Local-measurement-based quantum state tomography via neural networks

Tao Xin,[1,2,*] Sirui Lu,[2,*] Ningping Cao,[3,4,*] Galit Anikeeva,[4] Dawei Lu,[1,†] Jun Li,[1,‡] Guilu Long,[2,5] and Bei Zeng[3,4,1,§]

[1]*Shenzhen Institute for Quantum Science and Engineering and Department of Physics, Southern University of Science and Technology, Shenzhen 518055, China*
[2]*State Key Laboratory of Low-Dimensional Quantum Physics and Department of Physics, Tsinghua University, Beijing 100084, China*
[3]*Department of Mathematics & Statistics, University of Guelph, Guelph N1G 2W1, Ontario, Canada*
[4]*Institute for Quantum Computing, University of Waterloo, Waterloo N2L 3G1, Ontario, Canada*
[5]*Tsinghua National Laboratory of Information Science and Technology and The Innovative Center of Quantum Matter, Beijing 100084, China*

Quantum state tomography is a daunting challenge of experimental quantum computing even in moderate system size. One way to boost the efficiency of state tomography is via local measurements on reduced density matrices, but the reconstruction of the full state thereafter is hard. Here, we present a machine learning method to recover the full quantum state from its local information, where a fully-connected neural network is built to fulfill the task with up to seven qubits. In particular, we test the neural network model with a practical dataset, that in a 4-qubit nuclear magnetic resonance system our method yields global states via 2-local information with high accuracy. Our work paves the way towards scalable state tomography in large quantum systems.

Introduction— Quantum state tomography (QST) plays a vital role in validating and benchmarking quantum devices [1–5] because it can completely capture properties of an arbitrary quantum state. However, QST is not feasible for large systems because of its need for exponential resources. In recent years, there has been extensive research on methods for boosting the efficiency of QST. One of the promising candidates among these methods is QST via reduced density matrices (RDMs) [6–12]; because local measurements are convenient and accurate on many experimental platforms.

QST via RDMs is also a useful tool for characterizing ground states of local Hamiltonians. A many-body Hamiltonian H is k-local if $H = \sum_i H_k$, where each term H_i acts non-trivially on at most k-particles. In practical situations, we would mainly be interested in 2-local Hamiltonians, sometimes with certain interaction patterns, such as nearest-neighbor interactions on some lattices. It is known that the k-RDMs of a k-local Hamiltonian uniquely determine its unique ground state [11]. Therefore, for these ground states, one only needs k-local measurements (i.e., k-RDMs) for state tomography.

Although local measurements are efficient, in general, reconstructing the state from the measurement result is known to be hard [13]. To be more precise, it is easy to obtain the k-RDMs of any n-particle quantum state $|\psi\rangle$. However, even if $|\psi\rangle$ is uniquely determined by its k-RDMs, reconstructing $|\psi\rangle$ is hard. We remark that this is not due to the problem that $|\psi\rangle$ needs to be described by exponentially many parameters. In fact, in many cases, ground states of k-local Hamiltonians can be effectively represented by tensor product states [14].

The problem of state reconstruction has a natural connection to the regression problem of supervised learning. Regression analysis, in general, attempts to discover the relation between inputs and outputs, i.e., to recover the mathematical model. Machine learning techniques have been applied to QST in various cases, such as in Refs. [15, 16]. In our case, by knowing the Hamiltonian H, it is relatively easy to get the ground state $|\psi_H\rangle$ and then k-local measurements of $|\psi_H\rangle$. Therefore, the data for tuning our reverse engineering model is accessible, which allows us to practically realize QST through supervised learning. The data generation process is shown in Figure 1.

In this work, we build a fully-connected feedforward neural network for 4-qubit ground states of 2-local Hamiltonians. Our trained 4-qubit network not only predicts the test dataset with high fidelity but also reconstruct NMR experimental states accurately. We use the 4-qubit case to demonstrate the potential of using neural networks to realize QST via RDMs. The versatile framework of neural networks for ground states of k-local Hamiltonians could be extended to more qubits.

