Quantum autoencoders using mixed reference states

Check for updates

Hailan Ma[1,2], Gary J. Mooney[3], Ian R. Petersen[2], Lloyd C. L. Hollenberg ●[3] & Daoyi Dong ●[2,3] ✉

One of the fundamental tasks in quantum information theory is quantum data compression, which can be realized via quantum autoencoders that first compress quantum states to low-dimensional ones and then recover to the original ones with a reference state. When taking a pure reference state, there exists an upper bound for the encoding fidelity, which limits the compression of states with high entropy. To overcome the entropy inconsistency, we allow the reference state to be a mixed state and propose a cost function that combines the encoding fidelity and the quantum mutual information. We consider the reference states to be a mixture of maximally mixed states and pure states and propose three strategies for setting the ratio of mixedness. Numerical simulations of different states and experimental implementations on IBM quantum computers illustrate the effectiveness of our approach.

Quantum machine learning which combines machine learning and quantum computation has grown into a booming research topic[1–7]. Quantum autoencoders (QAEs) inherit the spirit of classical autoencoders to compress information into a latent space such that the original data can be recovered from a reduced-dimension representation[8,9]. They have the potential to reduce the requirements of quantum communication channels[10] and the size of quantum gates[11,12] and thus have a practical value for various applications including quantum simulation[13], quantum communication and distributed computation in quantum networks[10,11].

Owing to the potential of QAEs in quantum information processing, there is a growing interest in designing different schemes to complete state compression tasks. An early work proposed a quantum generalization of a classical neural network[14] and another work designed an autoencoder framework using programmable circuits[15]. An enhanced QAE that encodes the feature vector of the input data into single-qubit rotation gates has been implemented in variational quantum circuits[16]. There have also been achievements in the implementation of QAEs on photonic systems[12,17,18]. Apart from data compression, QAEs have also been applied to other applications, such as state denoise[19,20] and error mitigation[21]. A novel method based on QAEs has been devised to prepare the quantum Gibbs state and estimate the quantum Fisher information[22]. A hybrid QAE has been proposed to identify the emergence of order in the latent space that can be utilized for clustering and semi-supervised classification[23]. Recently, the execution of a QAE-facilitated teleportation protocol has been implemented on a silicon photonic chip[24]. Furthermore, the exploration of QAEs to analyze datasets originating from industrial contexts demonstrates their potential in processing real-world data[25].

In traditional QAE schemes[15,17,18,26], pure states are utilized as reference states for recovering the initial state. For each state to be compressed, there exists an upper bound (hereafter, we call it QAE-pure bound) for the encoding fidelity, i.e., the overlap between the trash state and the reference state. Such a bound limits the compression of states with high entropy. To compensate for the entropy inconsistency between the initial state and the recovered state[27], we allow the reference state to be a mixed one. Instead of a fixed mixed state, we configure different reference states for effective compression of different states. To achieve this, we take advantage of quantum mutual information that measures the disentanglement to guide the optimization of the encoding transformation. Meanwhile, the conventional cost function in standard QAEs aims to decouple the initial states into two parts and provides information about the inner structure of initial states, which can be useful for setting the reference state. Hence, we design a novel cost function that combines the above two factors to guide the training of QAEs towards better performance. Inspired by the compression of tensor products of identical states on IBM quantum computers[28], we experimentally realize QAEs with mixed reference states on the IBM quantum simulator *ibmq_qasm_simulator* and quantum computer *ibmq_quito*.

In this work, we leverage mixed reference states to break the upper bound of compression rate imposed by conventional QAEs[26]. In particular, we use a mixture of a pure state and the maximally mixed state, with pure reference states being a special case. Similarly, our proposed cost function considers both quantum mutual information and the existing function in conventional QAEs that favor pure reference states. Consequently, our protocol exhibits flexibility, enabling its application to various quantum states. We have observed the direct relationship between the optimal value of the purity ratio in the reference state and the QAE-pure bound. This

[1]School of Engineering and Technology, University of New South Wales, Canberra, ACT, Australia. [2]CIICADA Lab, School of Engineering, The Australian National University, Canberra, ACT, Australia. [3]School of Physics, University of Melbourne, Parkville, VIC, Australia. ✉e-mail: daoyi.dong@anu.edu.au

discovery empowers us with the insights gained from the training of QAEs to adaptively configure mixed reference states for compressing different unknown states. Our experimental implementations on IBM quantum computers demonstrate the potential of the proposed protocol in saving valuable quantum resources for real applications.

Results

QAEs using mixed reference states

Schematic. As illustrated in Fig. 1, we define the trash qubits as subsystem A and the latent qubits as subsystem B, respectively. We denote the dimensions of the original space, the latent space, and the trash space as d, d_B, and d_A, respectively. The goal of a QAE is to compress $(n_A + n_B)$-qubit state ρ_0 into n_B-qubit state ρ_{latent} via an encoder map U_e and then recover to $(n_A + n_B)$-qubit state ρ_f via a decoder map U_d. After the encoding operation U_e, the trash state and the latent state are obtained as $\rho_{\text{trash}} = \text{Tr}_B(U_e\rho_0 U_e^\dagger)$ and $\rho_{\text{latent}} = \text{Tr}_A(U_e\rho_0 U_e^\dagger)$, respectively. Denote $F(\rho_1, \rho_2)$ as the state fidelity between ρ_1 and ρ_2[29]. The efficiency of this task can be quantified by the decoding fidelity between the original state and the recovered state, i.e., $\mathcal{F}_d = F(\rho_0, \rho_f)$ and the scheme is considered reliable when $\mathcal{F}_d$ approaches 1. During the whole process, a reference state is utilized for two aspects: (i) measure the encoding fidelity between the trash state and the reference state, denoted as $\mathcal{F}_e = F(\rho_{\text{trash}}, \rho_{\text{ref}})$; (ii) reproduce the initial states with the combination of the latent state and the reference state. When the unitary operation U_e perfectly disentangles ρ_0 into two parts as $U_e\rho_0 U_e^\dagger = \rho_{\text{latent}} \otimes \rho_{\text{ref}}$, the overlap between the trash state and the reference state can achieve unity[15,26] and the decoding fidelity $\mathcal{F}_d$ can also achieve unity. In this work, the unitary transformation U_e is realized through Hamitonian-based control (see Section Methods for detailed information).

When taking a pure state as reference states $\rho_{\text{ref}} = |\psi_{\text{ref}}\rangle\langle\psi_{\text{ref}}|$ for compressing the initial state ρ_0. There exists an upper bound (i.e., QAE-pure bound, abbreviated as $Q_{\text{pure}}^{\text{bound}}$) for the encoding fidelity between ρ_{trash} and ρ_{ref}. From the previous work[26], we have

$$Q_{\text{pure}}^{\text{bound}}(\rho_0) = \max_{U_e} F(\text{Tr}_A(U_e\rho_0 U_e^\dagger), |\psi_{\text{ref}}\rangle) = \sum_{k=1}^{d_B} \lambda_k(\rho_0), \quad (1)$$

where $\lambda_k(\rho)$ is the k-th (in descending order) eigenvalue of ρ. This bound is determined by eigenvalues of the initial state ρ_0, with no dependency on the pure reference state $|\psi_{\text{ref}}\rangle$[26]. Hence, in traditional QAEs, a common practice is to utilize a fixed pure state as the reference state, e.g., $|0\rangle\langle0|$. According to Eq. (1), if ρ_0 has a rank larger than d_B, the optimal encoder U_e can only decouple the largest d_B eigenvalues of ρ_0, whose sum is less than one[26,27]. As such, a high-rank state in this work means the rank of its density matrix is larger than d_B.

