_{a1b1}^+\rangle|\Phi_{a2b2}^+\rangle$ with

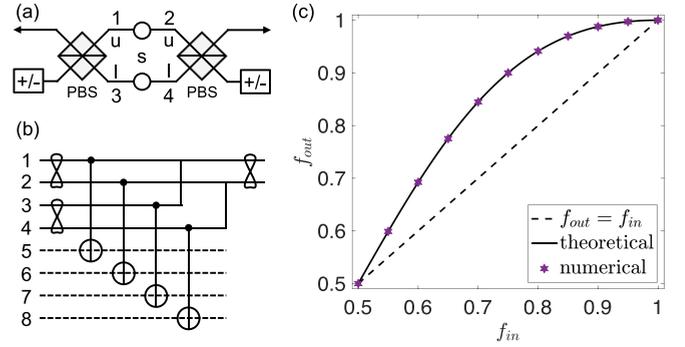


FIG. 2. The VQC learning of the PSBZ protocol. (a) The physical schematic diagrams of entanglement purification. “s” represents ideal entanglement source. “u” and “l” are upper and lower spatial DoF, respectively. (b) The quantum circuit version of PSBZ protocol. (c) The numerical and theoretical results of fidelities.

fidelity $f_{\text{in}}(1 - f_{\text{in}})$, and $|\Psi_{a1b1}^+\rangle|\Psi_{a2b2}^+\rangle$ with fidelity $(1 - f_{\text{in}})^2$. Different with protocols [7] and [8], the PSBZ protocol replaces the CNOT between two photons with the CNOT between the two DoFs of each photon. As shown in Fig. 2(a), two symmetric PBSs are used by both Alice and Bob. The whole process is described with quantum circuit language in Fig. 2(b). For simplicity of calculation, we label Alice and Bob’s photons with 1, 3, 5, 7 and 2, 4, 6, 8, respectively. Circuit lines 1 (3) and 2 (4) are polarization entangled. All the circuit lines 5, 6, 7, and 8 are spatial DoF and plotted together. The initial state of spatial DoF is $|0_50_61_71_8\rangle$ and 0 (1) stands for the upper (lower) path. The two PBSs play roles in applying four CNOT gates between the polarization and spatial DoFs of each photon. Actually, the CNOT gates on all spatial DoF produce the two pairs of four-qubit entangled state. Alice and Bob should measure their lower spatial DoF with basis $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and choose the case of all output with photons, i.e., so-called “four-mode cases” [19]. The detailed derivation process of the above protocol is given in Appendix A. If Alice and Bob get the result $|\text{++}\rangle$ or $|\text{--}\rangle$, the state of source pair is

$$\rho_{\text{upper}}^{\pm\pm} = f_{\text{out}}^{\pm\pm}|\Phi^+\rangle\langle\Phi^+| + (1 - f_{\text{out}}^{\pm\pm})|\Psi^+\rangle\langle\Psi^+|, \quad (4)$$

where the new fidelity is $f_{\text{out}}^{\pm\pm} = \frac{f_{\text{in}}^2}{f_{\text{in}}^2 + (1 - f_{\text{in}})^2}$. When the result is $|\text{+--}\rangle$ or $|\text{-+-}\rangle$, the output state will be translated to $|\Phi^+\rangle$, whose new fidelity is $f_{\text{out}}^{\pm\mp} = f_{\text{out}}^{\pm\pm}$, by applying a local phase flip gate on the one of the residual photons.

Using the VQC framework to learn the above entanglement purification protocol, we directly learn the ansatz by assuming a series of parametrized universal quantum gates including single-qubit arbitrary rotation gates and two-qubit CNOT gates. The input of ansatz is $\rho_{\text{in}} = \rho_{a1b1} \otimes \rho_{a2b2} \otimes |0_50_61_71_8\rangle\langle 0_50_61_71_8|$. Our goal is to optimize the fidelity of the final output state given by objective function $f_{\text{out}} = \langle\Phi^+|\rho_{\text{out}}|\Phi^+\rangle$ and the ρ_{out} is chosen with $\rho_{\text{out}} = \rho_{\text{upper}}^{\pm\pm}$ in our simulations. Numerical results are shown in Fig. 2(c). The points in the figure are our numerical fidelities and match the theoretical curve $f_{\text{out}}^{\pm\pm}$ very well. Each point shown in the figure is the best one chosen from learning results by

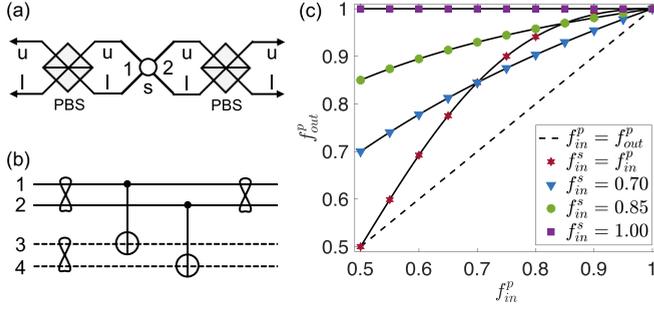


FIG. 3. The VQC learning of HHSZ+ and Simon-Pan protocols with hyperentanglement. (a) The physical schematic diagrams of the two protocols. (b) The quantum circuit version of the two protocols. Lines 1 (2) and 3 (4) are polarization and spatial DoFs of photon belonging to Alice (Bob), respectively. (c) The numerical and theoretical results of fidelities. The solid curves and markers are theoretical and numerical results, respectively.

randomly initializing the ansatz with 10 times as the one time learning might give a local optimal fidelity rather than global optimal one. The learned ansatz suggests the optimal fidelities of entanglement purification are the same as the PSBZ protocol in Fig. 2(b) for the input mixed state ρ_{in} . The learned local operations are not unique due to the learning of fidelity.

B. Hyperentanglement-based protocols

Another kind of entanglement purification protocols using multi-DoF is based on hyperentanglement. Only one pair of photons entangled in both polarization and spatial DoFs are required. The Bell states of polarization and spatial DoFs have the same form with Eq. (2) and are, respectively, labeled with $|\Phi_p^\pm\rangle$, $|\Psi_p^\pm\rangle$, $|\Phi_s^\pm\rangle$, and $|\Psi_s^\pm\rangle$. Subscripts “ p ” and “ s ” represent polarization and spatial DoFs, respectively. The typical entanglement purification protocols based on hyperentanglement are HHSZ+ [23], Simon-Pan [11], Li [12], Sheng-Deng [13], etc. We first study the HHSZ+ protocol whose initial state is described by $\rho_{in} = \rho_{in}^p \otimes \rho_{in}^s$, where the density operators ρ_{in}^p and ρ_{in}^s are

$$\rho_{in}^p = f_{in}^p |\Phi_p^+\rangle\langle\Phi_p^+| + (1 - f_{in}^p) |\Psi_p^+\rangle\langle\Psi_p^+| \quad (5)$$

and

$$\rho_{in}^s = f_{in}^s |\Phi_s^+\rangle\langle\Phi_s^+| + (1 - f_{in}^s) |\Psi_s^+\rangle\langle\Psi_s^+|. \quad (6)$$

