OPTICS

Resource-efficient high-dimensional subspace teleportation with a quantum autoencoder

Hui Zhang[1]†, Lingxiao Wan[1]†, Tobias Haug[2], Wai-Keong Mok[3], Stefano Paesani[4,5], Yuzhi Shi[6]*, Hong Cai[7], Lip Ket Chin[1], Muhammad Faeyz Karim[1], Limin Xiao[8], Xianshu Luo[9], Feng Gao[9], Bin Dong[9], Syed Assad[10], M. S. Kim[2], Anthony Laing[5], Leong Chuan Kwek[1,3,11]*, Ai Qun Liu[1]*

Quantum autoencoders serve as efficient means for quantum data compression. Here, we propose and demonstrate their use to reduce resource costs for quantum teleportation of subspaces in high-dimensional systems. We use a quantum autoencoder in a compress-teleport-decompress manner and report the first demonstration with qutrits using an integrated photonic platform for future scalability. The key strategy is to compress the dimensionality of input states by erasing redundant information and recover the initial states after chip-to-chip teleportation. Unsupervised machine learning is applied to train the on-chip autoencoder, enabling the compression and teleportation of any state from a high-dimensional subspace. Unknown states are decompressed at a high fidelity (~0.971), obtaining a total teleportation fidelity of ~0.894. Subspace encodings hold great potential as they support enhanced noise robustness and increased coherence. Laying the groundwork for machine learning techniques in quantum systems, our scheme opens previously unidentified paths toward high-dimensional quantum computing and networking.

INTRODUCTION

Data compression is a ubiquitous process that we use daily to transfer large audio and video files. Autoencoders serve as efficient means for large dataset compression (*1*). Given a training set, they work by learning an encoder map from the input data to a latent space and a decoder map from the latent space to the input space. Generalizing this learning process to quantum systems requires considering some fundamental differences between classical and quantum information. For example, while a complete description of the state of a classical system can, in principle, be achieved with the information obtained by measuring it, for quantum systems, this would in general require measurements on an infinite amount of copies of the system (*2*). Nevertheless, quantum compression protocols have been shown to be possible based on the Schur-Weyl transformation (*3*). An example is the quantum autoencoder (*4*–*7*), which was proposed to efficiently compress a particular dataset of quantum states using machine learning, and its concept has been demonstrated using bulk optics (*8*, *9*).

Through data compression, the quantum autoencoder can become a useful tool for transmitting high-dimensional quantum information between remote parties. One possible application is quantum teleportation, where removing the redundancies in input quantum states enables encoding the quantum information in states with lower dimensionality, reducing the resource costs required to teleport them. Quantum teleportation (*10*) is a crucial protocol for the physical implementation of many quantum communication and quantum computation schemes, such as quantum relays (*11*), quantum repeaters (*12*), and linear optical quantum computing (*13*, *14*). To date, quantum teleportation has been experimentally generalized by different physical systems such as photons (*15*–*17*), microwaves (*18*), optical coherent states (*19*), atoms (*20*), and nuclear magnetic states (*21*). Meanwhile, notable progress has been made recently in high-dimensional quantum systems, including chip-based qutrit manipulation (*22*), and high-quality transport of high-dimensional photonic entanglement (*23*, *24*).

Typical information (e.g., language) is not uniformly distributed anywhere in a Hilbert space, and high-dimensional subspaces offer several key advantages, including enhanced noise robustness and increased coherence. To harness subspace encodings in applications such as quantum communication/cryptographic protocols (*25*, *26*), quantum memory (*27*), and fault-tolerant computing (*28*), high-dimensional subspace teleportation is desired. However, most teleportation implementations only consider two-dimensional quantum systems, e.g., the chip-to-chip teleportation of qubits and entanglement swapping in photonic chips (*29*), as the teleportation of high-dimensional quantum states remains challenging because of additional resource costs. Bennett *et al.* (*30*) first proposed the teleportation of an unknown d-dimensional state by replacing the maximally entangled state of two qubits with that of two entangled qudits and by replacing the Bell measurements with generalized measurements involving a set of maximally entangled orthonormal or mutually unbiased bases. However, practical implementations of high-dimensional Bell measurement face significant challenges because of a limitation of linear optics (*31*): For a d-dimensional system, none of the d^2 Bell states in a complete orthonormal basis can be identified without additional ancillary photons, except for $d = 2$. Several theoretical approaches suggest that it becomes possible if additional resources are allowed (*32*–*34*), leading to two experiments (not integrated)

[1]Quantum Science and Engineering Centre (QSec), Nanyang Technological University, 50 Nanyang Ave., Singapore 639798, Singapore. [2]Quantum Optics and Laser Science, Imperial College London, Exhibition Road, London SW72AZ, UK. [3]Centre for Quantum Technologies, National University of Singapore, Block S15, 3 Science Drive 2, Singapore 117543, Singapore. [4]Center for Hybrid Quantum Networks (Hy-Q), Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark. [5]Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1TH, UK. [6]Institute of Precision Optical Engineering, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China. [7]Institute of Microelectronics, A*STAR (Agency for Science, Technology and Research), Singapore 138634, Singapore. [8]School of Information Science and Technology, Fudan University, Shanghai 200433, China. [9]Advanced Micro Foundry, 11 Science Park Road, Singapore 117685 Singapore. [10]Department of Quantum Science, Centre for Quantum Computation and Communication Technology, The Australian National University, Canberra, ACT 2600, Australia. [11]National Institute of Education, 1 Nanyang Walk, Singapore 637616 Singapore.
*Corresponding author. Email: yuzhi_shi@163.com (Y.S.); cqtklc@nus.edu.sg (L.C.K.); eaqliu@ntu.edu.sg (A.-Q.L.)
†These authors contribute equally to this work.

showing the teleportation of path-encoded qutrits (*35*, *36*). Both of them work in full qutrit space but require many ancilla photons and have significantly decreased success probabilities (scaling approximately as $1/d^2$).

Here, we propose a quantum autoencoder–facilitated teleportation (QAFT) protocol, which compresses high-dimensional subspaces into a lower dimension and teleports them to a receiver, where the original state is reconstructed at significantly lower resource overheads. As proof of principle, a silicon photonic chip implementation of a quantum autoencoder is demonstrated as well as a fully integrated and chip-to-chip demonstration of QAFT with qutrit subspaces. In teleporting qutrit subspaces, our protocol does not require ancillary photons and has a success probability of ½. The generation, teleportation, and measurement of photon states, as well as the training of the encoder, are all performed on-chip. Unsupervised machine learning methods are applied to train the autoencoder, enabling any two-dimensional subspace of qutrits to be compressed to qubits by taking a finite sample of qutrits randomly. Single-shot training is implemented to ensure that no qutrit is copied or measured twice to meet the real application scenario of the no-cloning theorem. We can freely change the parameters of the autoencoding unitary and adjust them to any desired subspace. The quantum autoencoder achieves almost lossless compression on qutrits to qubits, with a reconstruction fidelity of ~0.971. After training, we implement teleportation and reconstruct the high-dimensional states at the receiver's end, showing a high teleportation fidelity of ~0.894. While a single instance of our QAFT can only teleport a lower-dimensional subspace rather than the full state space, these subspace encodings can have strong potential in the development of high-dimensional quantum networks as they support significantly higher noises than direct encodings (*25*, *26*). Furthermore, the compression protocol is not limited to qutrit teleportation and can be easily extendable to higher dimensions. Therefore, our scheme will come in handy for quantum internet (*37*), cryptography, and transferring quantum states (*38*) in distributed architectures by reducing the requirements on quantum memory, quantum communication channels, and computation resources.

RESULTS

The QAFT protocol and chip design

The overarching idea of QAFT is that by training an autoencoder, the input states can be compressed into lower-dimensional systems (qubits in our implementation), teleported, and reconstructed by the decoder, as shown in Fig. 1A. The transmitter and the receiver each hold half of an Einstein-Podolsky-Rosen (EPR) pair. An encoder is trained on-chip to achieve lossless compression from qutrits to qubits. At the transmitter, the initial qutrit is compressed into a qubit by the encoder, and the Bell state measurement (BSM) is performed. Depending on the BSM result and the settings of the encoder, the receiver will set up the decoder and reconstruct the initial qutrit from the teleported qubit. The training of the autoencoder is crucial to the success of teleportation.