When compressing ρ_0 with high entropy, the trash state ρ_{trash} tends to have high entropy and consequentially have low overlap with a pure reference state (e.g., setting $\rho_{\text{ref}} = |0\rangle\langle0|$). In the decoding stage, the low entropy of a pure state may also limit the entropy of the recovered state (see ρ_f in Fig. 1). To overcome the entropy inconsistency between the initial state ρ_0 and the recovered state ρ_f[27], we remove the limitation of a pure reference state and allow the reference state ρ_{ref} to be mixed. The limitation of the conventional QAEs also motivates us to adopt different mixed states rather than a fixed state for compressing different initial states. In this way, the entropy in the reference state can assist the decoder in achieving a high fidelity for the recovered state. Here, the introduction of entropy offers an intuitive strategy for setting mixed reference states. To ensure that the recovered state ρ_f has high fidelity with ρ_0, additional efforts are required to optimize U_e. Instead of searching U_e and ρ_{ref} together (a full encoding and decoding procedure is required), we accomplish the task of QAEs with mixed reference states within two stages.

Cost function. In conventional QAEs, the reference state is fixed as a pure state, and $\mathcal{F}_e$ is utilized as the cost function to train

QAEs[15,17,18,26]. However, $\mathcal{F}_e$ is different from $\mathcal{F}_d$ which characterizes the effectiveness of QAEs in compressing and recovering quantum data. When allowing the reference state to be a mixed state, $\mathcal{F}_e$ can achieve one by setting $\rho_{\text{ref}} = \rho_{\text{trash}}$, whereas $\mathcal{F}_d$ is usually less than one and can reach one only when perfect disentanglement is realized[15]. Given that quantum mutual information (QMI) measures the correlation between subsystems of quantum states[30,31], it quantifies the amount of noise that is required to erase (destroy) the correlations completely. To facilitate QAEs using mixed reference states, we aim to disentangle $U_e\rho_0 U_e^\dagger$, which is achieved by minimizing the QMI, i.e., $\mathcal{I}(U_e\rho_0 U_e^\dagger)$ or equivalently maximizing

$$J_{\text{qmi}} = -\mathcal{I}(U_e\rho_0 U_e^\dagger), \quad (2)$$

where $\mathcal{I}(\rho) = S(\text{Tr}_A(\rho)) + S(\text{Tr}_B(\rho)) - S(\rho)$ denotes the QMI of ρ and $S(\rho) = -\text{Tr}(\rho \ln(\rho))$ denotes the von Neumann entropy of ρ. Generally $\mathcal{I}(\rho) \geq 0$, and $\mathcal{I}(\rho) = 0$ when $\rho = \text{Tr}_A(\rho) \otimes \text{Tr}_B(\rho)$.

According to existing research[26], when training QAEs using the overlap between the trash state and a pure reference state (e.g., taking the reference state as $\rho_{\text{ref}} = |0\rangle\langle0|)$, $J_{\text{pure}} = F(\text{Tr}_B(U_e\rho_0 U_e^\dagger), |0\rangle\langle0|)$ as the cost function, the compression rate of QAEs can approach the theoretical QAE-pure bound. Under that scheme, a high compression rate can be realized for low-rank states[26]. Although this cost function fails to disentangle a high-rank state with satisfactory performance, the optimization of J_{pure} leads to a direction of reorganizing the information of initial states into two parts. To combine the cases of low-rank states and high-rank states, we propose a cost function,

$$\Phi(w) = wJ_{\text{pure}} + (1 - w)J_{\text{qmi}}, \quad (3)$$

where $w \in [0, 1]$ controls the ratio of different factors. This protocol is termed QAE-qmi in this paper. Considering the potential of evolutionary strategy (ES) in conventional QAEs[26], we employ it to optimize the parameters of U_e.

Reference states. Recall the nature of QAEs lies in disentangling[15]. The encoding fidelity that measures the overlap between the trash states and the reference states can reach one (i.e., $\mathcal{F}_e = 1$) by setting $\rho_{\text{ref}} = \rho_{\text{trash}}$. Although $\mathcal{F}_d \leq \mathcal{F}_e$, in the general case, $\mathcal{F}_d$ can approach $\mathcal{F}_e$ and they both achieve one when perfect disentangling is realized[26]. When there is no limitation for the reference state, it is helpful to investigate the performance of QAE-qmi with $\rho_{\text{ref}} = \rho_{\text{trash}}$.

In practical applications, it may be useful to utilize reference states with some physical constraints. According to our previous study, a pure reference
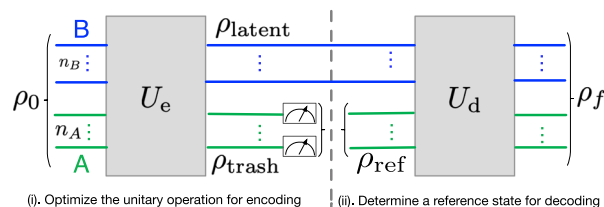

Fig. 1 | Quantum circuit representation of a QAE using a mixed reference state. The network includes two parts: (i) The encoder U_e reorganizes the $(n_A + n_B)$-qubit initial state ρ_0 into two parts, i.e., n_B-qubit state ρ_{latent} contains the useful information (blue lines) that represent the latent qubits and n_A-qubit state ρ_{trash} contains the superfluous information (green lines) that represent the trash qubits. Here, the trash state is obtained by tracing out the latent space of n_B qubits. The latent state is obtained by tracing out the trash space of n_A qubits. (ii) The decoder U_d recovers the state ρ_f by using the combination of the latent state and ancillary fresh qubits (initialized to the reference state). The goal of QAEs is to maximize the overlap between the recovered state ρ_f and the original state ρ_0. Note that, we use the common practice of taking the decoding as the inverse of the encoding of $U_d = U_e^\dagger$[17,26].

state, e.g., $|0\rangle\langle 0|$ is effective in compressing low-rank states, with the compression rate approaching the QAE-pure bound[26], whose value is usually high for low-rank states. For high-rank states with high entropy, the introduction of mixed reference states helps increase the entropy of the recovered states[27]. While the maximally mixed states I/d_A has the highest entropy among all states in $\mathcal{H}_A$ and is effective for increasing the entropy in the decoding stage. To achieve a good QAE for different quantum states, we take the following reference state

$$\rho_{\mathrm{mix}} = p_r|0\rangle\langle 0| + (1 - p_r)I/d_A, \qquad (4)$$

where p_r represents the ratio of the pure state and $(1 - p_r)$ represents the ratio of the mixed state in the reference state. I denotes the d_A-dimensional identity matrix. Different initial states with different inner structures may have different optimal reference states following the form of Eq. (4). When compressing initial states with high entropy, it is preferable to use low p_r that generates high entropy for ρ_{mix}. As such, it is desirable to specify an optimal p_r for different quantum states. Although mixed reference states cost additional memories, our method aims to achieve high fidelity between the recovered state and the initial state. This is particularly important for initial states with low QAE-pure bound. By constraining the mixed reference state in the form of Eq. (4), one can transmit the compressed latent representation and p_r to facilitate the subsequent recovery.