Theory—By the universal approximation theorem[17], every continuous function on the compact subsets of $\mathbb{R}^n$ can be approximated by a multi-layer feedforward neural network with a finite number of neurons, i.e., computational units. And by observing the relation between k-local Hamiltonian and lo-

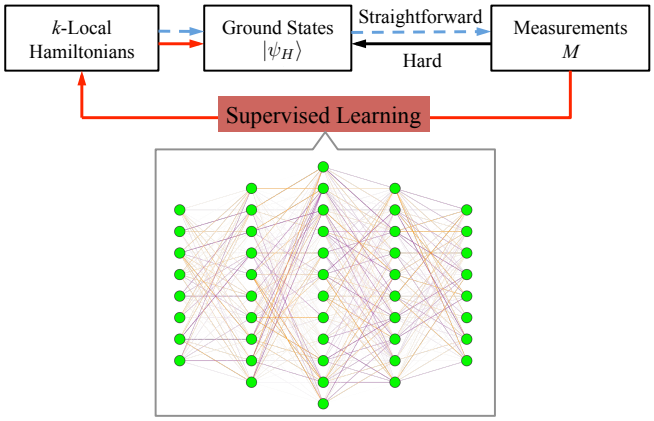


Figure 1. The blue dashed line is the method of generated training and testing data. The black arrow shows the direction of normal QST process. Red arrows demonstrate the procedure of our local tomography method. We first recover the Hamiltonian H through local measurement results $\mathbf{M}$ by using supervised learning neural network, then recover the ground state.

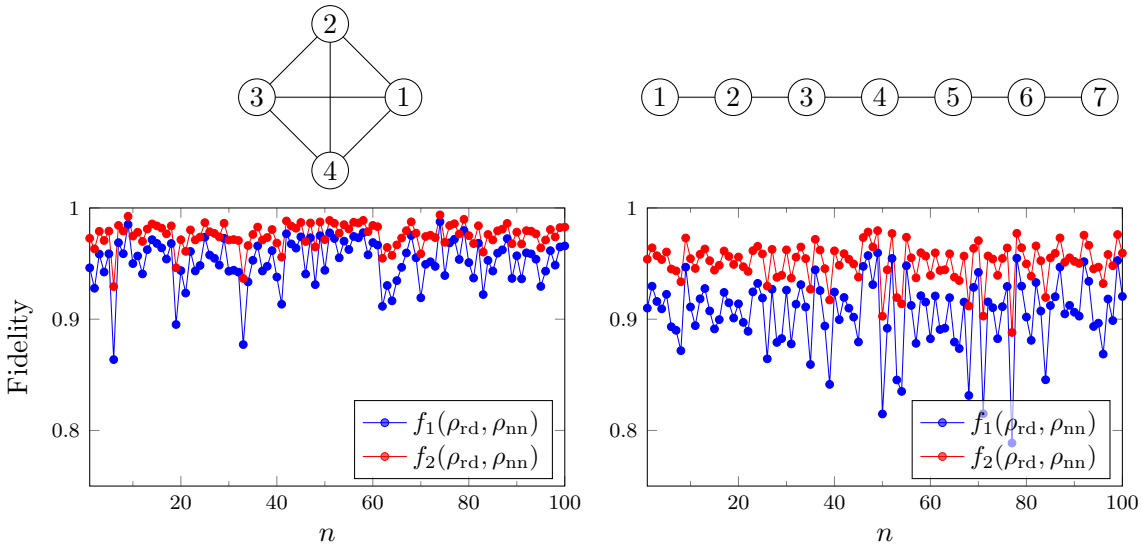


Figure 2. The left upper part is the configuration of our 4-qubit states. Each dot presents a qubit, and every qubit interacts with each other. The f_1 and f_2 of 100 random 4-qubit states $\rho_{\rm rd}$ and our neural network predictions $\rho_{\rm nn}$ are shown in the left. The configuration of our 7-qubit states is shown in the upper right part, which only nearest qubits have interactions. The f_1 and f_2 of 100 random 7-qubit states $\rho_{\rm rd}$ and our neural network predictions $\rho_{\rm nn}$ are shown in the right.

cal measurements of its ground state, as shown in Fig. 1, we are empowered to turn the tomography problem to a regression problem which fit perfectly into neural network framework.