Figure 1B depicts the multiphoton multiqubit chip implementing the QAFT protocol. Silicon photonics technology has matured into a versatile paradigm for different applications in quantum information processing, including photon pair generation (*39*–*42*), linear optical circuits (*43*, *44*), high-dimensional path entanglement (*45*), and error-protected quantum computation (*46*). Our chip is composed of two subchips with different functions, i.e., the transmitter and the receiver, which are coherently linked by a 10-m single-mode fiber after converting path-encoded qubits to polarization-encoded qubits. The transmitter circuit consists of a multiphoton generator, an EPR pair generator, the encoder of the QAFT, and a Bell projector for BSM, all of which are individually controllable and programmable. Single-wavelength pump light is divided and injected into three spiral sources to produce three pairs of signal photons (λ_s) and idler photons (λ_i) via spontaneous four-wave mixing (*39*). Asymmetric Mach-Zehnder interferometers (AMZIs) are applied in each pair of photon generation structures, one to filter the pump light and the other to separate the signal and idler photons (see transmission spectra in fig. S1). An EPR pair is generated from the top two pairs of photons after being routed through the crossing structure, as shown in Fig. 1B. See note S6A for the initial state generation. Photons are encoded as dual-rail qubits. Two paths of signal photons encode the first qubit of the EPR pair, and two idler paths encode the second qubit. The two qubits are maximally entangled in the form of the Bell state $|\phi^+\rangle$ with fidelities of 0.960 ± 0.004 (see fig. S1E). The idler photon of the bottom pair is heralded by measuring its signal photon and used to generate a path-encoded qutrit using the linear optical quantum circuit. The qutrit is then compressed to a qubit by the encoder, which is trained to retain all information on qubits modes such that, after teleportation, the initial qutrit can be decoded losslessly. At the transmitter, BSMs of the EPR qubit and the teleportee qubit are performed, yielding one of the two possible measurement outcomes $|\phi^+\rangle$ and $|\phi^-\rangle$, while $|\psi^+\rangle$ and $|\psi^-\rangle$ are indistinguishable by the linear-optical measurement results (see table S1). The central component of the BSM circuit is the reprogrammable two-qubit post-selected fusion gate, which entangles two qubits (that previously had never interacted), enabling the Bell state projection.

The first qubit of the EPR pair is transmitted through a 10-m optical fiber to the receiver. The path-encoded state in the waveguides to the polarization state in the optical fiber using a polarization rotator and splitter (PRS: see Materials and Methods). For higher-dimensional spaces, the path-encoded information can be encoded in time bin (*47*), orbital angular momentum (*48*), multicore fiber (*49*), or frequency information (*50*). The BSM outcome and the encoder settings are sent to the receiver through a classical information channel. The EPR qubit held by the receiver will be identical to the teleported qubit if the BSM result is $|\phi^+\rangle$. Alternatively, it can be recovered if the BSM result is $|\phi^-\rangle$ via a single-qubit rotation. Once the qubit is teleported, the decoder integrated on the receiver circuit will convert it back to the initial qutrit. As the encoder is a unitary transformation, the decoder is set up as the inverse of the encoder. The false-color micrograph of the fabricated chip is shown in Fig. 1C, in which the transmitter and receiver are integrated onto a single chip, forming a teleportation transceiver. If two transceiver chips perform a two-way communication, then each transceiver chip can either transmit qubits or receive qubits from the other transceiver chip. The experiment conducts fourfold photon detection, which includes a Bell pair, the heralding photon (as the trigger signal), and the teleportee photon, at a typical rate of ~30 counts/hour. Output photons are collected by the output fibers via polarization-independent edge couplers (see Materials and Methods) and then detected off-chip by six superconducting single-photon detectors (D1 to D6, ~0.85 efficiency). Among them, D1 to D4 are used in the BSMs, D5 is placed at the teleported photon, and D6 is used to detect the herald photon. During the training of the encoder, another single-photon detector (D7) is placed in the trash

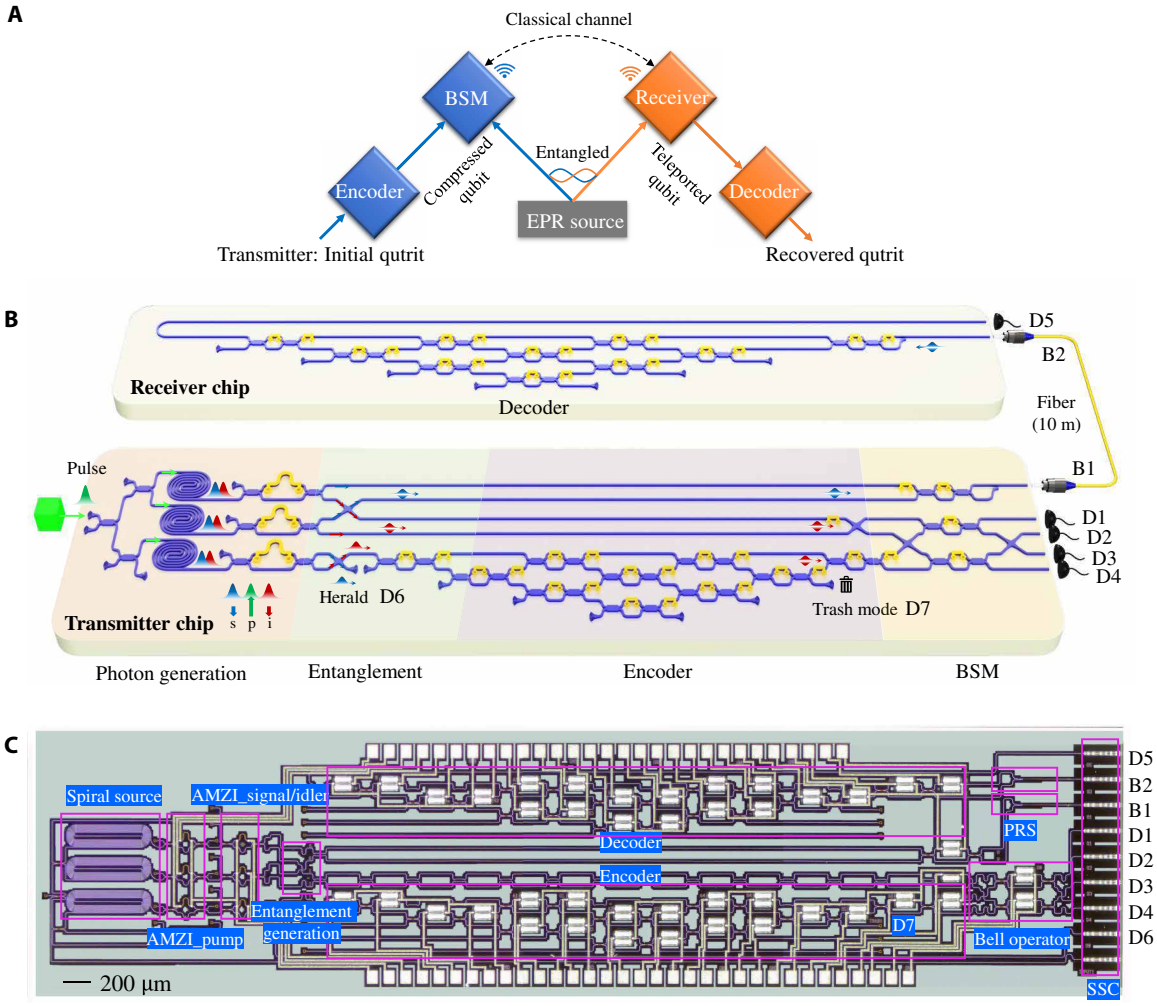

Fig. 1. The QAFT protocol in high dimensions. (**A**) Overview of the QAFT protocol. The transmitter and the receiver each hold one half of an EPR pair. An initial qutrit is compressed to a qubit by the trained encoder. The transmitter interacts the qubit with the half of an EPR pair and measures the two qubits in its possession. The BSM results and encoder settings are transmitted to the receiver via the classic channel. The decoder reconstructs the initial qutrit accordingly. (**B**) Design of the QAFT chip. The transmitter chip consists of a nondegenerate photon generator, entanglement, encoder, and Bell projector. The receiver chip integrates the decoder. The two subchips are coherently linked by a single mode fiber. At the transmitter, the qutrit is compressed by the encoder to a qubit, and Bell measurements are performed. At the receiver, the decoder reconstructs the initial qutrit state from the teleported qubit. p, pump; s, signal; i, idler. (**C**) False-color micrograph of the QAFT chip. The transmitter and receiver are integrated to form a two-way transceiver chip. Output photons (D1 to D6) are coupled by spot-size converters (SSCs) and detected off-chip by superconducting single-photon detectors. During the training of the encoder, another single-photon detector (D7) is placed at the trash mode to observe the photon occupancy. B1 and B2 are connected by a 10-m fiber, and a polarization rotator and splitter (PRS) are used to convert the path state in waveguide to the horizontal (vertical) state in optical fiber.

mode to observe the photon occupancy, as described in the autoencoder training section below. The number of used single-photon detectors does not increase with the dimension of the input qudit in the compression-teleportation-decompression protocol. Details on the device and setup are provided in note S5.