Now, we focus on determining a good p_r to recover quantum states with high decoding fidelity. Intuitively, quantum states with different inner structures (entropy) require different optimal p_r to achieve optimal decoding fidelity. Before deciding the optimal p_r for recovering the state, we first propose a grid-search strategy for setting p_r (marked as *grid*). Under a fixed cost function (e.g., $\Phi = 0.5(J_{\mathrm{pure}} + J_{\mathrm{qmi}})$), we define a candidate set (e.g., $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$) for p_r. Then, following the optimized U_e during the encoding process, the decoding process is performed with different p_r selected from the candidate pool. Eventually, the optimal one is determined regarding its decoding fidelity, and this value is defined as p_r^{grid}. In the following, we also introduce another two strategies: (i) *bound*: to leverage the QAE-pure bound when prior knowledge of the initial state is available; (ii) *guess*: to infer from the training process of QAEs. The *guess* strategy acts as a practical solution that can be implemented in different cases.

In summary, the optimization of QAEs using a mixed reference state is accomplished within two stages. Firstly, the training of QAEs using $\Phi(w) = wJ_{\mathrm{pure}} + (1 - w)J_{\mathrm{qmi}}$ is implemented in the encoding stage, while the mixed reference state is introduced in the decoding stage. After the optimization of U_e is finished by maximizing Φ, we determine a reference state for recovery to maximize the overlap between the initial state and the recovered state.

Numerical results
Quantum state settings.
For the compression task, we consider three classes of quantum states. Firstly, we consider thermal states as

$$\rho = \frac{e^{-\beta H}}{\mathrm{Tr}(e^{-\beta H})}, \qquad (5)$$

where β is the inverse temperature and H denotes the Hamitonian. Let $\sigma_z^j(\sigma_x^j)$ denote the composite value of $\sigma_z(\sigma_x)$ on the j-th qubit with identity matrices for the other qubits. For example, we investigate thermal states with the Hamiltonian of the one-dimensional transverse-field

$$H = -\left(\sum_j \sigma_z^j \sigma_z^{j+1} + \sum_j \sigma_x^j\right), \qquad (6)$$

with couplings set to 1. Then, we investigate Werner states, which are bipartite and are invariant under any unitary operator in the form of $U \otimes U$[32]. Let $|k\rangle$ and $|j\rangle$ be the computational basis for two bipartite subspaces,
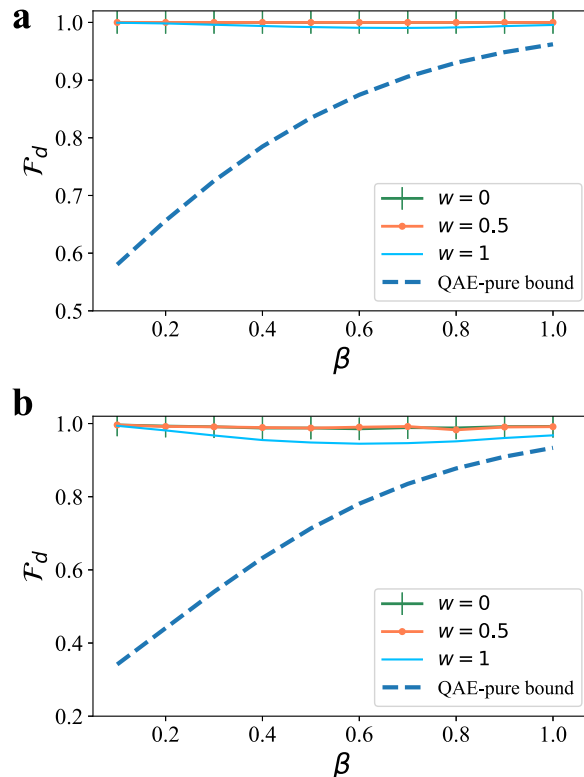

Fig. 2 | Comparison of QAE-qmi under different w using $\rho_{\mathrm{ref}} = \rho_{\mathrm{trash}}$. **a** for 2-qubit states, **b** for 4-qubit states. β controls the entropy of the thermal states, i.e., an increase of β leads to a decrease of entropy in quantum states. $\mathcal{F}_d$ represents the decoding fidelity between the initial state and the recovered state. w denotes the ratio of J_{pure} in the cost function $\Phi(w)$. The blue dashed line represents the theoretical upper bound of compression rate when only considering pure reference states[26].

respectively. A Werner state can be parameterized by

$$\rho(\alpha) = \frac{1}{d^2 - d\alpha}\left(I - \alpha\sum_{kj}|kj\rangle\langle jk|\right), \qquad (7)$$

where I denotes the d^2-dimensional identity matrix and α varies between -1 and 1. Additionally, we also consider the initial states that have a similar form to ρ_{mix} as

$$\rho(p_0) = p_0|\psi\rangle\langle\psi| + (1 - p_0)I/d, \qquad (8)$$

where I denotes the d-dimensional identity matrix and the value of p_0 controls the purity of the initial states. In this work, we randomly generate a pure state $|\psi\rangle$ and utilize it for different values of p_0. By computation, we find that states with high p_0 have a high QAE-pure bound. In this work, we consider compressing 2-qubit states into 1-qubit states and 4-qubit states into 2-qubit states. The density matrices of 2-qubit thermal states and Werner states are presented in Supplementary Discussion 1.

Investigation of different w. Firstly, we investigate the performance of $\rho_{\mathrm{ref}} = \rho_{\mathrm{trash}}$. In particular, we consider three different cases: (i) $w = 1$ considers the fidelity between the trash state and a fixed pure state; (ii) $w = 0$ considers QMI of the encoding state $U_e\rho_0 U_e^\dagger$; (iii) $w = 0.5$ considers both factors and acts as a more general function. Comparsion results for 2/4-qubit thermal states and 2/4-qubit Werner states are provided in Figs. 2 and 3, respectively. These results demonstrate that $\Phi = J_{\mathrm{qmi}}$ and $\Phi = 0.5(J_{\mathrm{pure}} + J_{\mathrm{qmi}})$ achieve a similar decoding fidelity, higher than that of $\Phi = J_{\mathrm{pure}}$. The gaps between $w = 0.5$ and $w = 1$ suggest that introducing

quantum mutual information to the cost function $\Phi(w)$ helps enhance the decoding fidelity for 4-qubit states. These results suggest that the introduction of QMI is effective in compressing and recovering quantum states, especially for states with low QAE-pure bounds (i.e., thermal states with β approaching 0 and Werner states with $|\alpha|$ approaching 0). The training curves of QAE-qmi under different w are summarized in Supplementary Discussion 2.

Then we turn to the case of setting reference states as Eq. (4). Although the disentanglement evaluation of J_{qmi} naturally follows the nature of QAEs, this function fails to achieve a perfect value of 1 for non-separable initial states. When considering $\rho_{ref} = \rho_{mix}$ for low-rank states, the optimization of J_{qmi} tends to bring the trash state far away from a pure state, in conflict with the intuition that pure reference states are enough for low-rank states. To combine the two scenarios together, w = 0.5 is a good solution.