Here, f_{in}^p and f_{in}^s are the fidelities of $|\Phi_p^+\rangle$ and $|\Phi_s^+\rangle$, respectively. The equivalent entanglement purification using time DoF is protocol [24]. Here, only one entangled pair and two PBSs are required for accomplishing the entanglement purification of the polarization DoF shown in Fig. 3(a) when the input fidelities satisfy $f_{in}^p > \frac{1}{2}$ and $f_{in}^s > \frac{1}{2}$. In the VQC framework, two polarization and another two spatial lines are introduced in Fig. 3(b). Compared with the PSBZ protocol, the physical devices are the same but the quantum circuit only has two CNOT gates in the HHSZ+ protocol. Shown in Appendix B, the output fidelity is $f_{out}^p = \frac{f_{in}^p f_{in}^s}{f_{in}^p f_{in}^s + (1 - f_{in}^p)(1 - f_{in}^s)}$ for the selection of $|0_3 0_4\rangle$ (two upper) or $|1_3 1_4\rangle$ (two lower). Three cases are considered as $f_{in}^p = f_{in}^s$, fixed $f_{in}^s \in (0.5, 1.0)$ and $f_{in}^s = 1$ in numerical simulations. For the fixed $f_{in}^s \in (0.5, 1.0)$, without loss of generality, we learn the output

fidelities by fixing the f_{in}^s with 0.70 and 0.85. When the fidelity satisfies $f_{in}^s = 1$, HHSZ+ protocol transforms to the Simon-Pan protocol. As shown in Fig. 3(c), the curve of $f_{in}^p = f_{in}^s$ is the same with the PSBZ protocol. Each point shown in the figure is also the best one chosen from learning results by random initializations with 10 times for avoiding local optimum. For instance, the case with input fidelities $f_{in}^s = 0.70$ and $f_{in}^p = 0.55$ is executed with given ansatz, the optimal output fidelity is $f_{out}^p = 0.7404$. But there exists a local optimal point $f_{out}^p = 0.70$ which suggests a swap operation to exchange the states between polarization and spatial DoFs. This operation will obtain a higher fidelity in polarization DoF, i.e., $f_{out}^p = 0.70 > f_{in}^p = 0.55$, but is not a global optimum. In $f_{in}^s = 1$, the output purified state is a pure entangled state. Results in Fig. 3(c) indicate all the assumed cases are discovered well by our VQC framework.

If we study the case considering phase errors, terms $|\Phi^-\rangle\langle\Phi^-|$ and $|\Psi^-\rangle\langle\Psi^-|$ are added in density operators. When spatial DoF is a pure entangled state, the corresponding protocols are the Li [12] and Sheng-Deng [13] protocols. As an instance, we simulate the Li protocol here. In Li’s protocol, all the output states appearing in spatial DoF $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ are $|\Phi_p^+\rangle$. To keep the same target with Li’s protocol, we use a new objective function with $f_{obj}^p = f_{00}^p * f_{01}^p * f_{10}^p * f_{11}^p$, where f_{00}^p , f_{01}^p , f_{10}^p , and f_{11}^p are the fidelities of output states in the spatial DoF $|00\rangle$ (two upper), $|01\rangle$ (one upper one lower), $|10\rangle$ (one upper one lower), and $|11\rangle$ (two lower), respectively. One can find that when the objective function f_{obj}^p is optimized to 1, all the four fidelities f_{00}^p , f_{01}^p , f_{10}^p , f_{11}^p are 1. The detailed explanations of the Li protocol are given in Appendix C. Compared with the Simon-Pan protocol, the operations in this case need three bilateral CNOT gates in Alice and Bob. The physical implementation and its corresponding quantum circuit of the Li protocol are plotted in Figs. 4(a) and 4(b), respectively. By the VQC learning, Li protocol is also learned well. Moreover, the results of optimized parameters in the ansatz also suggest some other different local operations whose final outputs are the same with those of Li. We show one of the equivalent protocols discovered by the VQC learning in Figs. 4(c) and 4(d) with quantum circuit and setup versions where four PBSs and two half-wave plates are used. This simple result can be learned by adaptively reducing and adjusting the gate operations in the ansatz. The learning curves of fidelities chosen with one of the best results from random initializations 10 times for each point are shown in Fig. 4(e). Here, the initial input fidelities of each state are assumed with $f_{in, \Phi^+}^p = (1 - f_{in, \Phi^+}^p)/2$, $f_{in, \Psi^+}^p = (1 - f_{in, \Phi^+}^p)/4$, $f_{in, \Psi^-}^p = (1 - f_{in, \Phi^+}^p)/4$, and $f_{in, \Phi^+}^s = 1$. With about 30 steps (actual steps not shown in the figure, breaking up the optimizing processes is usually more than 35 steps), all the 9 learning processes, i.e., input with $f_{in, \Phi^+}^p = \{0.55, 0.60, \dots, 0.90, 0.95\}$, achieve their optimal values.

IV. ENTANGLEMENT PURIFICATION OF MULTIPAIR AND MULTI-DOF INPUT

To verify the scalability of the framework, we simulate the entanglement purification of mutipair cases with multi-DoF and show the results of two pair (4 qubits), three pair (6

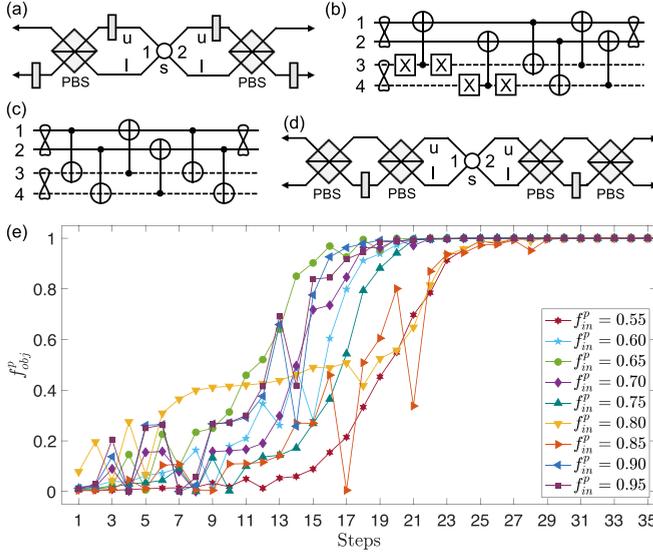


FIG. 4. The VQC learning of the Li protocol with hyperentanglement. (a) The physical schematic diagrams of the Li protocol. (b) The corresponding quantum circuit of the Li protocol. Lines 1 (2) and 3 (4) are polarization and spatial DoFs of photon belonging to Alice (Bob), respectively. (c) The quantum circuit of an example of new equivalent entanglement purification protocols learned by VQC. (d) The physical schematic diagrams of the new equivalent entanglement purification circuit in (c). (e) The learning curves of chosen fidelities based on VQC with a set of random initial parameters.

qubits), four pair (8 qubits), and five pair (10 qubits). As shown in Fig. 5(a), the model given by the case of five entangled pairs (10 qubits) is chosen with applying the two-qubit ansatz between nearest-neighbor qubits for simplicity, and one ancillary DoF is considered. The initial input entangled states are assumed as

$$\rho_{in} = \bigotimes_{i=1}^n \rho_{in}^i, \quad (7)$$

where n is considered with $n = 2, 3, 4, 5$ corresponding to 4, 6, 8, 10 qubits, and for simplicity, all the ρ_{in}^i have the same fidelities as

$$\rho_{in}^i = f_{in} |\Phi_i^+\rangle\langle\Phi_i^+| + (1 - f_{in}) |\Psi_i^+\rangle\langle\Psi_i^+|. \quad (8)$$