Implementation of the trainable quantum autoencoder

The quantum autoencoder is used to compress the quantum state from d-dimension to n-dimension ($d > n$). We implement an efficient autoencoder that reduces qutrits to qubits ($d = 3$ to $n = 2$) with its graphical representation shown in Fig. 2A. The basic architecture of the autoencoder is shown in Fig. 2B, which consists of an encoder (E) for compression and a decoder (D) for reconstruction. The intermediate nodes between two layers are called qubit modes, and those that are not used are called trash modes. The theory of quantum autoencoders is provided in note S4. There are two ways of training (*4*), based on trash state or reconstruction fidelity. The compression unitary U is obtained by minimizing their cost functions $L_t = 1 - F(|r\rangle, Tr_A U|\psi\rangle\langle\psi|U^\dagger)$ or $L_f = 1 - F(|\psi\rangle, U^\dagger(Tr_B U|\psi\rangle\langle\psi|U^\dagger)U)$, where the subscript A represents the qubit modes, B represents the trash mode, $|\psi\rangle$ is the input state, and $|r\rangle$ is the ideal trash state—a waveguide mode without photons; if the rank"? of the input state is smaller than the Hilbert space dimension of the compressed space, then both cost functions can be identically equal to zero (*28*). Therefore, in our case, minimizing photon occupancy at trash mode is sufficient to find the compression unitary. Lossless compression

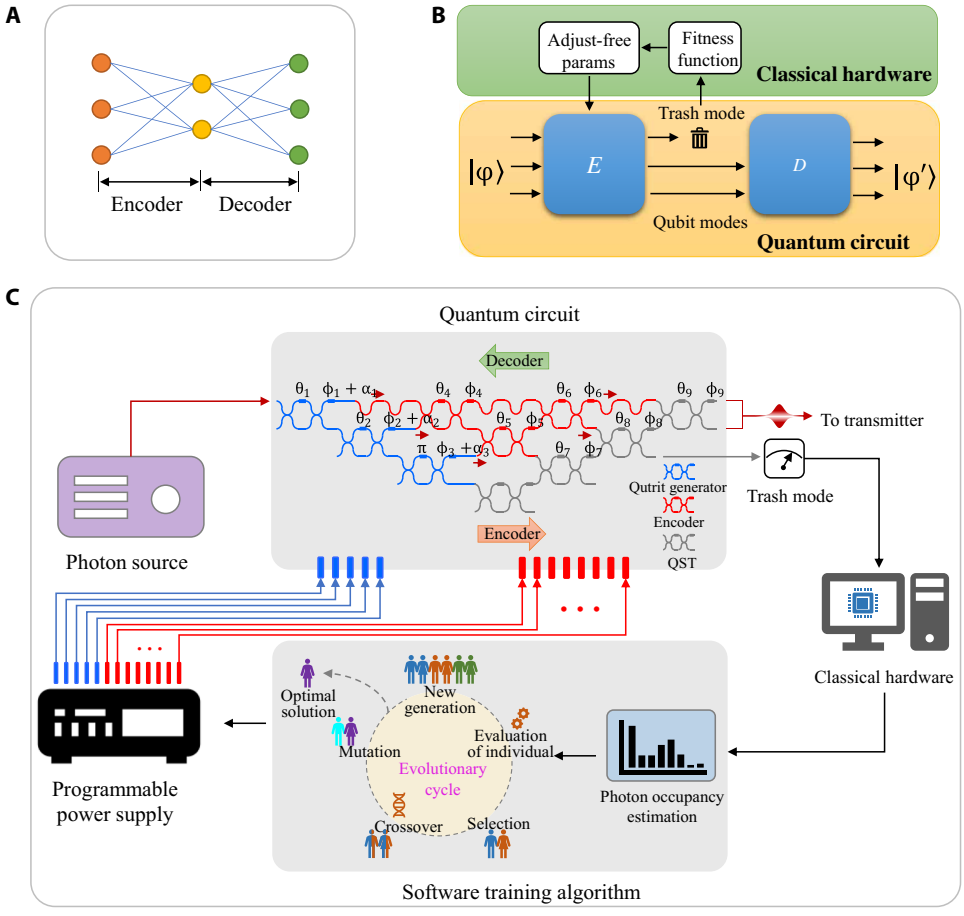

Fig. 2. Implementation of the trainable quantum autoencoder on a photonic chip. (**A**) Graphical representation of a 3-2-3 quantum autoencoder. Interconnected nodes represent space dimensions, and edges connecting adjacent layers represent the weights (a unitary transformation). The encoder shrinks high-dimensional data into lower-dimensional space, and then the decoder expands the space to reconstruct the original input states should the compression is lossless. (**B**) Architecture of the quantum autoencoder that consists of an encoder (*E*) and a decoder (*D*). The two modes that pass to the decoder are "qubit modes," and the one that is not used is "trash mode." The encoder is trained to retain all information on qubit modes while minimizing the occupation probability of the trash mode. The training parameters on the quantum circuit are optimized by a machine learning algorithm, according to the photon occupancy of the trash mode. (**C**) On-chip training process and the trainable encoder circuit. The qutrit generator (with five free parameters θ_{1-2} and ϕ_{1-3}) and the trainable encoder (with eight free parameters α_{1-3}, θ_{4-6}, and ϕ_{4-5}) are independently controlled. MZIs with different functionalities are marked by different colors, including the qutrit preparation (blue), encoder (red), and qutrit state tomography (gray). The decoder is the inverse of the encoder. The implementation of an evolutionary cycle involves the population initialization, the individual evaluation, and genetic operators (i.e., selection, crossover, and mutation). Individuals are evaluated by their photon occupancy at trash mode over randomly generated qutrits.

is achieved when the trash mode is unoccupied, and the initial qutrit can be reconstructed by the decoder. Conversely, any probability of measuring photons in the trash mode implies an imperfect compression and will result in the failure to reconstruct the initial qutrit when the trash mode is discarded.

Figure 2C depicts the quantum circuit of the autoencoder and our training system. The circuit realizes the functionalities of qutrit generator, encoder (decoder), and quantum state tomography. The qutrit generator produces arbitrary qutrits by programming three optical paths of the input photon using the three linked MZIs with five free parameters (θ_{1-2} and ϕ_{1-3}). The trainable encoder can be represented by a unitary matrix that has eight free parameters (α_{1-3}, θ_{4-6}, and ϕ_{4-5}). The on-chip encoder shows high stability and provides a simple way of controlling the unitary transformation by modulating the electrical powers of integrated heaters. Given that the encoder is unitary, the decoder is simply the inverse of the encoder. Quantum state tomography can be versatilely applied on input qutrits or reconstructed qutrits to validate the entire process.

An evolutionary optimization algorithm is used to iteratively update the encoder parameters, aiming to compress a subspace of qutrits. Given a training qutrit, the encoder should compress the qutrit into a qubit (that is pending for teleportation) while leaving the trash mode unoccupied. According to the no-cloning theorem, it is impossible to create independent and identical copies of any unknown quantum state. Practically, the qutrit generator does not know what state it produces and the sequence of output states. Therefore, a single-shot training is proposed where any qutrit will not be copied or sent twice, and the training algorithm will never measure the same qutrit twice. The photon occupancy is then estimated over a series of arbitrary input qutrits from the desired subspace. Each time after measuring a qutrit, the qutrit generator will be randomized such that the next qutrit is a different random qutrit. The single-shot training

balances incorporating new knowledge and retaining old knowledge during adjustment and has strong generalization (*51*, *52*). A clear advantage is visualized in fig. S2 by comparing the quality of compression without/with single-shot training. Single-shot training endows the encoder with stronger generalization capabilities whereby the obtained encoder can losslessly compress arbitrary qutrits from the subspace (where training instances are originally encoded) to qubits. After training, any further states from the subspace can be teleported and reconstructed at the receiver's end.

The genetic algorithm (GA) is adopted in the on-chip training. GA is inspired by the evolution theory in ecosystems, where individuals with higher adaptability have a higher probability of surviving and reproducing. As shown in Fig. 2C, an evolutionary cycle of the GA involves five stages, namely, population initialization, individual evaluation, selection, crossover, and mutation. The algorithm begins by creating a random initial population. Individuals in the initial population are implemented sequentially into the photonic chip, and their fitness values are measured. According to the fitness values, parents are selected to generate children. Some individuals with the best fitness values in the current population directly survive to the next population. Other children are produced either by the crossover—combining the vector entries of a pair of parents or by the mutation—making random changes to a single parent. The current population is then replaced with the children and will be reevaluated on the chip. The detailed settings of the GA implementation, including the population initialization and genetic operators (i.e., selection, crossover, and mutation) are provided in Materials and Methods.

Single-shot training of quantum autoencoder

Figure 3A shows the flow chart of the training process. The main components are the individually controllable photon activator, qutrit generator, and trainable encoder. Random qutrits are generated by a rotation matrix that maps random qubits to qutrits such that they belong to the same subspace and can be compressed by a common encoder. The subspace lies any two-dimensional space of the full Hilbert space. The training starts from the population initialization. Each population has 20 individuals, and each individual $\overrightarrow{\boldsymbol{P}_{\boldsymbol{k}}} = (I_1, I_2, \ldots I_8)$ is composed of eight free parameters, which are the electrical currents applied on the eight phase shifters, forming arbitrary unitary matrices. As an example, when the $\overrightarrow{\boldsymbol{P}_{\boldsymbol{1}}}$ in the initial population is applied to the encoder, a random qutrit will be generated at once for the evaluation of this individual. A total of 50 random qutrits are used to estimate the photon occupancy, with each qutrit measured only once. Each time after generating and measuring a qutrit, we will randomize the qutrit generator such that the next qutrit is a different random qutrit. The trash mode of the encoder is monitored by the counting logic, which outputs the required time (i.e., Δt) until the first click observed at the trash mode, which is inversely proportional to its photon occupancy (longer Δt leads to lower photon occupancy). If Δt is as long as the time required to detect the dark noise (which is $T = 10^{10}$ ps), then we can regard the photon occupancy of the trash mode as close to zero. The dark noise is the background counts recorded by the single-photon detector in the absence of any incident light. Accordingly, we design the fitness function of the GA as $f \propto 1 - e^{1-\Delta t/T}$ and the training objective as $f \to 0$. So far, the fitness evaluation of the individual $\overrightarrow{\boldsymbol{P}_{\boldsymbol{1}}}$ is achieved. Subsequently, the same process is repeated for the remaining individuals in the current population. If the fitness values satisfy the stopping criteria, then the best individual in the current population is returned as the optimal solution. Otherwise, the algorithm continues and the individuals in the current population are transformed by the genetic operators (i.e., selection, crossover, and mutation) to generate the offspring (a population of 20 new individuals). The offspring will be reevaluated until the algorithm converges.