We implement an example of low-rank states (with high QAE-pure bound) using $\rho_{ref} = \rho_{mix}$ under different values of p_r, with results in Fig. 4.
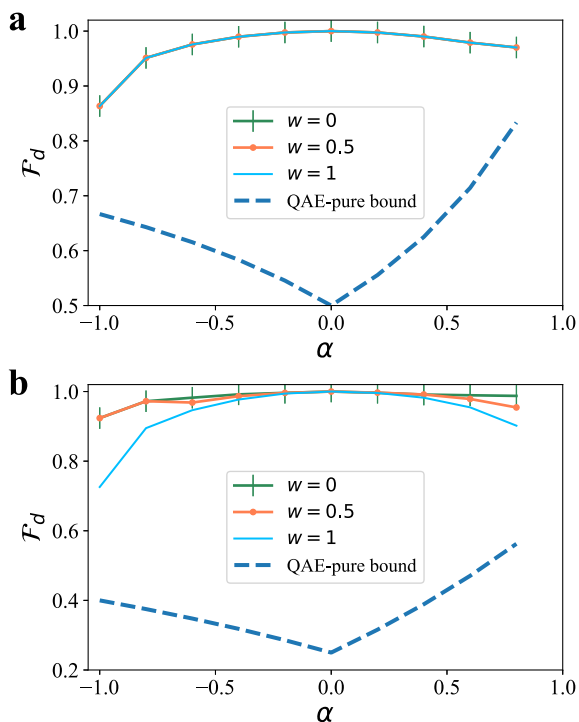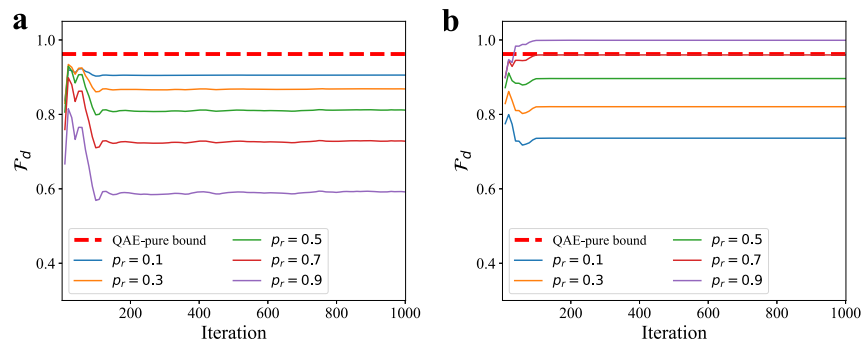
Fig. 3 | **Comparison of QAE-qmi under different w for Werner states under $\rho_{ref} = \rho_{trash}$. a** for 2-qubit states, **b** for 4-qubit states. w denotes the ratio of J_{pure} in the cost function $\Phi(w)$. $\mathcal{F}_d$ denotes the fidelity between the initial state and the recovered state. The blue dashed line represents the theoretical upper bound of compression rate when only considering pure reference states[26].

While, the best solution under w = 0 corresponds to a large p_r (indicating high entropy in ρ_{mix}) and the best solution under w = 0.5 corresponds to a small p_r (indicating low entropy in ρ_{mix}). The latter case is in line with our conjecture that compressing states with low entropy and high QAE-pure bound requires a reference state with high purity. As such, we consider $\Phi = 0.5(J_{pure} + J_{qmi})$ to be useful for both $\rho_{ref} = \rho_{trash}$ and $\rho_{ref} = \rho_{mix}$. Hereafter, without specific notation, QAE-qmi refers to the case of $\Phi = 0.5(J_{pure} + J_{qmi})$.

The cost function in Eq. (3) can be applied to compress initial states with different purities. According to Eq. (1), states with high purity tend to have eigenvalues close to 0 or 1, and have a high QAE-pure bound[26], suggesting that the conventional QAE approach (with only J_{pure}) still works. In the unified approach, it is preferable to adopt a large w for compressing initial states with high purities (e.g., thermal states with large β or Werner states with $|\alpha|$ approaching 1). Furthermore, when compressing initial states with high purity, for example, $\rho(p_0)$ with high p_0, the existence of QMI in the cost function with ratio w = 0.5 brings in a negative effect, which can be overcome by reducing the ratio of J_{qmi} e.g., w = 0.99. This originates from the two stages of our protocol: firstly optimize U_e via $\Phi(w)$ and then determine a reference state for recovering. A large ratio of J_{qmi} tends to conflict with the mechanism of setting reference states as $p_r|0\rangle\langle0| + (1 - p_r)I/d_A$. This conflict increases with the purity of states, as initial states with high purities favor a high ratio of J_{pure} and a high value of p_r. Nevertheless, we can avoid this by adjusting w. Please refer to Supplementary Discussion 4 for detailed information.

Investigation of different p_r. The introduction of mixed states aims to bring in appropriate entropy for recovering different states. Intuitively, high-rank states with high entropy and low QAE-pure bounds require more entropy for recovery, i.e., a low p_r. Recall that each state has its inner structure, and can be characterized by a QAE-pure bound[26]. Initially, we tested Q_{pure}^{bound} but found that its value does not align well with p_r^{grid}, and their performance regarding $\mathcal{F}d$ exhibits a significant gap. Subsequently, we find the square of Q_{pure}^{bound} aligns more closely with p_r^{grid}, and their decoding fidelities are comparably close. Hence, we propose a second strategy (marked as *bound*) for setting p_r by leveraging the QAE-pure bound of the initial state ρ_0, and we have $p_r^{bound} = (Q_{pure}^{bound}(\rho_0))^2$.

To validate our conjecture that the optimal p_r^{grid} tends to approach the square of the QAE-pure bound, we compare the two strategies, with their actual values of p_r and the associated decoding fidelity $\mathcal{F}_d$ in one figure. From the results in Figs. 5, 6, it is clear that p_r^{grid} has the same trend as p_r^{bound}, and their decoding fidelities are close to each other under different parameters. For states in the form of $\rho(p_0)$, we observe that the best p_r found using the *grid* strategy (i.e., p_r^{grid}) tends to be approaching p_0. The comparison of the two strategies (*grid* vs *bound*) regrading p_r and $\mathcal{F}_d$ is summarized in Fig. 7. The two curves are close to each other with increasing p_0. The decoding fidelities for the two strategies of setting p_r achieve similar values.