We assume that the residual pair is kept when we chose the case of $|0\rangle^{\otimes 2(n-1)}$ or $|1\rangle^{\otimes 2(n-1)}$ of the $n - 1$ nonlocal entangled pairs. The output density matrix calculated in objective function is chosen with $\rho_{out} = \rho_{out}^1 + \rho_{out}^0$, where ρ_{out}^1 (ρ_{out}^0) is residual state corresponding to measurement outcome $|0\rangle^{\otimes 2(n-1)}$ ($|1\rangle^{\otimes 2(n-1)}$). In numerical simulations, input fidelities $\{0.5, 0.55, 0.6, \dots, 0.95, 1.0\}$ are performed. For each point of the input fidelities, we also simulate it for 10 times by initializing the parameters of ansatz randomly and choose the best fidelity as the final output result rather than expectation. For the case of five pair (10 qubits), when the output optimal numerical result is larger than the optimal fidelity of the four-pair case, the simulation is ended at this round for saving time. All the purification results shown in Fig. 5(b) indicate that the output fidelity can be increased by using more input pairs in domain $f_{in} \in (0.5, 1.0)$. For instance of $f_{in} = 0.65$ shown in

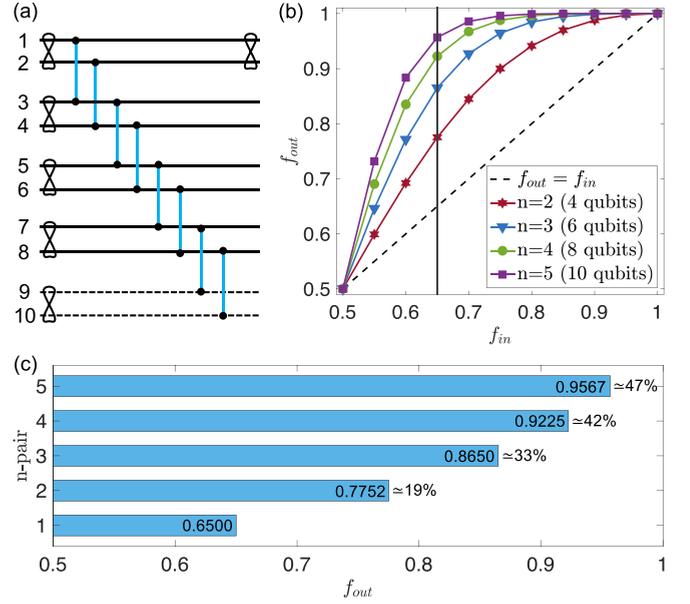


FIG. 5. Entanglement purification for the cases of multipair with multi-DoF. (a) The purification model and ansatz structure of five entangled pairs (10 qubits). The connection labeled with a vertical line and two dots between two circuit lines is two-qubit ansatz. (b) The learning fidelities of all cases from two to five input entangled pairs. (c) The learning fidelities of $f_{in} = 0.65$ from two to five input entangled pairs.

Fig. 5(c), the fidelity is purified to $f_{out} \simeq 0.7752$ [purification rate $pr = (f_{out} - f_{in})/f_{in} \simeq 19\%$] in the two-pair case and this result can be further enhanced to $f_{out} \simeq 0.8650$ ($pr \simeq 33\%$), $f_{out} \simeq 0.9225$ ($pr \simeq 42\%$), and $f_{out} \simeq 0.9567$ ($pr \simeq 47\%$) in the cases of three pair, four pair, and five pair, respectively. Our calculations indicate that the entanglement purification with multipair cases can be carried out with only one round. In conventional protocols, the purifications of multipair mixed entangled states are usually based on a recursive process in which two-pair case is carried out in each round. For the instance of four entangled pairs' input, the conventional recursive process is two rounds described by 4 pairs \rightarrow 2 pairs \rightarrow 1 pairs, but this process is only one round in our protocol, i.e., 4 pairs \rightarrow 1 pairs.

V. ENTANGLEMENT PURIFICATION WITH NOISE

To simulate practical cases, we consider the three types of noisy channels, i.e., phase-damping, amplitude-damping, and symmetric depolarizing channels, in our simulations of entanglement purification. The noisy channels are described by applying Kraus operators E_k to density matrix by conjugation as follows:

$$\rho = \mathbb{N}(\rho) = \sum_k E_k \rho E_k^\dagger. \quad (9)$$

The Kraus operators of phase-damping channel \mathbb{N}_{pd} are given by

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{bmatrix}. \quad (10)$$

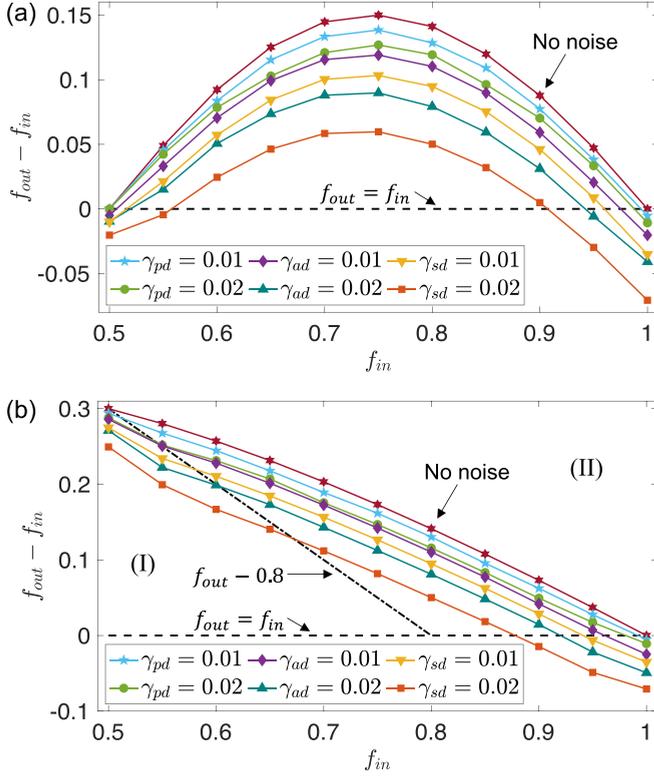


FIG. 6. The simulation of entanglement purification with noise. (a) The results of the case with the input entangled states $f_{in}^i = f_{in}^j = f_{in}$. (b) The results of the case with the input entangled states $f_{in}^i = f_{in}^j$ and $f_{in}^j = 0.80$.

For amplitude-damping channel \mathbb{N}_{ad} , the Kraus operators has form of

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, E_2 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}. \quad (11)$$

Symmetric depolarizing channel is described by a set of Pauli operators as

$$\mathbb{N}_{sd}(\rho) = \left(1 - \frac{3\gamma}{4}\right)\rho + \frac{\gamma}{4}(\sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z). \quad (12)$$

The parameter γ in the above expressions stands for the magnitude of the noise.