Figure 3B shows the timing diagram of the counting logic. Two signal channels, a laser channel and a photon channel (monitored at the trash mode), are connected to the counting logic. We create a gate that starts at one of the laser pulses with a period of 10^{10} ps, which is the minimum required time to detect the dark noise. If no photon signal is detected during this period, then the photon signal is regarded as a dark noise. The gated stream in the recording is composed of clicks in both the laser channel and photon channel. The output of the time tagger is the time difference between the first photon click and the first laser click, i.e., $\Delta t = t_{\text{photon}} - t_{\text{laser}}$. The photon clicks after the first click will be discarded. If a photon click appears before the first laser click, then the photon will also be discarded.

Figure 3 (C to G) reports the results of training the quantum autoencoder in the single-shot setting. Figure 3C shows the evolution of the time spent until the first click is observed in the trash mode, showing that the time taken to observe the first click is usually very short for first-generation individuals. The final generation takes a significantly longer time, mostly approaching the dark noise level of 10^{10} ps. Occasionally, a good individual will appear early on, but we do not regard this as "convergence." For example, the first good individual (i.e., $\Delta t \approx \Delta t_{\text{dark noise}}$) is observed around the 10th generation; however, most of the others fall short of this level, showing large statistical fluctuations. On the contrary, in the final convergence, the time variations among individuals are greatly reduced (see the statistics in fig. S5), whereby we consider a convergence is reached. Figure 3D shows the evolution of the best and the average fitness value within each generation. The fitness value approaches 0 when Δt approaches $\Delta t_{\text{dark noise}}$. The fitness value of the final best encoder is 0.016, and the average fitness of the final population is 0.082. Figure 3 (E and F) shows the comparison of the initial and the final generation regarding the time taken to the first click and the fitness value, respectively. The statistical information of the 20 individuals in the initial and the final generation is shown in Fig. 3G. In the box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points that are not considered outliers. The "+" markers represent the outliers. Most individuals in the final generation have almost the same current value, with an average SD of 0.016 mA, implying a high-quality convergence. With the encoder setup, random qutrits are sent over for compression. The same training protocol that relies on a finite number of training samples is applied to different subspaces individually (see fig. S6). As seen from the tomography results of the initial qutrits and the compressed states, the device achieves a high-quality compression from initial qutrits to the compressed states, reducing the occupation probability of the trash mode to 0.023 ±0.011. When we reconstruct the qutrit by using the decoder (i.e., the inverse of the trained encoder), the fidelity $F = \langle \varphi | \rho | \varphi \rangle$ between the initial qutrit $|\varphi\rangle$ and the reconstructed density matrix ρ is 0.971 ± 0.013.

Theoretically, the compression can be completely lossless through training, with zero photon occupancy at the trash mode. Experimentally, factors that limit the quality of convergence arise from the hardware and optimization algorithms. Hardware limitations are inherently insurmountable, e.g., the visibility of MZI (>98% in our

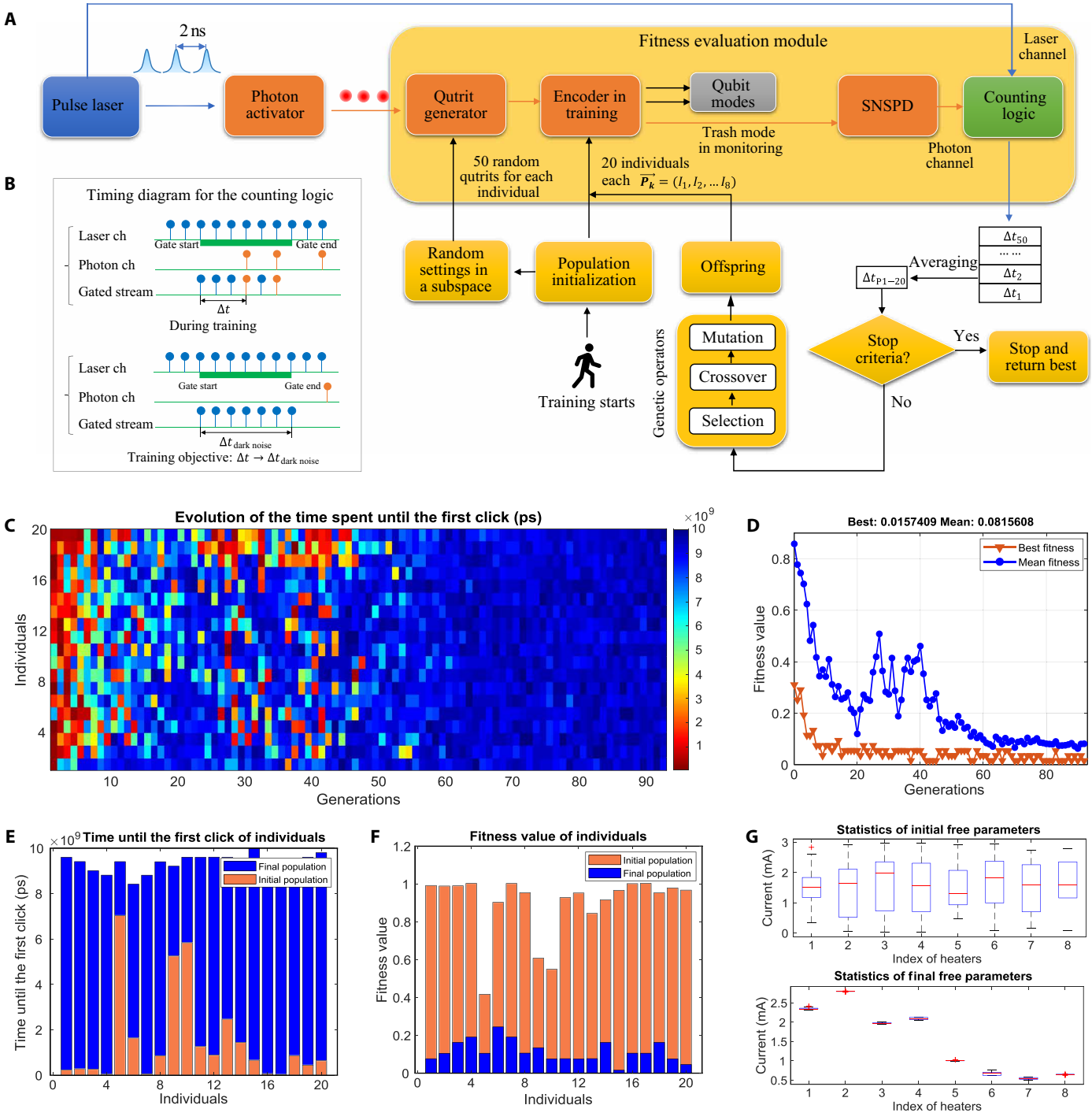

Fig. 3. Single-shot training of the quantum autoencoder. (**A**) Flowchart of single-shot training, which starts from population initialization. Individuals are evaluated on chip by the photon occupancy, which is inversely proportional to the time taken (i.e., Δt) for the first click observed at trash mode. (**B**) Timing diagram of the counting logic. Two signal channels, the laser channel and the photon channel (at the trash mode), are monitored. A gate with a period of $T = 10^{10}$ ps (i.e., the time required to detect dark noise) is created from one of the laser pulses. The time difference Δt between the first photon click and the first laser click is returned as the output. The training objective is $\Delta t \rightarrow T$, so that the fitness $f \propto 1 - e^{\Delta t/T}$ can approach $f = 0$. (**C**) The evolution of the time spent until the first click at the trash mode, which intuitively shows the convergence. Most individuals in the final generation achieve photon occupancy comparable to dark noise. (**D**) Evolution of the best and average fitness value. The fitness value of the final best encoder is 0.016, and the average fitness of the final population is 0.082. (**E** and **F**) Contrast of the time taken for the first click and fitness value in the initial and final generation. (**G**) Statistics of the 20 individuals in the initial and the final generation, respectively. Most individuals in the final generation have almost the same current value, with an average SD of 0.016 mA.

chip). Algorithm factors can be optimized by hyperparameters like the initialization of population, population size, and genetic operators. A common convergence problem is the premature convergence, where GA cannot produce better offspring although it has not reached the optimum. The premature convergence is typically caused by loss of population diversity, strong selective preference toward the best solution, and slight mutation. Our attempts include generating a diverse initial population and adopting the tournament selection operator (which selects the best from a few randomly selected individuals, rather than the best in the entire population). Other ways to alleviate this problem include removing similar individuals, increasing population size, and adopting adaptive mutation operators to regain the genetic variation.