Based on the observation that taking p_r as the square of QAE-pure bound achieves good performance, J_{pure} among the unified cost function is approaching the QAE-pure bound during the training process[26]. We

Fig. 4 | **An example of QAE-qmi with $\rho_{ref} = \rho_{mix}$ under different p_r for 2-qubit thermal states. a** for w = 0, **b** for w = 0.5. p_r represents the ratio of the pure state in the reference state ρ_{mix} and $(1 - p_r)$ represents the ratio of the mixed state in the reference state ρ_{mix}. $\mathcal{F}_d$ represents the decoding fidelity between the initial and recovered states. The red dashed line represents the theoretical upper bound of compression rate when only considering pure reference states[26]. In this figure, some selected values of p_r from the candidate pool are used to clearly demonstrate the trends of different values.
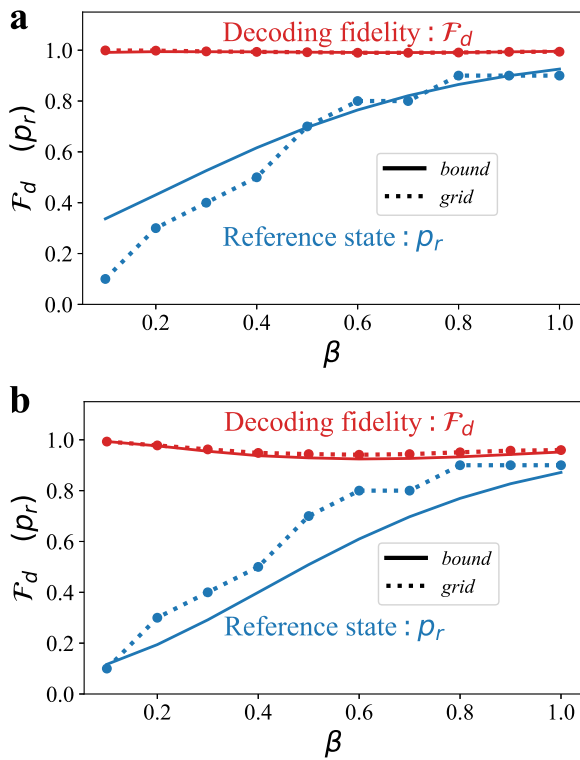
Fig. 5 | Comparison of two strategies of setting p_r for compressing and recovering thermal states. a for 2-qubit states, **b** for 4-qubit states. p_r represents the ratio of the pure state and $(1 - p_r)$ represents the ratio of the mixed state in the reference state. $\mathcal{F}_d$ represents the fidelity between the initial state and the recovered state. The solid blue line (dashed blue line) corresponds to the actual value of p_r. The solid red line (dashed red line) represents the value of $\mathcal{F}_d$ through the bound information (grid searching from a set of values).

Fig. 6 | Comparison results of two strategies of setting p_r when compressing and recovering Werner states. a for 2-qubit states, **b** for 4-qubit states. $\mathcal{F}_d$ represents the fidelity between the initial and recovered states. p_r represents the ratio of the pure state and $(1 - p_r)$ represents the ratio of the mixed state in the reference state. The solid blue line (dashed blue line) corresponds to the actual value of p_r. The solid red line (dashed red line) represents the value of $\mathcal{F}_d$ through the bound information (grid searching from a set of values).

propose a third strategy (marked as *guess*) for adaptively configuring the reference states as $p_r^{\text{guess}} = J_{\text{pure}}^2$. The comparison results of different strategies of setting p_r are summarized in Supplementary Discussion 3, demonstrating that the manual and automatic ways of setting p_r are effective in compressing and recovering different quantum states. By now, we have three strategies for setting p_r, and their comparison for compressing thermal states and Werner states is summarized in Supplementary Discussion 3. Under the strategy of $p_r = p_r^{\text{guess}}$, we further compare the performance of QAE-qmi under different w, with results shown in Supplementary Discussion 4. The decline of $\mathcal{F}_d$ when p_0 approaches one reveals that $\Phi = 0.5(J_{\text{pure}} + J_{\text{qmi}})$ hinders the compression of quantum states with high p_0.

Experimental results
Generally, it is assumed that quantum circuits deal with pure states. We need to find a solution to generate mixed states, which is essential in preparing initial states and reference states for encoding and decoding, respectively. We use the technique of *purification*[29] to associate a mixed state with a pure state in a large space. Given a state ρ_K of a quantum system K, it is possible to introduce another system R, and define a pure state $|\psi\rangle_{KR}$ for the joint system KR such that $\rho_K = \text{Tr}_R(|\psi\rangle\langle\psi|_{KR})$. The pure state $|\psi\rangle_{KR}$ reduces to ρ_K when we look at the system K alone. This mathematical procedure can be done for any state. Please refer to Supplementary Discussion 5 for detailed information about the construction of $|\psi\rangle_{KR}$ for arbitrary ρ_K.

The quantum circuit for compressing 2-qubit states into 1-qubit states is depicted in Fig. 8, where four qubits $q_0q_1q_2q_3$ are utilized to generate mixed states on q_2q_3, on which the encoding gate and the decoding gate are performed. The circuit can be divided into five parts: (a) prepare the initial state, (b) perform the encoding operation, (c) prepare the reference state, (d)
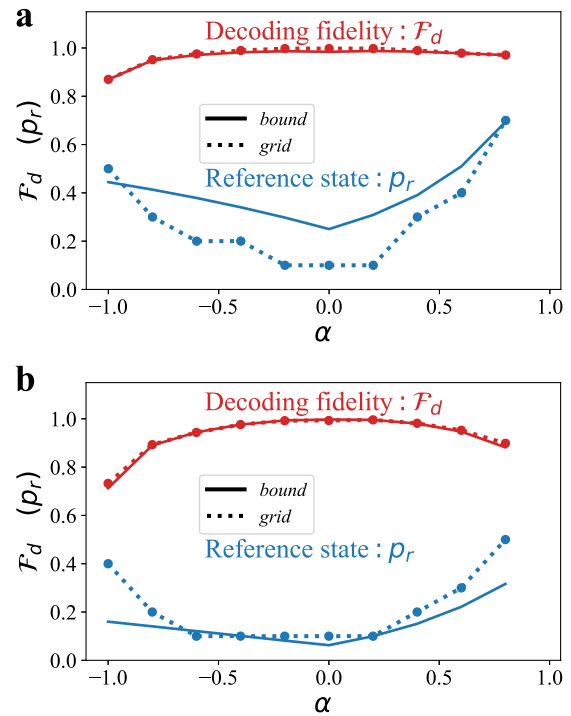
perform the decoding operation, (e) perform quantum measurements to obtain the density matrix of the recovered state. A set of complete measurements is required to specify the density matrix of a quantum state[7]. In Fig. 8, only a special case of local measurement of $\sigma_z \otimes \sigma_z$ on q_2q_3 is performed. Adding some gates (such as the Hadamard gate and the $S^\dagger$ gate) before the measurement part (i.e., between (d) and (e)) helps realize other measurements. Hence, the quantum circuits are repeated several times until a complete measurement is accomplished. Feeding the measured data to the built-in function for *quantum state tomography* in qiskit[33], the density matrix of the recovered state is finally obtained.

In this work, we do not perform the optimization loops on quantum computers. Instead, we take the encoding transformation U_e and the reference state in the form of Eq. (4) that are learned numerically on classical computers, and then deploy them on IBM quantum simulators and quantum computers, respectively. Note that, each green block in Fig. 8 represents quantum circuits composed of a sequence of quantum gates to achieve unitary operations. Please refer to Supplementary Discussion 5 for the transpiled circuits for the green blocks.