Two-qubit gates have higher noise rate than one-qubit ones in physical considerations [40]; for simplicity, the noise is applied only after CNOT gates in our simulations. The two cases of initial input fidelities, i.e., $f_{in}^i = f_{in}^j$ and $f_{in}^i = f_{in}^j$ with fixed $f_{in}^j = 0.80$ are simulated by considering noisy channel with $\gamma = 0, 0.01, 0.02$. The learning results of entanglement purification with noise are shown in Figs. 6(a) and 6(b). The measurement choices of all types of noise channel are the same, i.e., choosing $|00\rangle$ or $|11\rangle$ of ancillary DoF, and each data point shown in the figure is the best result of random initializations with 10 times. The objective function is chosen with fidelity $f_{out} = \langle \Phi^+ | (\rho_{out}^{00} + \rho_{out}^{11}) | \Phi^+ \rangle$, where ρ_{out}^{00} (ρ_{out}^{11}) is the output state of measurement outcome 00 (11). γ_{pd} , γ_{ad} , and γ_{sd} represent the magnitude of the noise of the phase-damping, amplitude-damping, and symmetric depolarizing channels,

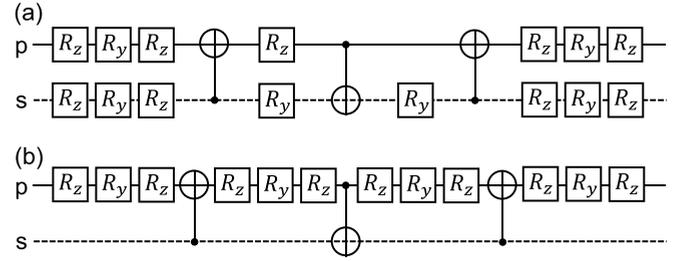


FIG. 7. The quantum circuits of the ansatz used in our simulations between two different DoFs. (a) A universal quantum circuit for two-qubit operations. (b) A limited ansatz with no single-qubit operations in spatial DoF. R_y and R_z are single-qubit rotation gates about the y and z axes, respectively.

respectively. In the case of $f_{in}^i = f_{in}^j$, we calculate the result of $f_{out} - f_{in}$ called net fidelity. While in the case of fixed $f_{in}^j = 0.8$, we choose the net fidelity $f_{out} - f_{in}$ in which f_{in} is chosen with the modulated input fidelity f_{in}^i rather than the fixed one f_{in}^j . The results in the two figures show that noise reduces the quality of entanglement purification. As the noise (γ) increases, the net fidelities decrease. If the net fidelity is less than zero where the area below dashed line is $f_{out} = f_{in}$ in Fig. 6(a), it indicates the purification fails. In Fig. 6(b), the case is little different: we give two dashed lines, i.e., the line $f_{out} = f_{in}$ and $f_{out} = 0.8$. In some cases of multi-DoF purification, for instance in linear optics, one usually uses ancillary DoF, i.e., path, to purify the primary one, i.e., polarization. If the output fidelity of polarization is greater than the input, the purification is valid although the output fidelity is less than the input fidelity of spatial DoF. We give a dashed line $f_{out} = f_{in}$ as the critical line for the above case and a line $f_{out} = 0.8$ for a common case whose requirement is f_{out} greater than all the input fidelities. So in our simulations, the second dashed line is $f_{out} = 0.8$ as the fidelity of input ancillary DoF is 0.8. In Fig. 6(b), we mark two valid areas I + II and II for the requirements of $f_{out} > f_{in}$ and $f_{out} > f_{in}$ and 0.8, respectively. Comparing three types of noise, the phase-damping and symmetric depolarizing channels have the smallest and biggest influence on entanglement purification, respectively.

VI. DISCUSSION AND SUMMARY

The circuit of the ansatz with parameters $\theta_u = [\theta_1, \theta_2, \dots, \theta_{15}]$ used in our simulations is given in Fig. 7(a). It is a universal quantum circuit for two-qubit operations with a series of single-qubit operations and three CNOT gates [41,42]. In some practical cases, the operations can not be realized with experimental conditions. Therefore, we introduce a limited ansatz with $\theta_l = [\theta_1, \theta_2, \dots, \theta_{12}]$ by assuming the single-qubit operations are difficult in the spatial DoF in Fig. 7(b). In practical numerical simulations, we learned the optimal fidelities by using a universal ansatz and also can repeat the process with a limited one according to experimental conditions if necessary. The parameter vector θ_u in the ansatz is assigned with random numbers in domain $[0, 2\pi]$, so the learning curves are usually different in each round, but the final stationary results of fidelities are almost the same (local optimum

can be detoured by repeating learning process more times). Our numerical simulations are mainly performed by using MINDQUANTUM [43]. The optimizer is chosen with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) [44–47].

In conclusion, we have proposed an effective VQC framework for learning the entanglement purification with multi-DoF by properly introducing additional circuit lines. To verify our framework, the well-known protocols of entanglement purification in linear optics, such as PSBZ, HHSZ+, Simon-Pan, etc., are learned with alternative operations and the fidelities match well with the theoretical values. To show the scalability, we simulate the cases of multipair with multi-DoF well and show that the entanglement purification of multipair can be carried out with only one round. Moreover, the influence of noise is performed. The results indicate that our VQC learning method is effective in designing the protocols of the entanglement purification in multi-DoF. Our work introduces an effective way to design entanglement purification by variational quantum learning with near-term quantum devices.

APPENDIX A: PSBZ PROTOCOL IN QUANTUM CIRCUIT LANGUAGE

The density operator of an initial mixed state is described as

$$\rho_{12} = \rho_{34} = f_{\text{in}} |\Phi_{ab}^+\rangle \langle \Phi_{ab}^+| + (1 - f_{\text{in}}) |\Psi_{ab}^+\rangle \langle \Psi_{ab}^+|. \quad (\text{A1})$$

Therefore, the system composed of two entangled pairs has a density operator given by

$$\begin{aligned} \rho_{1-4} &= \rho_{12} \otimes \rho_{34} \\ &= [f_{\text{in}}^2 |\Phi_{12}^+\rangle \langle \Phi_{12}^+| \otimes |\Phi_{34}^+\rangle \langle \Phi_{34}^+| \\ &\quad + f_{\text{in}}(1 - f_{\text{in}}) |\Phi_{12}^+\rangle \langle \Phi_{12}^+| \otimes |\Psi_{34}^+\rangle \langle \Psi_{34}^+| \\ &\quad + f_{\text{in}}(1 - f_{\text{in}}) |\Psi_{12}^+\rangle \langle \Psi_{12}^+| \otimes |\Phi_{34}^+\rangle \langle \Phi_{34}^+| \\ &\quad + (1 - f_{\text{in}})^2 |\Psi_{12}^+\rangle \langle \Psi_{12}^+| \otimes |\Psi_{34}^+\rangle \langle \Psi_{34}^+|]. \quad (\text{A2}) \end{aligned}$$

The state of spatial DoF at the beginning is a product state expressed with $\rho_{5-8} = |0_5 0_6 1_7 1_8\rangle \langle 0_5 0_6 1_7 1_8|$, where 0 and 1 stand for upper and lower paths, respectively. The density operator of the whole system can be written as

$$\begin{aligned} \rho_{1-8} &= \rho_{1-4} \otimes \rho_{5-8} \\ &= f_1 (|0_1 0_2 0_3 0_4 0_5 0_6 1_7 1_8\rangle + |1_1 1_2 0_3 0_4 0_5 0_6 1_7 1_8\rangle \\ &\quad + |0_1 0_2 1_3 1_4 0_5 0_6 1_7 1_8\rangle + |1_1 1_2 1_3 1_4 0_5 0_6 1_7 1_8\rangle) \langle \dots | \\ &\quad + f_2 (|0_1 0_2 0_3 1_4 0_5 0_6 1_7 0_8\rangle + |1_1 1_2 0_3 1_4 0_5 0_6 1_7 0_8\rangle \\ &\quad + |0_1 0_2 1_3 0_4 0_5 0_6 1_7 0_8\rangle + |1_1 1_2 1_3 0_4 0_5 0_6 1_7 0_8\rangle) \langle \dots | \\ &\quad + f_2 (|0_1 1_2 0_3 0_4 0_5 0_6 1_7 1_8\rangle + |1_1 0_2 0_3 0_4 0_5 0_6 1_7 1_8\rangle \\ &\quad + |0_1 1_2 1_3 1_4 0_5 0_6 1_7 1_8\rangle + |1_1 0_2 1_3 1_4 0_5 0_6 1_7 1_8\rangle) \langle \dots | \\ &\quad + f_3 (|0_1 1_2 0_3 1_4 0_5 0_6 1_7 0_8\rangle + |1_1 0_2 0_3 1_4 0_5 0_6 1_7 0_8\rangle \\ &\quad + |0_1 1_2 1_3 0_4 0_5 0_6 1_7 0_8\rangle + |1_1 0_2 1_3 0_4 0_5 0_6 1_7 0_8\rangle) \langle \dots |. \end{aligned}$$