Bell projection and teleportation of qutrits

In the teleportation protocol, an unknown quantum state can be transmitted to another location by locally collapsing the state and remotely reconstructing it. This requires access to Bell states and Bell measurements. A Bell projector is used to entangle initially separable qubits and measure qubits in the Bell basis. The schematic diagram for the Bell projector devised for dual-rail qubits is depicted in Fig. 1B (and fig. S9). We denote the three generated qubits from top to bottom as "the 1st qubit," "the 2nd qubit," and "the 3rd qubit" and denote the four detectors of the BSM from top to bottom as "D1," "D2," "D3," and "D4." By deliberately designing the circuitry, the Bell states $|\phi^{\pm}\rangle$ can be distinguished from the others, as shown in table S1. We distinguish $|\phi^{+}\rangle$ when observing joint clicks in {D1, D2} or {D3, D4} and distinguish $|\phi^{-}\rangle$ when observing joint clicks in {D2, D3} or {D1, D4}.

Chip-to-chip teleportation is implemented while ensuring the preservation of coherent teleportation between two chips. The $|\phi^{+}\rangle_{1,2}$ is created on the transmitter's end, and the 1st qubit is distributed to the receiver circuit via a 10-m optical fiber. The Bell measurement is performed at the 2nd qubit and the 3rd qubit on the transmitter's end, projecting the state into the $|\phi^{+}\rangle$ basis. This process achieves the teleportation of the 3rd qubit's state to the 1st qubit. We remark that the states of the entangled channel after distribution remain highly coherent with negligible degradation of fidelity. We first validate the chip-to-chip teleportation by implementing teleportation of several single-qubit states (i.e., $|H\rangle$, $|V\rangle$, $|D\rangle$, $|A\rangle$, $|R\rangle$, and $|L\rangle$) between the transmitter and the receiver and show the experimental data in Fig. 4A. The quantum state tomography on the teleported qubits reports an average fidelity of $\overline{F_1} = 0.914 \pm 0.022$. This high fidelity manifests the feasibility of the chip in achieving chip-to-chip teleportation.

Last, we demonstrate the QAFT protocol with chip-to-chip teleportation by teleporting several randomly generated qutrits. The initial qutrit is compressed to a qubit by a lossless encoder obtained from training, and the qubit is teleported from the 3rd qubit to the 1st qubit. Then, at the receiver circuit, the teleported 1st qubit is reconstructed to the qutrit by the decoder. The decoder is built according to the inverse of the encoder. The implementation of state tomography of qutrits is provided in note S11. The mean fidelity between the density matrices of the initial qutrit and the transferred qutrit is $\overline{F_2} = 0.894 \pm 0.026$ as shown in Fig. 4B. All error bars refer to ±1 SD estimated from Poissonian photon-counting statistics through Monte-Carlo simulation. One of the reasons for the degradation in the teleportation fidelity from $\overline{F_1}$ to $\overline{F_2}$ can be attributed to the reconstruction ability of the encoder. Nevertheless, the high fidelity proves that the QAFT can successfully compress the input qutrits using the lossless encoder and reconstructs them through the decoder after long-distance teleportation.

DISCUSSION

In this work, GA is chosen over other evolutionary algorithms (e.g., simulated annealing algorithm and particle swarm optimization) because it achieves the best convergence with the minimal experimental time cost (see note S12). In addition, GA has great potential for quantum photonic applications thanks to its following features: (i) It is physics-agnostic, without requiring explicit knowledge about the chip structures nor the analytical models. (ii) GA takes the most basic chip variables for learning and adapts according to the real response of the chip. Therefore, factors that conventionally affect the accuracy of off-chip modeling can be incorporated into the training process and effectively eliminated. From the noise study (in note S13), the GA-based adaptive learning exhibits strong noise resilience to the dominant noise in practice like random noise and thermal crosstalk. (ii) Its gradient-free feature makes on-chip training feasible, as calculating gradients on a chip is prohibitive (*53*–*56*), for both symbolic differentiation and finite difference approximation (*57*). The shift rule (*58*) is inherently exact, but it relies on excessive measurements and suffers from shot-noise error and experimental noise, because the gradient needs to be estimated from the device with finite samples. An immediate application of GA is to optimize hybrid quantum-classical systems, e.g., variational quantum eigensolvers (*59*), where optimization is carried out by a classical optimizer and the cost function is evaluated on a parametrized quantum circuit and hybrid quantum computers (*60*, *61*). The scale of GA can reach 4 million free parameters (*62*).

In many scenarios of quantum technology, incoming quantum data are concentrated in a small subspace within a large Hilbert space. This typicality of quantum data is a common feature in many communication applications. Given a certain subspace of qutrits for encoding, experimentally, the average fidelity between states before encoding and after decoding is as high as 0.97, which proves that such subspace is to a good approximation typical according to Wilde's framework for quantum typicality (*63*). The autoencoder can also enhance the noise tolerance when considering noisy input states (fig. S17), suggesting that it improves error tolerance in terms of typicality compared to direct qubit transmission. Here, the autoencoder projects the incoming noisy density matrix onto its dominant eigenvector that recovers the ideal state (*64*). The autoencoder is scalable for qudits with higher dimensions, with a compression loss of no more than 0.01, and a reconstruction fidelity of 0.99 (note S2). The effect of most compression protocols is to remove the redundancy of input data, and the redundancy here comes from the qubit encoding in a higher-dimensional system. Encoding qubits in higher-dimensional space has several key advantages that are harnessed for many applications, including increasing coherence for quantum memory, higher phase sensitivity for quantum sensing, and enhanced noise resilience for fault-tolerant computation. The autoencoder can also serve to denoise quantum data. We present examples of practical applications in the following sections.

Multiple subspaces

The autoencoder unitary encodes a two-dimensional subspace of a qutrit into a qubit for teleportation. However, we can easily change

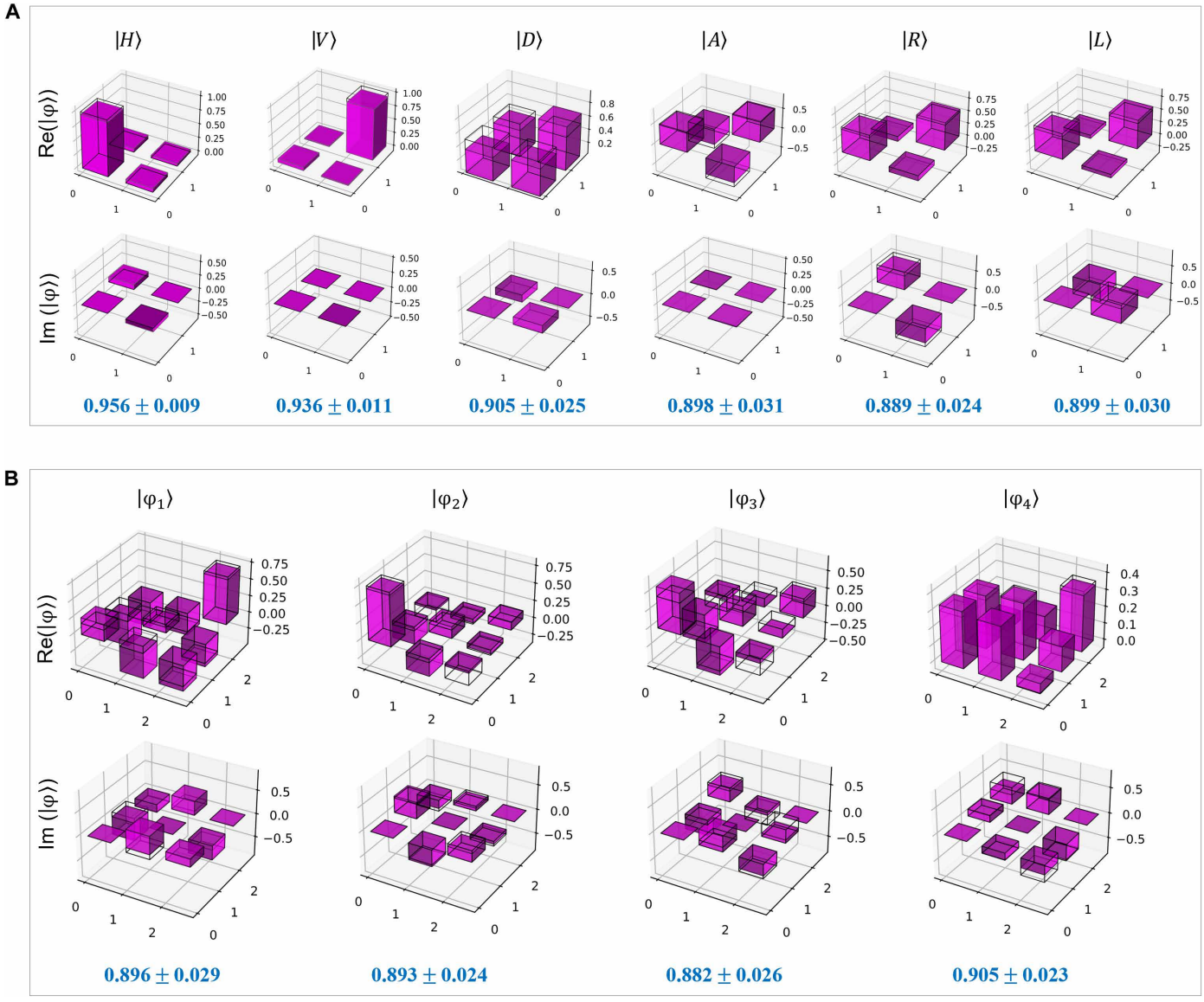

Fig. 4. BSM and quantum state tomography results for chip-to-chip teleportation. (**A**) Chip-to-chip teleportation results. A total of six elementary states were teleported from the transmitter to the receiver, respectively. The density matrices of the six states were constructed by full quantum state tomography on the receiver chip. Reconstructed density matrices are shown along with the measured fidelities, reporting a mean fidelity of 0.914 ± 0.022. (**B**) Qutrit encoding and decoding. The teleported qubit was decoded to reconstruct the initially generated qutrit on the receiver chip. The decoder is built according to the classical information transmitted to the receiver chip. Full quantum state tomography was applied to reconstruct the density matrix and report a mean fidelity of 0.894 ± 0.026. All error bars refer to ±1 SD estimated from Poissonian photon-counting statistics through Monte-Carlo simulation.