We implement the procedure of compressing and recovering 2-qubit states on *ibmq_qasm_simulator* and *ibmq_quito*, with 8192 shots. Each compression task is run 6 times on *ibmq_quito*. The comparison results are summarized in Fig. 9. The results of the simulators are in agreement with the theoretical results obtained from classical computers. However, gaps exist between the results of *ibmq_qasm_simulator* and *ibmq_quito*. In particular, the gap becomes apparent for thermal states with increasing β and Werner states with α approaching -1 or 1. The underlying reason may be that compressing states close to maximally mixed states with a low QAE-pure bound, presents a large space for improvement through the introduction of mixed reference states. By

contrast, initial states closer to pure states can achieve a high compression rate by utilizing pure reference states, while pure states can be affected by various noise sources, e.g., CNOT noise. Then, we visualize the density matrices of the initial states and the recovered states obtained from *ibmq_quito* for Werner states in Supplementary Discussion 5.

Discussion

In this paper, we have investigated the performance of QAEs with mixed reference states. One may consider employing a fixed mixed state as the reference state as in conventional QAEs. However, it is challenging to decide on a fixed reference state for different initial states, and the proof that
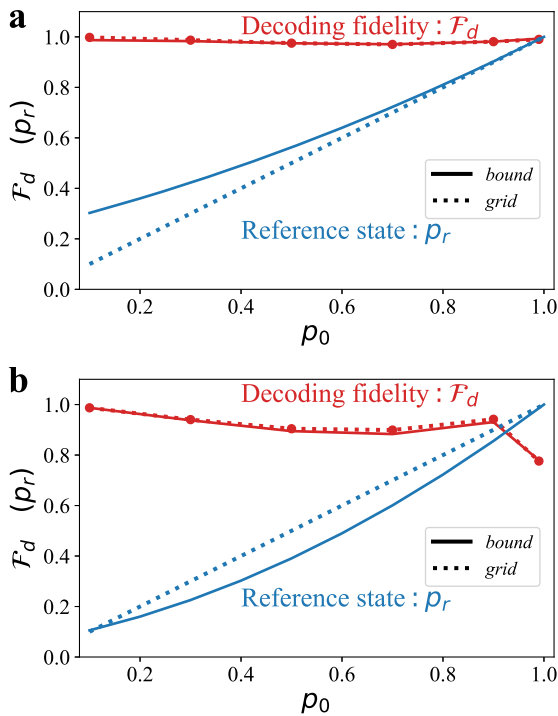

Fig. 7 | Comparison of two strategies of setting p_r for compressing maximally mixed states blended with pure states. a for 2-qubit states, **b** for 4-qubit states. p_r represents the ratio of the pure state and $(1 - p_r)$ represents the ratio of the mixed state in the reference state. $\mathcal{F}_d$ represents the fidelity between the initial state and the recovered state. The solid blue line (dashed blue line) corresponds to the actual value of p_r. The solid red line (dashed red line) represents the value of $\mathcal{F}_d$ through the bound information (grid searching from a set of values).

arbitrary mixed reference states yield similar bounds remains elusive. The way of merging all the operations into the encoder fails to reveal features of QAEs with mixed reference states. By comparsion, the adaptive configuration of mixed reference states provides a clear clue about improving fidelity via appropriate entropy compensations. Then, we summarize the characteristics of our protocol as follows.

(i) The proposed function of $\Phi(w) = wJ_{\text{pure}} + (1 - w)J_{\text{qmi}}$ combines the approximate QAE-pure bound function that reflects the inner structure of the initial states and the quantum mutual information that measures the correlation between subsystems. It is a general function that can be applied to both low-rank states and high-rank states. As demonstrated by the numerical results, training QAEs using $\Phi(w)$ achieves high decoding fidelity under different reference setting rules including $\rho_{\text{ref}} = \rho_{\text{trash}}$ and $\rho_{\text{ref}} = \rho_{\text{mix}}$. In addition, it has been found that for initial states with high QAE-pure bounds (e.g., large p_0 in Eq. (8)), it is preferable to increase w in $\Phi(w) = wJ_{\text{pure}} + (1 - w)J_{\text{qmi}}$, giving more importance to the approximate QAE-pure bound. This is consistent with the fact that pure reference states together with J_{pure} can realize a good compression rate for low-rank states[26]. However, it is crucial to recognize the inherent limitation of employing QMI, whose value depends on the von Neumann entropy that is not an observable. Consequently, this limitation restricts its applications in experimental settings. In our future work, we will explore an approximate function that can be more readily implemented in experimental contexts.

(ii) The numerical results demonstrate that setting the reference state in the form of Eq. (4) helps enhance the decoding fidelity for high-rank states. Due to the special form of the reference states, it is intuitive that different initial states may rely on different optimal purity ratios p_r that help maintain the entropy consistency between the initial states and the recovered states. As demonstrated by the numerical results in Figs. 5, 6, and 7, the optimal p_r using via the *grid* strategy is close to the square of QAE-pure bound for thermal states, Werner states and maximally mixed states blended with pure states. Such findings provide hints for adaptively setting reference states for different quantum states.

(iii) When limiting the reference states to the form of Eq. (4), we can take advantage of the prior information of the initial states to determine a mixed reference state that achieves a high $\mathcal{F}_d$. For example, $(Q_{\text{pure}}^{\text{bound}}(\rho_0))^2$ can be determined before the training process of QAEs. If no prior knowledge is available, we can also approximate the QAE-pure bound by inferring from the training process and taking $F(\rho_{\text{trash}}, |0\rangle)^2$. From this perspective, our protocol may have wide applications in practical quantum applications. For example, the compressed latent representations can be utilized to effectively denoise errors in the original states[19], or act as intermediate states to facilitate high-dimensional subspace teleportation[24].

Our work illustrates the effectiveness of QAEs using mixed reference states under different constraints and thus provides implications for
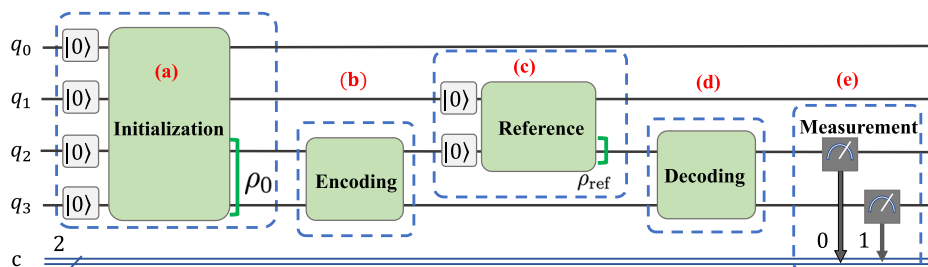