Here, the symbols $\langle \dots |$ mean they are the bras of corresponding kets in their left brackets. The fidelities f_1 , f_2 , and f_3 are $f_1 = f_{\text{in}}^2$, $f_2 = f_{\text{in}}(1 - f_{\text{in}})$, and $f_3 = (1 - f_{\text{in}})^2$. PBS has the function of a CNOT gate whose control qubit is polarization

and the target is spatial DoF as follows:

$$\begin{aligned} U_{\text{CNOT}}^{\text{PBS}} &= \bigotimes_{i=1}^4 (|0_i 0_{i+4}\rangle \langle 0_i 0_{i+4}| + |0_i 1_{i+4}\rangle \langle 0_i 1_{i+4}| \\ &\quad + |1_i 1_{i+4}\rangle \langle 1_i 0_{i+4}| + |1_i 0_{i+4}\rangle \langle 1_i 1_{i+4}|). \quad (\text{A3}) \end{aligned}$$

Here, the states 0 and 1 represent V and H polarization, respectively. After the CNOT gate operations on the states of all the spatial DoF, the two entangled pairs become

$$\begin{aligned} \rho_{1-8}^{\text{CNOT}} &= U_{\text{CNOT}}^{\text{PBS}} \rho_{1-8} U_{\text{CNOT}}^{\text{PBS}\dagger} \\ &= f_1 (|0_1 0_2 0_3 0_4 0_5 0_6 1_7 1_8\rangle + |1_1 1_2 0_3 0_4 1_5 1_6 1_7 1_8\rangle \\ &\quad + |0_1 0_2 1_3 1_4 0_5 0_6 0_7 0_8\rangle + |1_1 1_2 1_3 1_4 1_5 1_6 0_7 0_8\rangle) \langle \dots | \\ &\quad + f_2 (|0_1 0_2 0_3 1_4 0_5 0_6 1_7 0_8\rangle + |1_1 1_2 0_3 1_4 1_5 1_6 1_7 0_8\rangle \\ &\quad + |0_1 0_2 1_3 0_4 0_5 0_6 0_7 1_8\rangle + |1_1 1_2 1_3 0_4 1_5 1_6 0_7 1_8\rangle) \langle \dots | \\ &\quad + f_2 (|0_1 1_2 0_3 0_4 0_5 1_6 1_7 1_8\rangle + |1_1 0_2 0_3 0_4 1_5 0_6 1_7 1_8\rangle \\ &\quad + |0_1 1_2 1_3 1_4 0_5 1_6 0_7 0_8\rangle + |1_1 0_2 1_3 1_4 1_5 0_6 0_7 0_8\rangle) \langle \dots | \\ &\quad + f_3 (|0_1 1_2 0_3 1_4 0_5 1_6 1_7 0_8\rangle + |1_1 0_2 0_3 1_4 1_5 0_6 1_7 0_8\rangle \\ &\quad + |0_1 1_2 1_3 0_4 0_5 1_6 0_7 1_8\rangle + |1_1 0_2 1_3 0_4 1_5 0_6 0_7 1_8\rangle) \langle \dots |. \quad (\text{A4}) \end{aligned}$$

The four states of spatial DoF $|0_5 0_6 1_7 1_8\rangle$, $|1_5 1_6 0_7 0_8\rangle$, $|0_5 1_6 1_7 0_8\rangle$, $|1_5 0_6 0_7 1_8\rangle$ stand for the cases where each output port has one photon, i.e., so-called ‘‘four-mode cases’’ [19]. When we make a postselection of the above case, the states are projected into

$$\begin{aligned} \rho_{1-8}^{\text{CNOT}} &= f_1 (|0_1 0_2 0_3 0_4 0_5 0_6 1_7 1_8\rangle \\ &\quad + |1_1 1_2 1_3 1_4 1_5 1_6 0_7 0_8\rangle) \langle \dots | \\ &\quad + f_3 (|0_1 1_2 0_3 1_4 0_5 1_6 1_7 0_8\rangle \\ &\quad + |1_1 0_2 1_3 0_4 1_5 0_6 0_7 1_8\rangle) \langle \dots |. \end{aligned}$$

For the convenience of analysis, we rewrite the state in order from left upper and lower to right upper and lower of the spatial DoF, i.e., $|0_5 1_6 0_7 1_8\rangle$. Under the new order, the above density matrix becomes

$$\begin{aligned} \rho_{1-8}^{\text{CNOT}} &= f_1 (|0_1 0_2 0_3 0_4 0_5 1_6 0_7 1_8\rangle + |1_1 1_2 1_3 1_4 0_5 1_6 0_7 1_8\rangle) \langle \dots | \\ &\quad + f_3 (|0_1 0_2 1_3 1_4 0_5 1_6 0_7 1_8\rangle \\ &\quad + |1_1 1_2 0_3 0_4 0_5 1_6 0_7 1_8\rangle) \langle \dots |. \end{aligned}$$

In the PSBZ protocol, this measurement is implemented by measuring the photons in two lower paths with bases $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Therefore, when the measurement result in Alice and Bob is $|+\rangle|+\rangle$ or $|-\rangle|-\rangle$, the state is chosen with

$$\begin{aligned} \rho_{\text{upper}}^{\pm\pm} &= \rho_{\text{upper}}^{++} = \rho_{\text{upper}}^{--} \\ &= f_1 (|0_1 0_3\rangle + |1_1 1_3\rangle) \langle \dots | \otimes |0_5 0_7\rangle \langle \dots | \\ &\quad + f_3 (|0_1 1_3\rangle + |1_1 0_3\rangle) \langle \dots | \otimes |0_5 0_7\rangle \langle \dots |. \quad (\text{A5}) \end{aligned}$$

Omitting the spatial DoF, the residual state of the upper path is given by

$$\rho_{\text{upper}}^{\pm\pm} = f_{\text{in}}^2 |\Phi^+\rangle \langle \Phi^+| + (1 - f_{\text{in}})^2 |\Psi^+\rangle \langle \Psi^+|.$$

If the measurement result is $|+-\rangle$ or $| - + \rangle$, the state achieved by Alice and Bob is

$$\begin{aligned}\rho_{\text{upper}}^{\pm\mp} &= \rho_{\text{upper}}^{+-} = \rho_{\text{upper}}^{-+} \\ &= f_{\text{in}}^2 |\Phi^-\rangle \langle \Phi^-| + (1 - f_{\text{in}})^2 |\Psi^-\rangle \langle \Psi^-|.\end{aligned}$$

For this case, Alice or Bob should make a local phase flip gate on her/his photon to obtain target state $|\Phi^+\rangle$.