the parameters of the autoencoding unitary to adjust it to any desired subspace by using previous knowledge or learning from the given qutrits. As a practical example, the input qutrits are drawn from different classes of two-dimensional subspaces, with the encoder individually trained for each subspace using a small set of states. During operation, our setup acts as an adaptive compressor that can identify the appropriate autoencoder to be used by simply measuring the trash mode for a small sample of the states. This could potentially be applied to quantum communication/cryptographic protocols where the qutrit states are restricted to a set of multiple two-dimensional subspaces. For example, these scenarios have been recently proposed to build quantum networks with significantly enhanced noise robustness (*25*). In quantum key distribution, a protocol is introduced to use the subspaces of the high-dimensional system to certify entanglement in the presence of strong noise in the data (*26*). Our method greatly reduces the resource requirements in these schemes.

Storage of qubits in higher-dimensional systems

To process and store quantum information with high fidelity, the quantum information is often encoded within nontrivial subsystems of high-dimensional objects that offer convenient properties such as low coherence or easy manipulation operations. Often, only a small fraction of the total Hilbert space is used in these quantum memories. Examples include processing the information of a single qubit

within a system composed of two qubits with Heisenberg interaction (*65*) or storing qubit information with the ground state subspace of a multilevel system to suppress decoherence (*27*). Here, a common challenge is the transport of information from the quantum memory to another location where it is needed, often over long distances. To facilitate the transport, the quantum state has to be compressed, teleported, and decoded again. Our work is a proof-of-principle experiment on how to accomplish this task efficiently.

Error correction

Fault tolerance to noise is crucial to running quantum algorithms with potential advantages over classical computers. Fault tolerance is achieved by encoding the logical state of a single qubit within a configuration of multiple physical qubits and has been demonstrated experimentally recently (*66*). The logical state spans only a single qubit space. Thus, it can be compressed into a single qubit via an autoencoder and then teleported via our protocol. Because the teleportation between photonic chips can be achieved with low noise (*29*), the resource savings from compressing the quantum data likely outweigh the temporary removal of fault tolerance. Our proof-of-principle demonstration could potentially be used to exchange information between fault-tolerant quantum computers, which is essential for distributed quantum computing or the quantum internet.

Noise mitigation

Besides dimensionality reduction, quantum autoencoders can also be used to denoise quantum states. The autoencoder compresses incoming quantum states to a lower dimension, stripping away unneeded information. By training the circuit on a set of noisy states belonging to a subspace, the quantum autoencoder can learn to discard noise and reconstruct the original pure state with high fidelity. Proposals have shown high fidelities in reconstructing noisy GHZ states (*67*, *68*). Last, by using the information from the trash mode, one could further improve the performance of autoencoders subject to noise and enhance their noise resilience (*28*).

In conclusion, we propose and demonstrate a QAFT protocol that teleports arbitrarily generated qutrits on a photonic chip by training a quantum autoencoder that compresses the input qutrits to qubits. The proposal is generic and can be extended to higher dimensions. The generation, teleportation, and measurement of photon states, as well as the training of the encoder, are all performed on a single silicon chip. During training, the qutrits are arbitrarily generated from a subspace, and each qutrit undergoes only a single measurement. After the training, the encoder achieves almost lossless compression from the qutrit to qubit, with a reconstruction fidelity of ~0.971. Different encoders can be trained for qutrit from different subspaces. A teleportation fidelity of ~0.894 is achieved between the initial input qutrit and the teleported qutrit. The integrated silicon photonics technology endows the implementation of QAFT with high stability and easier scalability. The quantum autoencoder that allows data compression is beneficial to quantum communication by reducing the requirements on quantum memory, quantum communication channels, and the size of quantum gates. Our scheme uses an autoencoder for the teleportation of quantum states: It will come in handy for quantum internet, cryptography, and transfer quantum computer states. Our work also paves way for interdisciplinary quantum machine learning (*69*) and large-scale integrated photonic quantum technology (*40*) for quantum communication and quantum computing with high complexities.

MATERIALS AND METHODS

Experimental setup

The experimental setup is depicted in fig. S8. Following procedures previously described in (*70*), we bond the chip to a printed circuit board and provide independent control of each phase shifter by an electronic current driver with 1-kHz frequency and 12-bit resolution (Qontrol Devices Inc.). Laser pulses are generated by an Ultrafast Optical Clock device (PriTel Inc.) with a repetition rate of 500 MHz, a central wavelength of 1550.12 nm, and a bandwidth of 2 nm. The laser pulses are amplified by an Erbium-doped fiber amplifier and filtered off-chip via a wavelength-division multiplexer (WDM; >100-dB extinction ratio and 0.8-nm bandwidth) and a filter WDM (FWDM) that passes wavelength at 1550 ± 1 nm. The pump power is 0.4 mW. The typical rate of the fourfold photon detection is about 30 counts/hour. The pump, signal, and idler photons are located at λ_p = 1550.12 nm, λ_s = 1544.53 nm, and λ_i = 1555.75 nm, respectively. The input single pump light is coupled into the chip by a one-dimensional subwavelength grating coupler. A polarization controller is used to maximize the coupling efficiency of the fiber to chip. The input power is monitored via a photodetector placed at the reflection port of the FWDM. Output photons are coupled out of the chip via an eight-mode edge coupling fiber array and filtered via an FWDM and a WDM. The purpose of the filtering operation is to remove the residual pump photons and enhance the photon indistinguishability. The output photons enter an eight-channel (superconducting nanowire single-photon detectors) SNSPD (Photon Spot Inc.; dark counts of 100 Hz and 85% efficiency). Polarization controllers are placed before the SNSPDs. Counting logic from Swabian Instruments GmbH is used to count the single-photon events and process them on the fly. The counting device supports more than 40 million events/s. A Peltier controlled by Thorlabs TED200C is used to stabilize the temperature of the chip and reduce the heat fluctuations caused by the ambient temperature and the heat cross-talk within the chip.

Characterization of phase shifters

The phase shifters are used extensively to control the AMZIs and MZIs. The power dissipation caused by applying electrical power to the TiN heater will heat the underneath optical waveguide to change the refractive index and induce a phase shift. The *I*-*V* (current-voltage) characteristics of each heater were calibrated to observe the equivalent resistance. The characterization of each phase shifter was done by varying the applied current while measuring the optical power at the output port. The collected measurement data were fitted with a cosine function. An average R^2 (coefficient of determination) value of 0.99 was achieved in the fittings, and the average visibility was 99.85%. In this chip fabrication, the process of deep-etched trenches is adopted, and the heating efficiency is significantly improved. For heaters on MZIs, the 3-mA electrical current is sufficient for a phase shift range of 3π. For heaters on AMZIs, 7.5 mA induces a shift of more than one free spectral range (FSR) in the transmission. The FSRs for the pump filter (FSR_{pump}) and the signal/idler filter (FSR_{signal}) are 11.62 and 23.40 nm, respectively. The average extinction ratio for AMZI is more than 25 dB (*70*).

PRS and spot-size converter

The on-chip PRS is specially designed to realize the conversion of dual-rail qubit on-chip to polarization encoding in optical fiber. The path state $|0\rangle$ ($|1\rangle$) is first converted and combined to the transverse electric (TE) [transverse magnetic (TM)] mode $|TE0\rangle$ ($|TM0\rangle$) in

a single waveguide and then to the polarization state $|H\rangle$ ($|V\rangle$) in the optical fiber (see details in note S8). Because the waveguide cross section is designed for $|\mathrm{TE0}\rangle$ mode propagation, $|\mathrm{TM0}\rangle$ mode would experience a slightly higher loss. The phase shifter connected to PRS is thus calibrated by observing the interference fringe of intensity at the combined waveguide output. Suppose the ratio ($|\mathrm{TM0}\rangle/|\mathrm{TE0}\rangle$) of propagation loss is α, the intensity at the combined output is $\frac{1}{2}(1+\alpha^2)-\frac{1}{2}(1-\alpha^2)\cos\theta$, where θ is the phase shifter. According to the measured fringe of MZI for PRS on the encoder and decoder chips, the visibilities $V_{\mathrm{en}} = 3.0\%$ and $V_{\mathrm{de}} = 5.6\%$ are achieved, respectively, and the ratio α is estimated to be 0.97. To couple the two propagation modes simultaneously, we adopt suspended spot-size converters (SSCs) as the polarization-independent edge coupling structure. Using proper matching oils, the SSC can realize a coupling loss of less than −1.3 dB per facet with cleaved single-mode fiber. For TE and TM polarization, the polarization-dependent loss of less than 0.5 dB is achieved. The one-dimensional grating coupler is not adoptable here because it rejects the TM mode at an extinction ratio of ~20 dB. A polarization controller is placed between the transmitter circuit and the receiver circuit, and the polarization alignment of the two circuits is done by sending $|0\rangle$ and $|1\rangle$ as the calibration references (*70*).