Fig. 8 | Quantum circuits for compressing 2-qubit states into 1-qubit states and recovering to 2-qubit states using mixed reference states. Firstly, four qubits are required to generate a 4-qubit pure state in **a**, whose partial trace on q_0q_1 leads to a mixed state ρ_0 on q_2q_3. Hereafter, the encoding operation in **b** and the decoding operation are performed on q_2q_3, where q_2 contains the trash state ρ_{trash} and q_3

contains the latent state ρ_{latent}. Then, q_1q_2 are reused to generate a pure state whose partial trace on q_1 leads to a reference state ρ_{ref} on q_2 in **c**. Then, the decoding operation in **d** is performed on the combination of q_2 with ρ_{ref} and q_3 with ρ_{latent}. In **e**, measurements are performed on q_2q_3 for quantum state tomography to obtain the density matrix of the recovered state.
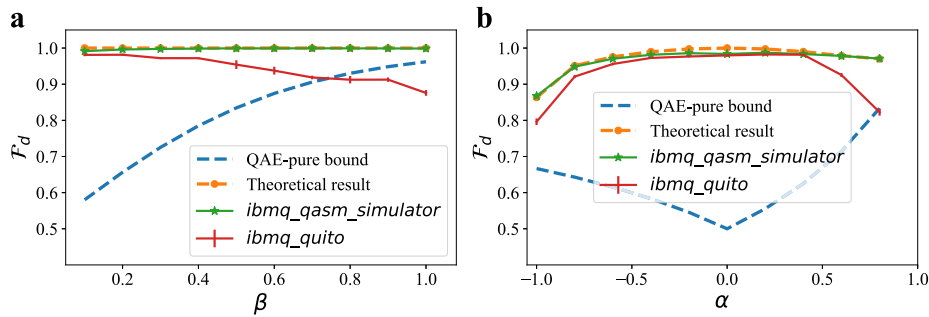
Fig. 9 | Results for QAEs on *ibmq_qasm_simulator* and *ibmq_quito* for encoding and decoding 2-qubit states. a for thermal states, **b** for Werner states. $\mathcal{F}_d$ represents the fidelity between the initial state and the recovered state. The blue dashed line represents the upper bound of compression rate when only considering pure reference states[26]. The orange line represents the theoretical results of simulating QAEs on classical computers. The green line represents the numerical results obtained from the IBM simulator *ibmq_qasm_simulator*. The red line with error bars represents the average results of 6 times together with standard deviation when quantum circuits are sampled with 8192 shots on IBM Quantum computer *ibmq_quito*.

practical applications. More work remains to be investigated in the future. For example, other forms of mixed reference states are worthy of further exploration. Imperfections in quantum system models are not considered in this work. Our future work will also include general quantum channels to deal with decoherence for mixed quantum states.

Methods

Quantum control model

Here, we use the density matrix $\rho(t)$ (which is a Hermitian, positive semi-definite matrix satisfying $\text{Tr}(\rho(t)) = 1$) to describe the state of a closed quantum system. The evolution equation for $\rho(t)$ can be described by the quantum Liouville equation[34]

$$i\dot{\rho}(t) = \rho(t)H(t) - H(t)\rho(t). \tag{9}$$

When we use control fields $\{u_j(t)\}_{j=1}^M$ to manipulate the system, the system Hamiltonian in Eq. (9) can be divided into two parts, i.e., $H(t) = H_0 + H_c(t) = H_0 + \sum_{j=1}^M u_j(t)H_j$, where H_0 is the time-independent free Hamiltonian of the system, $H_c(t)$ is the control Hamiltonian representing the interaction of the system with the control fields. For such a control system, its solution is given as $\rho(t) = U(t)\rho_0 U^{\dagger}(t)$ with $U(0) = I$, where the propagator $U(t)$ is formulated as follows:

$$\frac{d}{dt}U(t) = -i\left(H_0 + \sum_{j=1}^M u_j(t)H_j\right)U(t). \tag{10}$$

For the compression task, we consider spin chain models with

$$H_0 = \sum_{i=0}^{n_A+n_B-1} \sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \sigma_i^z \sigma_{i+1}^z.$$

Chains with Heisenberg coupling are known to be controllable given at least two noncommuting controls acting on the first or the last spin in the chain[35,36], we exert control fields on the first two qubits towards X and Y directions[37], with the control Hamiltonian as

$$H_c = \sum_{k=0,1} u_k^x \sigma_k^x + u_k^y \sigma_k^y.$$

As such, there are four control fields to be designed. We use piece-wise control fields, which means that the total control time T is equally divided into different periods, with each having $dt = T/N$ duration times. In this work, the total control time $T = 20$ is equally divided into 100 pieces. The

bound of control fields is set as $[-10, 10]$. Then, the encoding map for QAEs can be obtained by $U_e = U(T)$ following Eq. (10).

Training QAEs using learning algorithms

In this work, the training of QAEs is reduced to searching for an optimal U_e that maximizes $\Phi(w) = wJ_{\text{pure}} + (1 - w)J_{\text{qmi}}, w \in [0, 1]$. After the training is completed, injecting a mixed state ρ_{ref} to the decoder helps maintain the entropy consistency between the initial state ρ_0 and the recovered state $\rho_f = U_e^{\dagger}(\rho_{\text{latent}} \otimes \rho_{\text{ref}})U_e$. Finally, the overlap between the recovered states and the original states is measured to evaluate the efficiency of QAEs. Denote the parameters for the encoder as a vector $\boldsymbol{\theta}$. The procedure of QAE-qmi using mixed reference states is as follows:

1. Randomly initialize $\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ represents the control fields for the systems
2. Apply $U_e(\boldsymbol{\theta})$ to the initial states ρ_0
3. Measure ρ_{trash} and ρ_{latent} and compute the cost function $\Phi(w) = wF(\rho_{\text{trash}}, |0\rangle\langle 0|) - (1 - w)\mathcal{I}(U_e\rho_0 U_e^{\dagger})$
4. Perform the optimization of $\Phi(w)$ using a learning algorithm and obtain a better control parameter $\boldsymbol{\theta}$
5. Repeat steps 2-4 until convergence
6. Report the classical information $\boldsymbol{\theta}$ and store the latent state ρ_{latent}
7. Determine a suitable reference state ρ_{ref} using different strategies (e.g., $\rho_{\text{ref}} = \rho_{\text{trash}}$ or $\rho_{\text{ref}} = \rho_{\text{mix}}$) and prepare the reference state
8. Perform $U_e^{\dagger}(\boldsymbol{\theta})$ on the combined state $\rho_{\text{latent}} \otimes \rho_{\text{ref}}$ and obtain the recovered state as ρ_f

The key is to optimize the cost function of Φ using learning algorithms. Evolutionary strategy (ES) methods exhibit an advantage in exploring unknown environments in games[38] and have been applied in optimizing quantum control issues[39]. The comparison results in our previous work suggest that ES has the potential to optimize QAEs towards the theoretical upper bounds with high efficiency[26]. In this work, we utilize ES to optimize the cost function $\Phi(w)$.

ES is a black-box optimization method that utilizes heuristic search procedures inspired by natural evolution. At every iteration ("generation"), a population of parameter vectors ("genotypes") is perturbed ("mutated") and their objective function value ("fitness") is evaluated. The highest-scoring parameter vectors are then recombined to form the population for the next generation, and this procedure is iterated until convergence[38]. The detailed description for the ES method is provided in Supplementary Method 1.

It is worth noting that, the initialization process in Step 1 can be formulated as $\boldsymbol{\theta} = \mathbf{u}_{\min} + \text{Rand}[0, 1](\mathbf{u}_{\max} - \mathbf{u}_{\min})$, where $\text{Rand}[0, 1]$ is a function to generate random numbers uniformly distributed between 0 and 1 to meet the physical restriction of control fields. In addition, boundary checks and resetting values are required for every step that involves new

parameters to guarantee that newly generated parameters lie in the constrained field. For the parameter setting of the ES method, we set the population size as NP = 40 for 2-qubit states and NP = 50 for 4-qubit states. The perturbation factor is set as $\delta = 0.01$. The learning rate is set as $\chi_1 = 0.5$. The momentum factor is set as $\chi_2 = 0.9$. The perturbation factor is decayed as $\delta \leftarrow 0.98\delta$ every 100 training iterations.