APPENDIX B: HHSZ+ AND SIMON-PAN PROTOCOLS

Considering the HHSZ+ protocol [23] in which a hyperentanglement in both polarization and spatial DoFs is used, the state of a system is $\rho_{\text{in}} = \rho_p \otimes \rho_s$. Here, ρ_p and ρ_s are the density operators of polarization and spatial DoFs, respectively. The Bell states of the spatial DoF are given by

$$\begin{aligned}|\Phi_s^\pm\rangle &= \frac{1}{\sqrt{2}}(|0_3 0_4\rangle \pm |1_3 1_4\rangle), \\ |\Psi_s^\pm\rangle &= \frac{1}{\sqrt{2}}(|0_3 1_4\rangle \pm |1_3 0_4\rangle).\end{aligned}\quad (\text{B1})$$

As shown in Fig. 3(b), we use subscripts 1 (3) and 2 (4) to label the polarization (spatial) DoF of Alice and Bob, respectively. Hyperentanglement is distributed to Alice and Bob, and will be a mixed one via a noisy channel given by

$$\rho_{\text{in}}^p = f_{\text{in}}^p |\Phi_p^+\rangle \langle \Phi_p^+| + (1 - f_{\text{in}}^p) |\Psi_p^+\rangle \langle \Psi_p^+| \quad (\text{B2})$$

and

$$\rho_{\text{in}}^s = f_{\text{in}}^s |\Phi_s^+\rangle \langle \Phi_s^+| + (1 - f_{\text{in}}^s) |\Psi_s^+\rangle \langle \Psi_s^+|. \quad (\text{B3})$$

The coefficients f_{in}^p and f_{in}^s are the initial fidelities of polarization and spatial DoFs, respectively. The fidelities satisfy conditions $f_{\text{in}}^p > \frac{1}{2}$ and $f_{\text{in}}^s > \frac{1}{2}$. Therefore, the state of the whole system is

$$\begin{aligned}\rho_{1-4} &= \rho_{12} \otimes \rho_{34} \\ &= \frac{1}{4} [f_1(|0_1 0_2\rangle + |1_1 1_2\rangle)(|0_3 0_4\rangle + |1_3 1_4\rangle) \langle \dots | \\ &\quad + f_2(|0_1 0_2\rangle + |1_1 1_2\rangle)(|0_3 1_4\rangle + |1_3 0_4\rangle) \langle \dots | \\ &\quad + f_3(|0_1 1_2\rangle + |1_1 0_2\rangle)(|0_3 0_4\rangle + |1_3 1_4\rangle) \langle \dots | \\ &\quad + f_4(|0_1 1_2\rangle + |1_1 0_2\rangle)(|0_3 1_4\rangle + |1_3 0_4\rangle) \langle \dots |],\end{aligned}$$

where the fidelities f_1 , f_2 , f_3 , and f_4 are $f_1 = f_{\text{in}}^p f_{\text{in}}^s$, $f_2 = f_{\text{in}}^p (1 - f_{\text{in}}^s)$, $f_3 = f_{\text{in}}^s (1 - f_{\text{in}}^p)$, and $f_4 = (1 - f_{\text{in}}^p)(1 - f_{\text{in}}^s)$. To purify the above state, the bilateral CNOT gates realized by two PBSs are applied as

$$\begin{aligned}U_{\text{CNOT}}^{\text{PBS}} &= \bigotimes_{i=1}^2 (|0_i 0_{i+2}\rangle \langle 0_i 0_{i+2}| + |0_i 1_{i+2}\rangle \langle 0_i 1_{i+2}| \\ &\quad + |1_i 1_{i+2}\rangle \langle 1_i 0_{i+2}| + |1_i 0_{i+2}\rangle \langle 1_i 1_{i+2}|).\end{aligned}\quad (\text{B4})$$

After the CNOT operations, the system is transferred to

$$\begin{aligned}\rho_{1-4}^{\text{CNOT}} &= U_{\text{CNOT}}^{\text{PBS}} \rho_{1-4} U_{\text{CNOT}}^{\text{PBS}\dagger} \\ &= \frac{1}{4} [f_1(|0_1 0_2 0_3 0_4\rangle + |1_1 1_2 1_3 1_4\rangle \\ &\quad + |0_1 0_2 1_3 1_4\rangle + |1_1 1_2 0_3 0_4\rangle) \langle \dots | \\ &\quad + f_2(|0_1 0_2 0_3 1_4\rangle + |1_1 1_2 1_3 0_4\rangle \\ &\quad + |0_1 0_2 1_3 0_4\rangle + |1_1 1_2 0_3 1_4\rangle) \langle \dots |]\end{aligned}$$

$$\begin{aligned}&+ f_3(|0_1 1_2 0_3 1_4\rangle + |1_1 0_2 1_3 0_4\rangle \\ &+ |0_1 1_2 1_3 0_4\rangle + |1_1 0_2 0_3 1_4\rangle) \langle \dots | \\ &+ f_4(|0_1 1_2 0_3 0_4\rangle + |1_1 0_2 1_3 1_4\rangle \\ &+ |0_1 1_2 1_3 1_4\rangle + |1_1 0_2 0_3 0_4\rangle) \langle \dots |].\end{aligned}\quad (\text{B5})$$

Analyzing the above density operator, we find that there are four cases, $0_3 0_4$, $0_3 1_4$, $1_3 0_4$, and $1_3 1_4$, for obtaining the final entangled pair. The cases $0_3 0_4$ (two upper) and $1_3 1_4$ (two lower) induce the system with $\rho_{\text{out}}^p = f_{\text{in}}^p f_{\text{in}}^s |\Phi_p^+\rangle \langle \Phi_p^+| + (1 - f_{\text{in}}^p)(1 - f_{\text{in}}^s) |\Psi_p^+\rangle \langle \Psi_p^+|$. The fidelity of the residual entangled pair $|\Phi_p^+\rangle$ is $f_{\text{out}}^p = \frac{f_{\text{in}}^p f_{\text{in}}^s}{f_{\text{in}}^p f_{\text{in}}^s + (1 - f_{\text{in}}^p)(1 - f_{\text{in}}^s)}$. With conditions $f_{\text{in}}^p > \frac{1}{2}$ and $f_{\text{in}}^s > \frac{1}{2}$, it has higher fidelity, i.e., $f_{\text{out}}^p > f_{\text{in}}^p$ and $f_{\text{out}}^s > f_{\text{in}}^s$.

In some special cases, such as strong robustness of spatial or time DoFs in experiments [25,26], the distributed state of this robust DoF could be a pure state, e.g., $f_{\text{in}}^s = 1$ in Eq. (B3). The entanglement purification of this case is the Simon-Pan protocol [11]. In Eq. (B5), if the spatial fidelity is $f_{\text{in}}^s = 1$, the density operator after the CNOTs is

$$\begin{aligned}\rho_{1-4}^{\text{CNOT}} &= U_{\text{CNOT}}^{\text{PBS}} \rho_{1-4} U_{\text{CNOT}}^{\text{PBS}\dagger} \\ &= \frac{1}{4} [f_{\text{in}}^p (|0_1 0_2 0_3 0_4\rangle + |1_1 1_2 1_3 1_4\rangle \\ &\quad + |0_1 0_2 1_3 1_4\rangle + |1_1 1_2 0_3 0_4\rangle) \langle \dots | \\ &\quad + (1 - f_{\text{in}}^p) (|0_1 1_2 0_3 1_4\rangle + |1_1 0_2 1_3 0_4\rangle \\ &\quad + |0_1 1_2 1_3 0_4\rangle + |1_1 0_2 0_3 1_4\rangle) \langle \dots |].\end{aligned}\quad (\text{B6})$$

One can see that the cases $0_3 0_4$ and $1_3 1_4$ with probability f_{in}^p will produce $|\Phi_p^+\rangle$, and the $0_3 1_4$ and $1_3 0_4$ with probability $1 - f_{\text{in}}^p$ obtain $|\Psi_p^+\rangle$ which can be transformed to $|\Phi_p^+\rangle$ by a local bit-flip operation. The fidelity of purified state is 1.