Detailed settings of GA

The initial population is created with a uniform distribution, which is in the range of 0 to 3 mA (that allows a tuning range of 2π on the phase shifters). The initial population has 20 individuals, and each individual represents eight free parameters, which are the electrical currents supplied on the eight phase shifters of the trainable encoder. Then, the individuals of the initial population are implemented onto the photonic chip and their fitness values are acquired. According to the fitness values, we select qualified individuals as parents to generate children (that constitute the next population). We use a tournament-based selection operator, and the concrete process is as follows: First, the algorithm chooses some size of players at random (the size in our implementation is 4), and then the best individual is chosen, out of that set of four individuals, as a parent. Each parent is selected in this way, and an individual can be selected more than once as a parent. Then, we start to create new populations. The best of the current population is selected as the "elite" that directly survives to the new population. In the single-shot setting, it is not guaranteed that the elite will have the same fitness value when evaluating the new generation because the input qutrit state is constantly changing. Then, 80% of the new population (other than the elite) is created by the crossover operator. Ensuring that the created children still fall into the valid range of 0 to 3, the crossover operator is based on intermediate recombination, which creates children by taking a weighted average of the parents, e.g., child = parent1 + $\boldsymbol{w}\times$ (parent2 – parent1), where $\boldsymbol{w}$ is a vector of scaling factors chosen uniformly at random over the interval [−0.25, 0.75]. The other children are created by the mutation operator, which replaces a mutated variable with a new random variable in the bounds. The mutation provides genetic diversity and enables the algorithm to search a broader space. Because the selected size of the population is 20 per generation, the population is composed of 1 elite child, 15 crossover children, and 4 mutation children.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abn9783

REFERENCES AND NOTES

1. Y. Bengio, L. Yao, G. Alain, P. Vincent, Generalized denoising auto-encoders as generative models. arXiv:1305.6663 (2013).
2. M. A. Nielsen, I. L. Chuang, Quantum Computation and Quantum Information: 10th Anniversary Edition (Cambridge University Press, 2011).
3. L. A. Rozema, D. H. Mahler, A. Hayat, P. S. Turner, A. M. Steinberg, Quantum data compression of a qubit ensemble. *Phys. Rev. Lett.* **113**, 160504 (2014).
4. J. Romero, J. P. Olson, A. Aspuru-Guzik, Quantum autoencoders for efficient compression of quantum data. *Quantum Sci. Technol.* **2**, 045001 (2017).
5. K. H. Wan, O. Dahlsten, H. Kristjánsson, R. Gardner, M. Kim, Quantum generalisation of feedforward neural networks. *Npj Quantum Inf.* **3**, 36 (2017).
6. L. Lamata, U. Alvarez-Rodriguez, J. D. Martín-Guerrero, M. Sanz, E. Solano, Quantum autoencoders via quantum adders with genetic algorithms. *Quantum Sci. Technol.* **4**, 014007 (2019).
7. Y. Ding, L. Lamata, M. Sanz, X. Chen, E. Solano, Experimental implementation of a quantum autoencoder via quantum adders. *Adv. Quantum Technol.* **2**, 1800065 (2019).
8. C.-J. Huang, H. Ma, Q. Yin, J.-F. Tang, D. Dong, C. Chen, G.-Y. Xiang, C.-F. Li, G.-C. Guo, Realization of a quantum autoencoder for lossless compression of quantum data. *Phys. Rev. A* **102**, 032412 (2020).
9. A. Pepper, N. Tischler, G. J. Pryde, Experimental realization of a quantum autoencoder: The compression of qutrits via machine learning. *Phys. Rev. Lett.* **122**, 060501 (2019).
10. S. Pirandola, J. Eisert, C. Weedbrook, A. Furusawa, S. L. Braunstein, Advances in quantum teleportation. *Nat. Photonics* **9**, 641–652 (2015).
11. B. Jacobs, T. Pittman, J. Franson, Quantum relays and noise suppression using linear optics. *Phys. Rev. A* **66**, 052307 (2002).
12. Z.-D. Li, R. Zhang, X.-F. Yin, L.-Z. Liu, Y. Hu, Y.-Q. Fang, Y.-Y. Fei, X. Jiang, J. Zhang, L. Li, Experimental quantum repeater without quantum memory. *Nat. Photonics* **13**, 644–648 (2019).
13. R. Prevedel, P. Walther, F. Tiefenbacher, P. Böhi, R. Kaltenbaek, T. Jennewein, A. Zeilinger, High-speed linear optics quantum computing using active feed-forward. *Nature* **445**, 65–69 (2007).
14. S. Tanzilli, W. Tittel, M. Halder, O. Alibart, P. Baldi, N. Gisin, H. Zbinden, A photonic quantum information interface. *Nature* **437**, 116–120 (2005).
15. D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, A. Zeilinger, Experimental quantum teleportation. *Nature* **390**, 575–579 (1997).
16. X.-L. Wang, X.-D. Cai, Z.-E. Su, M.-C. Chen, D. Wu, L. Li, N.-L. Liu, C.-Y. Lu, J.-W. Pan, Quantum teleportation of multiple degrees of freedom of a single photon. *Nature* **518**, 516–519 (2015).
17. B. J. Metcalf, J. B. Spring, P. C. Humphreys, N. Thomas-Peter, M. Barbieri, W. S. Kolthammer, X.-M. Jin, N. K. Langford, D. Kundys, J. C. Gates, Quantum teleportation on a photonic chip. *Nat. Photonics* **8**, 770–774 (2014).
18. K. Fedorov, M. Renger, S. Pogorzalek, R. Di Candia, Q. Chen, Y. Nojiri, K. Inomata, Y. Nakamura, M. Partanen, A. Marx, Experimental quantum teleportation of propagating microwaves. *Sci. Adv.* **7**, eabk0891 (2021).
19. A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, E. S. Polzik, Unconditional quantum teleportation. *Science* **282**, 706–709 (1998).
20. S. Bose, P. Knight, M. Plenio, Proposal for teleportation of an atomic state via cavity decay. *Phys. Rev. Lett.* **83**, 5158–5161 (1999).
21. M. A. Nielsen, E. Knill, R. Laflamme, Complete quantum teleportation using nuclear magnetic resonance. *Nature* **396**, 52–55 (1998).
22. L. Lu, L. Xia, Z. Chen, L. Chen, T. Yu, T. Tao, W. Ma, Y. Pan, X. Cai, Y. Lu, Three-dimensional entanglement on a silicon chip. *Npj Quantum Inf.* **6**, 30 (2020).
23. N. H. Valencia, V. Srivastav, M. Pivoluska, M. Huber, N. Friis, W. McCutcheon, M. Malik, High-dimensional pixel entanglement: Efficient generation and certification. *Quantum* **4**, 376 (2020).
24. N. H. Valencia, S. Goel, W. McCutcheon, H. Defienne, M. Malik, Unscrambling entanglement through a complex medium. *Nat. Phys.* **16**, 1112–1116 (2020).
25. D. Bacco, J. F. Bulmer, M. Erhard, M. Huber, S. Paesani, Proposal for practical multidimensional quantum networks. *Phys. Rev. A* **104**, 052618 (2021).
26. M. Doda, M. Huber, G. Murta, M. Pivoluska, M. Plesch, C. Vlachou, Quantum key distribution overcoming extreme noise: Simultaneous subspace coding using high-dimensional entanglement. *Phys. Rev. Appl.* **15**, 034003 (2021).
27. A. I. Lvovsky, B. C. Sanders, W. Tittel, Optical quantum memory. *Nat. Photonics* **3**, 706–714 (2009).
28. C. Cao, X. Wang, Noise-assisted quantum autoencoder. *Phys. Rev. Appl.* **15**, 054012 (2021).
29. D. Llewellyn, Y. Ding, I. I. Faruque, S. Paesani, D. Bacco, R. Santagati, Y.-J. Qian, Y. Li, Y.-F. Xiao, M. Huber, Chip-to-chip quantum teleportation and multi-photon entanglement in silicon. *Nat. Phys.* **16**, 148–153 (2020).
30. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899 (1993).