Data availability

References

1. Biamonte, J. et al. Quantum machine learning. *Nature* **549**, 195 (2017).
2. Dong, D., Chen, C., Li, H. & Tarn, T.-J. Quantum reinforcement learning. *IEEE Trans. Syst. Man. Cybern, Part B (Cybernetics)* **38**, 1207–1220 (2008).
3. Huang, H.-Y. et al. Power of data in quantum machine learning. *Nat. Commun.* **12**, 2631 (2021).
4. Cerezo, M., Verdon, G., Huang, H.-Y., Cincio, L. & Coles, P. J. Challenges and opportunities in quantum machine learning. *Nat. Comput. Sci.* **2**, 567–576 (2022).
5. Niu, M. Y., Boixo, S., Smelyanskiy, V. N. & Neven, H. Universal quantum control through deep reinforcement learning. *npj Quantum Inf.* **5**, 33 (2019).
6. Li, J.-A. et al. Quantum reinforcement learning during human decision-making. *Nat. Hum. Behav.* **4**, 294–307 (2020).
7. Dong, D. & Petersen, I. R. Quantum estimation, control and learning: opportunities and challenges. *Annu. Rev. Control.* **54**, 243–251 (2022).
8. Pu, Y. et al. Variational autoencoder for deep learning of images, labels and captions. In *Adv. Neural Inform. Process. Syst.*, 2352–2360 (2016).
9. Bartůšková, L. et al. Optical implementation of the encoding of two qubits to a single qutrit. *Phys. Rev. A* **74**, 022325 (2006).
10. Steinbrecher, G. R., Olson, J. P., Englund, D. & Carolan, J. Quantum optical neural networks. *npj Quantum Inf.* **5**, 60 (2019).
11. Lamata, L., Alvarez-Rodriguez, U., Martín-Guerrero, J. D., Sanz, M. & Solano, E. Quantum autoencoders via quantum adders with genetic algorithms. *Quantum Mach. Learn.: Sci. Technol.* **4**, 014007 (2018).
12. Ding, Y., Lamata, L., Sanz, M., Chen, X. & Solano, E. Experimental implementation of a quantum autoencoder via quantum adders. *Adv. Quantum Technol.* **2**, 1800065 (2019).
13. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).
14. Wan, K. H., Dahlsten, O., Kristjánsson, H., Gardner, R. & Kim, M. Quantum generalisation of feedforward neural networks. *npj Quantum Inf.* **3**, 36 (2017).
15. Romero, J., Olson, J. P. & Aspuru-Guzik, A. Quantum autoencoders for efficient compression of quantum data. *Quantum Sci. Technol.* **2**, 045001 (2017).
16. Bravo-Prieto, C. Quantum autoencoders with enhanced data encoding. *Mach. Learn.: Sci. Technol.* **2**, 035028 (2021).
17. Pepper, A., Tischler, N. & Pryde, G. J. Experimental realization of a quantum autoencoder: the compression of qutrits via machine learning. *Phys. Rev. Lett.* **122**, 060501 (2019).
18. Huang, C.-J. et al. Realization of a quantum autoencoder for lossless compression of quantum data. *Phys. Rev. A* **102**, 032412 (2020).
19. Bondarenko, D. & Feldmann, P. Quantum autoencoders to denoise quantum data. *Phys. Rev. Lett.* **124**, 130502 (2020).
20. Achache, T., Horesh, L. & Smolin, J. Denoising quantum states with quantum autoencoders–theory and applications. *Preprint at* [https://arxiv.org/pdf/2012.14714] (2020).
21. Zhang, X.-M. et al. Generic detection-based error mitigation using quantum autoencoders. *Phys. Rev. A* **103**, L040403 (2021).
22. Du, Y. & Tao, D. On exploring practical potentials of quantum autoencoder with advantages. *Preprint at* [https://arxiv.org/pdf/2106.15432] (2021).
23. Srikumar, M., Hill, C. D. & Hollenberg, L. C. Clustering and enhanced classification using a hybrid quantum autoencoder. *Quantum Sci. Technol.* **7**, 015020 (2021).
24. Zhang, H. et al. Resource-efficient high-dimensional subspace teleportation with a quantum autoencoder. *Sci. Adv.* **8**, 9783 (2022).
25. Mangini, S. et al. Quantum neural network autoencoder and classifier applied to an industrial case study. *Quantum Mach. Intell.* **4**, 13 (2022).
26. Ma, H. et al. On compression rate of quantum autoencoders: Control design, numerical and experimental realization. *Automatica* **147**, 110659 (2023).
27. Cao, C. & Wang, X. Noise-assisted quantum autoencoder. *Phys. Rev. Appl.* **15**, 054012 (2021).
28. Pivoluska, M. & Plesch, M. Implementation of quantum compression on IBM quantum computers. *Sci. Rep.* **12**, 5841 (2022).
29. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).
30. Wilde, M. M. From classical to quantum Shannon theory. *Preprint at* [https://arxiv.org/pdf/1106.1445] (2011).
31. Watrous, J. *The Theory of Quantum Information* (Cambridge University Press, 2018).
32. Lyons, D. W., Skelton, A. M. & Walck, S. N. Werner state structure and entanglement classification. *Adv. Math. Phys.* **2012** (2012).
33. Qiskit library for quantum state tomography, https://qiskit.org/ecosystem/experiments/stubs/qiskit_experiments.library.tomography.statetomography.html (2021).
34. Dong, D. & Petersen, I. R. Quantum control theory and applications: a survey. *IET Control Theory Appl.* **4**, 2651–2671 (2010).
35. Burgarth, D., Bose, S., Bruder, C. & Giovannetti, V. Local controllability of quantum networks. *Phys. Rev. A* **79**, 060305 (2009).
36. Wang, X., Burgarth, D. & Schirmer, S. Subspace controllability of spin-1/2 chains with symmetries. *Phys. Rev. A* **94**, 052319 (2016).
37. Dong, D. & Petersen, I. R. *Learning and Robust Control in Quantum Technology* (Springer Nature, Switzerland AG, 2023).
38. Salimans, T., Ho, J., Chen, X., Sidor, S. & Sutskever, I. Evolution strategies as a scalable alternative to reinforcement learning. *Preprint at* [https://arxiv.org/pdf/1703.03864] (2017).
39. Shir, O. M. & Bäck, T. Niching with derandomized evolution strategies in artificial and real-world landscapes. *Nat. Comput.* **8**, 171–196 (2009).

Author contributions

H.M., D.D., and I.R.P. developed the theory. H.M. performed the numerical simulation and analyzed data with the assistance of D.D. and I.R.P. H.M.,

G.J.M., and L.C.L.H. designed the experiment. H.M. and G.J.M. analyzed the experimental data with the help of other authors. All authors contributed in writing the paper.

Competing interests

The authors declare no competing interests.

Additional information