APPENDIX C: LI AND SHENG-DENG PROTOCOLS

The entanglement purification using hyperentanglement for cases considering the both bit-flip and phase errors of polarization are the Li [12] and Sheng-Deng [13] protocols. When considering phase errors, the density operator will be added with the two terms of phase error and becomes

$$\begin{aligned}\rho_{\text{in}}^p &= f_{\text{in}}^{p1} |\Phi_p^+\rangle \langle \Phi_p^+| + f_{\text{in}}^{p2} |\Psi_p^+\rangle \langle \Psi_p^+| \\ &\quad + f_{\text{in}}^{p3} |\Phi_p^-\rangle \langle \Phi_p^-| + f_{\text{in}}^{p4} |\Psi_p^-\rangle \langle \Psi_p^-|,\end{aligned}\quad (\text{C1})$$

where the fidelities satisfy $f_{\text{in}}^{p4} = 1 - f_{\text{in}}^{p1} - f_{\text{in}}^{p2} - f_{\text{in}}^{p3}$. And the entangled spatial DoF is the ideal case $\rho_s = |\Phi_s^+\rangle \langle \Phi_s^+|$. As shown in Figs. 4(a) and 4(b), the four half-wave plates used in the upper and lower paths act as CNOT gates whose control and target qubit are spatial and polarization DoFs, respectively. The two kinds of CNOT operations are

$$\begin{aligned}U_{\text{CNOT}}^u &= \bigotimes_{i=1}^2 (|1_i 0_{i+2}\rangle \langle 0_i 0_{i+2}| + |0_i 1_{i+2}\rangle \langle 0_i 1_{i+2}| \\ &\quad + |0_i 0_{i+2}\rangle \langle 1_i 0_{i+2}| + |1_i 1_{i+2}\rangle \langle 1_i 1_{i+2}|)\end{aligned}\quad (\text{C2})$$

and

$$U_{\text{CNOT}}^l = \bigotimes_{i=1}^2 (|0_i 0_{i+2}\rangle\langle 0_i 0_{i+2}| + |1_i 1_{i+2}\rangle\langle 0_i 1_{i+2}| + |1_i 0_{i+2}\rangle\langle 1_i 0_{i+2}| + |0_i 1_{i+2}\rangle\langle 1_i 1_{i+2}|). \quad (\text{C3})$$

Here, U_{CNOT}^u and U_{CNOT}^l are corresponding to Pauli- X gates (half-wave plates) in upper and lower paths, respectively. So, with a series of operations in sequence of bilateral upper-CNOT, PBS-CNOT, and lower-CNOT gates shown in Figs. 4(a) and 4(b), the entangled pair is governed by U_{lpu} as follows:

$$U_{lpu} = U_{\text{CNOT}}^l U_{\text{CNOT}}^{\text{PBS}} U_{\text{CNOT}}^u = \bigotimes_{i=1}^2 (|0_i 0_{i+2}\rangle\langle 1_i 0_{i+2}| + |1_i 1_{i+2}\rangle\langle 0_i 1_{i+2}| + |1_i 0_{i+2}\rangle\langle 1_i 1_{i+2}| + |0_i 1_{i+2}\rangle\langle 0_i 0_{i+2}|). \quad (\text{C4})$$

The final state of system is calculated as

$$\begin{aligned} \rho &= U_{lpu}(\rho_p \otimes \rho_s)U_{lpu}^\dagger \\ &= f_{\text{in}}^{p1}(|0_1 0_2 1_3 1_4\rangle + |0_1 0_2 0_3 0_4\rangle \\ &\quad + |1_1 1_2 1_3 1_4\rangle + |1_1 1_2 0_3 0_4\rangle)\langle \dots | \\ &\quad + f_{\text{in}}^{p2}(|0_1 0_2 1_3 0_4\rangle + |0_1 0_2 0_3 1_4\rangle \\ &\quad + |1_1 1_2 1_3 0_4\rangle + |1_1 1_2 0_3 1_4\rangle)\langle \dots | \\ &\quad + f_{\text{in}}^{p3}(|0_1 0_2 1_3 1_4\rangle - |0_1 0_2 0_3 0_4\rangle \\ &\quad + |1_1 1_2 1_3 1_4\rangle - |1_1 1_2 0_3 0_4\rangle)\langle \dots | \\ &\quad + f_{\text{in}}^{p4}(|0_1 0_2 1_3 0_4\rangle - |0_1 0_2 0_3 1_4\rangle \\ &\quad + |1_1 1_2 1_3 0_4\rangle - |1_1 1_2 0_3 1_4\rangle)\langle \dots |. \end{aligned} \quad (\text{C5})$$

Analyzing the above density operator, one can find that all the four cases of spatial DoF, i.e., $0_3 0_4$, $0_3 1_4$, $1_3 0_4$, and $1_3 1_4$, will obtain $|\Phi_p^+\rangle$ with fidelity 1.

The learning results show that the three gates arranged with the sequences of p-s-CNOT (p is control bit and s is target) \otimes s-p-CNOT (s is control bit and p is target) \otimes p-s-CNOT and s-p-CNOT \otimes p-s-CNOT \otimes s-p-CNOT can realize the same goal with the Li protocol. Some kinds of gate combinations are given as follows. The lpl gate is

$$U_{lpl} = U_{\text{CNOT}}^l U_{\text{CNOT}}^{\text{PBS}} U_{\text{CNOT}}^l = \bigotimes_{i=1}^2 (|0_i 0_{i+2}\rangle\langle 0_i 0_{i+2}| + |1_i 1_{i+2}\rangle\langle 1_i 1_{i+2}| + |1_i 0_{i+2}\rangle\langle 0_i 1_{i+2}| + |0_i 1_{i+2}\rangle\langle 1_i 0_{i+2}|). \quad (\text{C6})$$

The above lpl gate equals to plp gate, i.e., $U_{plp} = U_{\text{CNOT}}^{\text{PBS}} U_{\text{CNOT}}^l U_{\text{CNOT}}^{\text{PBS}} = U_{lpl}$. The plp gate is shown as an example in Figs. 4(c) and 4(d). The upu gate is expressed by

$$U_{upu} = U_{\text{CNOT}}^u U_{\text{CNOT}}^{\text{PBS}} U_{\text{CNOT}}^u = \bigotimes_{i=1}^2 (|1_i 0_{i+2}\rangle\langle 1_i 0_{i+2}| + |0_i 1_{i+2}\rangle\langle 0_i 1_{i+2}| + |0_i 0_{i+2}\rangle\langle 1_i 1_{i+2}| + |1_i 1_{i+2}\rangle\langle 0_i 0_{i+2}|). \quad (\text{C7})$$

It is also equal to pup gate as $U_{pup} = U_{\text{CNOT}}^{\text{PBS}} U_{\text{CNOT}}^u U_{\text{CNOT}}^{\text{PBS}} = U_{upu}$. Another gate upl is written by

$$U_{upl} = U_{\text{CNOT}}^u U_{\text{CNOT}}^{\text{PBS}} U_{\text{CNOT}}^l = \bigotimes_{i=1}^2 (|1_i 0_{i+2}\rangle\langle 0_i 0_{i+2}| + |0_i 1_{i+2}\rangle\langle 1_i 1_{i+2}| + |0_i 0_{i+2}\rangle\langle 0_i 1_{i+2}| + |1_i 1_{i+2}\rangle\langle 1_i 0_{i+2}|). \quad (\text{C8})$$