31. J. Calsamiglia, Generalized measurements by linear elements. *Phys. Rev. A* **65**, 030301 (2002).
32. S. K. Goyal, T. Konrad, Teleporting photonic qudits using multimode quantum scissors. *Sci. Rep.* **3**, 3548 (2013).
33. H. Zhang, C. Zhang, X.-M. Hu, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, Arbitrary two-particle high-dimensional Bell-state measurement by auxiliary entanglement. *Phys. Rev. A* **99**, 052301 (2019).
34. C. Zhang, J. F. Chen, C. Cui, J. P. Dowling, Z. Y. Ou, T. Byrnes, Quantum teleportation of photonic qudits using linear optics. *Phys. Rev. A* **100**, 032330 (2019).
35. Y.-H. Luo, H.-S. Zhong, M. Erhard, X.-L. Wang, L.-C. Peng, M. Krenn, X. Jiang, L. Li, N.-L. Liu, C.-Y. Lu, Quantum teleportation in high dimensions. *Phys. Rev. Lett.* **123**, 070505 (2019).
36. X.-M. Hu, C. Zhang, B.-H. Liu, Y. Cai, X.-J. Ye, Y. Guo, W.-B. Xing, C.-X. Huang, Y.-F. Huang, C.-F. Li, Experimental high-dimensional quantum teleportation. *Phys. Rev. Lett.* **125**, 230501 (2020).
37. A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini, G. Bianchi, Quantum internet: Networking challenges in distributed quantum computing. *IEEE Netw.* **34**, 137–143 (2019).
38. D. Cozzolino, B. Da Lio, D. Bacco, L. K. Oxenløwe, High-dimensional quantum communication: Benefits, progress, and future challenges. *Adv. Quantum Technol.* **2**, 1900038 (2019).
39. J. W. Silverstone, D. Bonneau, K. Ohira, N. Suzuki, H. Yoshida, N. Iizuka, M. Ezaki, C. M. Natarajan, M. G. Tanner, R. H. Hadfield, On-chip quantum interference between silicon photon-pair sources. *Nat. Photonics* **8**, 104–108 (2014).
40. S. Paesani, Y. Ding, R. Santagati, L. Chakhmakhchyan, C. Vigliar, K. Rottwitt, L. K. Oxenløwe, J. Wang, M. G. Thompson, A. Laing, Generation and sampling of quantum states of light in a silicon chip. *Nat. Phys.* **15**, 925–929 (2019).
41. J. C. Adcock, C. Vigliar, R. Santagati, J. W. Silverstone, M. G. Thompson, Programmable four-photon graph states on a silicon chip. *Nat. Commun.* **10**, 3528 (2019).
42. J. Wang, D. Bonneau, M. Villa, J. W. Silverstone, R. Santagati, S. Miki, T. Yamashita, M. Fujiwara, M. Sasaki, H. Terai, Chip-to-chip quantum photonic interconnect by path-polarization interconversion. *Optica* **3**, 407–413 (2016).
43. J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, Universal linear optics. *Science* **349**, 711–716 (2015).
44. N. C. Harris, G. R. Steinbrecher, M. Prabhu, Y. Lahini, J. Mower, D. Bunandar, C. Chen, F. N. Wong, T. Baehr-Jones, M. Hochberg, Quantum transport simulations in a programmable nanophotonic processor. *Nat. Photonics* **11**, 447–452 (2017).
45. J. Wang, S. Paesani, Y. Ding, R. Santagati, P. Skrzypczyk, A. Salavrakos, J. Tura, R. Augusiak, L. Mančinska, D. Bacco, Multidimensional quantum entanglement with large-scale integrated optics. *Science* **360**, 285–291 (2018).
46. C. Vigliar, S. Paesani, Y. Ding, J. C. Adcock, J. Wang, S. Morley-Short, D. Bacco, L. K. Oxenløwe, M. G. Thompson, J. G. Rarity, Error-protected qubits in a silicon photonic chip. *Nat. Phys.* **17**, 1137–1143 (2021).
47. Y. He, X. Ding, Z.-E. Su, H.-L. Huang, J. Qin, C. Wang, S. Unsleber, C. Chen, H. Wang, Y.-M. He, Time-bin-encoded boson sampling with a single-photon device. *Phys. Rev. Lett.* **118**, 190501 (2017).
48. X. Cai, J. Wang, M. J. Strain, B. Johnson-Morris, J. Zhu, M. Sorel, J. L. O'Brien, M. G. Thompson, S. Yu, Integrated compact optical vortex beam emitters. *Science* **338**, 363–366 (2012).
49. E. Hugues-Salas, O. Alia, R. Wang, K. Rajkumar, G. T. Kanellos, R. Nejabati, D. Simeonidou, 11.2 tb/s classical channel coexistence with dv-qkd over a 7-core multicore fiber. *J. Light. Technol.* **38**, 5064–5070 (2020).
50. C. Reimer, M. Kues, P. Roztocki, B. Wetzel, F. Grazioso, B. E. Little, S. T. Chu, T. Johnston, Y. Bromberg, L. Caspani, Generation of multiphoton entangled quantum states by means of integrated frequency combs. *Science* **351**, 1176–1180 (2016).
51. C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, I.-H. Yeh, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. (2020), pp. 390–391.
52. J. He, R. Mao, Z. Shao, F. Zhu, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (IEEE, 2020), pp. 13926–13935.
53. T. W. Hughes, M. Minkov, Y. Shi, S. Fan, Training of photonic neural networks through in situ backpropagation and gradient measurement. *Optica* **5**, 864–871 (2018).
54. L. G. Wright, T. Onodera, M. M. Stein, T. Wang, D. T. Schachter, Z. Hu, P. L. McMahon, Deep physical neural networks trained with backpropagation. *Nature* **601**, 549–555 (2022).
55. H. Zhang, M. Gu, X. Jiang, J. Thompson, H. Cai, S. Paesani, R. Santagati, A. Laing, Y. Zhang, M. Yung, An optical neural chip for implementing complex-valued neural network. *Nat. Commun.* **12**, 457 (2021).
56. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017).
57. A. G. Baydin, B. A. Pearlmutter, A. A. Radul, J. M. Siskind, Automatic differentiation in machine learning: A survey. *J. Mach. Learn. Res* **18**, 1–43 (2018).
58. K. Mitarai, M. Negoro, M. Kitagawa, K. Fujii, Quantum circuit learning. *Phys. Rev. A* **98**, 032309 (2018).
59. A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. O'brien, A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
60. D. Zhu, N. M. Linke, M. Benedetti, K. A. Landsman, N. H. Nguyen, C. H. Alderete, A. Perdomo-Ortiz, N. Korda, A. Garfoot, C. Brecque, Training of quantum circuits on a hybrid quantum computer. *Sci. Adv.* **5**, eaaw9918 (2019).
61. K. Beer, D. Bondarenko, T. Farrelly, T. J. Osborne, R. Salzmann, D. Scheiermann, R. Wolf, Training deep quantum neural networks. *Nat. Commun.* **11**, 808 (2020).
62. F. P. Such, V. Madhavan, E. Conti, J. Lehman, K. O. Stanley, J. Clune, Deep neuroevolution: Genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. arXiv:1712.06567 (2017).
63. M. M. Wilde, *Quantum information theory* (Cambridge Univ. Press, 2013).
64. B. Koczor, The dominant eigenvector of a noisy quantum state. *New J. Phys.* **23**, 123047 (2021).
65. J. Levy, Universal quantum computation with spin-1/2 pairs and Heisenberg exchange. *Phys. Rev. Lett.* **89**, 147902 (2002).
66. L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, Fault-tolerant control of an error-corrected qubit. *Nature* , 1–6 (2021).
67. D. Bondarenko, P. Feldmann, Quantum autoencoders to denoise quantum data. *Phys. Rev. Lett.* **124**, 130502 (2020).
68. T. Achache, L. Horesh, J. Smolin, Denoising quantum states with quantum autoencoders--Theory and applications. arXiv:2012.14714 (2020).
69. L. Hu, S.-H. Wu, W. Cai, Y. Ma, X. Mu, Y. Xu, H. Wang, Y. Song, D.-L. Deng, C.-L. Zou, Quantum generative adversarial learning in a superconducting quantum circuit. *Sci. Adv.* **5**, eaav2761 (2019).
70. H. Zhang, L. Wan, T. Haug, W.-K. Mok, H. Cai, Y. Shi, L. K. Chin, M. Karim, L. Xiao, G. Lo, Chip-to-chip high-dimensional teleportation via a quantum autoencoder. Research Square 10.21203/rs.3.rs-809022/v1 (2021).

Acknowledgments

Funding: This work was supported by the Singapore Ministry of Education (MOE) Tier 3 grant (MOE2017-T3-1-001), Singapore National Research Foundation (NRF) National Natural Science Foundation of China (NSFC) joint grant (NRF2017NRF-NSFC002-014), Samsung GRC project and the UK Hub in Quantum Computing and Simulation, and part of the U.K. National Quantum Technologies Programme with funding from UKRI EPSRC grant EP/T001062/1.
Author contributions: H.Z., L.W., T.H., W.-K.M., and L.C.K. jointly conceived the idea. H.Z. and L.W. designed the chip and built the experimental setup. H.C., B.D., F.G., and X.L. fabricated the silicon photonic chip. H.Z. and L.W. performed the experiments. S.P., A.L., M.F.K., and L.X. assisted the setup and experiment. W.-K.M., T.H., S.A., M.S.K., and L.C.K. assisted with the theory. All authors contributed to the discussion of experimental results. L.C.K. and A.Q.L. supervised and coordinated all the work. H.Z., Y.S., L.K.C., and A.Q.L. wrote the manuscript with contributions from all coauthors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 12 January 2022
Accepted 23 August 2022
Published 7 October 2022
10.1126/sciadv.abn9783