-
- [1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
- [2] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, Quantum state transfer and entanglement distribution among distant nodes in a quantum network, *Phys. Rev. Lett.* **78**, 3221 (1997).
- [3] H. J. Kimble, The quantum internet, *Nature (London)* **453**, 1023 (2008).
- [4] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: A vision for the road ahead, *Science* **362**, eaam9288 (2018).
- [5] H. Zhang, Y. Li, C. Zhang, and T. Huang, Connection-oriented and connectionless quantum internet considering quantum repeaters, [arXiv:2208.03930](https://arxiv.org/abs/2208.03930).
- [6] G. L. Long, D. Pan, Y. Sheng, Q. Xue, J. Lu, and L. Hanzo, An evolutionary pathway for the quantum internet relying on secure classical repeaters, *IEEE Network* **36**, 82 (2022).
- [7] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, *Phys. Rev. Lett.* **76**, 722 (1996).
- [8] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Quantum privacy amplification and the security of quantum cryptography over noisy channels, *Phys. Rev. Lett.* **77**, 2818 (1996).
- [9] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824 (1996).
- [10] J. W. Pan, C. Simon, C. Brukner, and A. Zeilinger, Entanglement purification for quantum communication, *Nature (London)* **410**, 1067 (2001).
- [11] C. Simon and J. W. Pan, Polarization entanglement purification using spatial entanglement, *Phys. Rev. Lett.* **89**, 257901 (2002).
- [12] X. H. Li, Deterministic polarization-entanglement purification using spatial entanglement, *Phys. Rev. A* **82**, 044304 (2010).
- [13] Y. B. Sheng and F. G. Deng, One-step deterministic polarization-entanglement purification using spatial entanglement, *Phys. Rev. A* **82**, 044305 (2010).
- [14] Y. B. Sheng and F. G. Deng, Deterministic entanglement purification and complete nonlocal Bell-state analysis with hyperentanglement, *Phys. Rev. A* **81**, 032307 (2010).
- [15] B. C. Ren, F. F. Du, and F. G. Deng, Two-step hyperentanglement purification with the quantum-state-joining method, *Phys. Rev. A* **90**, 052309 (2014).

- [16] G. Y. Wang, Q. Liu, and F. G. Deng, Hyperentanglement purification for two-photon six-qubit quantum systems, *Phys. Rev. A* **94**, 032319 (2016).
- [17] H. Zhang, Q. Liu, X.-S. Xu, J. Xiong, A. Alsaedi, T. Hayat, and F.-G. Deng, Polarization entanglement purification of nonlocal microwave photons based on the cross-Kerr effect in circuit QED, *Phys. Rev. A* **96**, 052330 (2017).
- [18] S. Krastanov, V. V. Albert, and L. Jiang, Optimized entanglement purification, *Quantum* **3**, 123 (2019).
- [19] J. W. Pan, S. Gasparoni, R. Ursin, G. Weihs, and A. Zeilinger, Experimental entanglement purification of arbitrary unknown states, *Nature (London)* **423**, 417 (2003).
- [20] R. Reichle, D. Leibfried, E. Knill, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Experimental purification of two-atom entanglement, *Nature (London)* **443**, 838 (2006).
- [21] L. K. Chen, H. L. Yong, P. Xu, X. C. Yao, T. Xiang, Z. D. Li, C. Liu, H. Lu, N. L. Liu, L. Li *et al.*, Experimental nested purification for a linear optical quantum repeater, *Nat. Photon.* **11**, 695 (2017).
- [22] N. Kalb, A. A. Reiserer, P. C. Humphreys, J. J. W. Bakermans, S. J. Kamerling, N. H. Nickerson, S. C. Benjamin, D. J. Twitchen, M. Markham, and R. Hanson, Entanglement distillation between solid-state quantum network nodes, *Science* **356**, 928 (2017).
- [23] X. M. Hu, C. X. Huang, Y. B. Sheng, L. Zhou, B. H. Liu, Y. Guo, C. Zhang, W. B. Xing, Y. F. Huang, C. F. Li, and G. C. Guo, Long-distance entanglement purification for quantum communication, *Phys. Rev. Lett.* **126**, 010503 (2021).
- [24] S. Ecker, P. Sohr, L. Bulla, M. Huber, M. Bohmann, and R. Ursin, Experimental single-copy entanglement distillation, *Phys. Rev. Lett.* **127**, 040506 (2021).
- [25] C. X. Huang, X. M. Hu, B. H. Liu, L. Zhou, Y. B. Sheng, C. F. Li, and G. C. Guo, Experimental one-step deterministic polarization entanglement purification, *Sci. Bull.* **67**, 593 (2022).
- [26] S. Ecker, P. Sohr, L. Bulla, R. Ursin, and M. Bohmann, Remotely establishing polarization entanglement over noisy polarization channels, *Phys. Rev. Appl.* **17**, 034009 (2022).
- [27] H. Yan, Y. Zhong, H. S. Chang, A. Bienfait, M. H. Chou, C. R. Conner, E. Dumur, J. Grebel, R. G. Povey, and A. N. Cleland, Entanglement purification and protection in a superconducting quantum network, *Phys. Rev. Lett.* **128**, 080504 (2022).
- [28] P. S. Yan, L. Zhou, W. Zhong, and Y. B. Sheng, Advances in quantum entanglement purification, *Sci. China Phys. Mech. Astron.* **66**, 250301 (2023).
- [29] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and J. Biamonte, Quantum machine learning, *Nature (London)* **549**, 195 (2017).
- [30] V. Dunjko and H. J. Briegel, Machine learning & artificial intelligence in the quantum domain: A review of recent progress, *Rep. Prog. Phys.* **81**, 074001 (2018).
- [31] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [32] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke *et al.*, Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).
- [33] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.* **70**, 1895 (1993).
- [34] H. J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Quantum repeaters: The role of imperfect local operations in quantum communication, *Phys. Rev. Lett.* **81**, 5932 (1998).
- [35] J. Wallnöfer, A. A. Melnikov, W. Dür, and H. J. Briegel, Machine learning for long-distance quantum communication, *PRX Quantum* **1**, 010301(R) (2020).
- [36] A. Peruzzo, J. McClean, P. Shadbolt, M. Yung, X. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, 4213 (2014).
- [37] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, *New J. Phys.* **18**, 023023 (2016).
- [38] 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, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [39] X. Zhao, B. Zhao, Z. Wang, Z. Song, and X. Wang, Practical distributed quantum information processing with LOCCNet, *npj Quantum Inf.* **7**, 159 (2021).
- [40] E. Fontana, N. Fitzpatrick, D. M. Ramo, R. Duncan, and I. Rungger, Evaluating the noise resilience of variational quantum algorithms, *Phys. Rev. A* **104**, 022403 (2021).
- [41] G. Vidal and C. M. Dawson, Universal quantum circuit for two-qubit transformations with three controlled-NOT gate, *Phys. Rev. A* **69**, 010301(R) (2004).
- [42] F. Vatan and C. Williams, Optimal quantum circuits for general two-qubit gates, *Phys. Rev. A* **69**, 032315 (2004).
- [43] MindQuantum Developer, MINDQUANTUM: A general software library supporting the development of applications for quantum computation, <https://gitee.com/mindspore/mindquantum>.
- [44] C. G. Broyden, The convergence of a class of double-rank minimization algorithms: 2. The new algorithm, *IMA J. Appl. Math.* **6**, 222 (1970).
- [45] R. Fletcher, A new approach to variable metric algorithms, *Comput. J.* **13**, 317 (1970).
- [46] D. Goldfarb, A family of variable-metric methods derived by variational means, *Math. Comp.* **24**, 23 (1970).
- [47] D. F. Shanno, Conditioning of quasi-Newton methods for function minimization, *Math. Comp.* **24**, 647 (1970).