In particular, we construct a deep neural network for 4-qubit ground states of 2-local Hamiltonians as follows:

$$H = \sum_{i=1}^{4} \sum_{1\le k\le 3} \lambda_k^{(i)} \sigma_k^{(i)} + \sum_{1\le i<j\le 4} \sum_{1\le n,m\le 3} \lambda_{nm}^{(ij)} \sigma_n^{(i)} \otimes \sigma_m^{(j)}, \tag{1}$$

where $\sigma_k, \sigma_n, \sigma_m \in \Delta$, $\Delta = \{\sigma_1 = \sigma_x, \sigma_2 = \sigma_y, \sigma_3 = \sigma_z, \sigma_4 = I\}$, $\lambda_k^{(i)}$ and $\lambda_{nm}^{(ij)}$ are constants of the Hamiltonian. The configuration of the ground states is illustrated in Fig. 2. The number of parameters of the local observables of ground states determines the amount of network input units. The input layer has 66 neurons, since the cardinality of the set of measurement results $\mathbf{M} = \{s_{m,n}^{(i,j)} : s_{m,n}^{(i,j)} = \mathrm{Tr}(\mathrm{Tr}_{(i,j)}\, \rho \cdot \sigma_m \otimes \sigma_n), 1 \le i < j \le 4, 1 \le n, m \le 4, n + m \ne 8\}$ is 66. Our network then follows three fully connected hidden layers, in which every neuron in the previous layer is connected to every neuron in the next layer, and information only passes forward. The number of output units equals the number of parameters of our 2-local Hamiltonian, which is 66 in our case.

Our training data consist of the 50000 randomly generated 2-local Hamiltonians as output and the corresponding local measurements of its ground states as input. The test data are formed by 5000 pairs of Hamiltonians and local measurement results.

We train the network by stochastic gradient descent algorithm with the cosine proximity, $\cos(\theta) = (x \cdot y)/(\|x\| \cdot \|y\|)$, as the loss function, where x is the prediction of the neural network and y is the desired output value. We find the cosine proximity function fits our scenario better than the more commonly chosen loss functions such as mean square error or mean absolute error. The reason is that parameters of the random Hamiltonian are close to zero. We also apply L_2 regularizers to avoid overfitting. Figure 1 illustrates our process.

Two different fidelities are used to measure the distance between random generated states $\rho_{\rm rd}$ and our neural network predictions $\rho_{\rm nn}$:

$$f_1(\rho_1, \rho_2) = \frac{\mathrm{tr}(\rho_1\rho_2)}{\sqrt{\mathrm{tr}(\rho_1^2)} \cdot \sqrt{\mathrm{tr}(\rho_2^2)}}, \tag{2}$$

$$f_2(\rho_1, \rho_2) = \mathrm{tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}. \tag{3}$$

Our machine learning process is implemented using Tensorflow[18]. We choose the optimizer having the best performance over almost all the built-in optimizers in Tensorflow: AdamOptimizer (Adaptive Moment Estimation)[19]. The batch size is 256.

After supervised learning on training data, our neural network is capable of predicting the 4-qubit output of the test set with fidelity f_1 of 95.3% and fidelity f_2 of 97.6%. Figure 2 illustrates the two fidelities between 100 random states $\rho_{\rm rd}$ and our neural network predictions $\rho_{\rm nn}$.

Our framework also applies to recover 7-qubits ground states of nearest 2-local Hamitionians. A similar neural network has been trained, predicting the 5000 random gendered testset with fidelity f_1 of 90.3% and fidelity f_2 of 94.9%. The configuration of our 7-qubit states and fidelities is shown in Fig. 2.

Experiment— To demonstrate the applications of our trained machine learning model in quantum simulation, we experimentally perform the dynamics of the random Hamilto-

nian and attempt to reconstruct the final quantum state by local RDMs using four-qubit nuclear magnetic resonance (NMR) platform [20–23]. The four-qubit sample is ^{13}C-labeled trans-crotonic acid dissolved in d6-acetone, where C_1 to C_4 are encoded as the four work qubits, and the rest spin-half nuclei are decoupled throughout all experiments. Figure. 3 presents the parameters and structure of this molecule. Under the weak-coupling approximation, the Hamiltonian of the system can be written as,

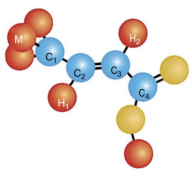
$$\mathcal{H}_{\text{int}} = \sum_{j=1}^{4} \pi(\nu_j - \nu_0)\sigma_z^j + \sum_{j<k,=1}^{4} \frac{\pi}{2} J_{jk}\sigma_z^j\sigma_z^k, \qquad (4)$$

where ν_j and J_{jk} are the chemical shifts and the J-coupling strengths, respectively. ν_0 is the reference frequency of ^{13}C channel in the NMR platform. All experiments were carried out on a Bruker AVANCE 400 MHz spectrometer at room temperature. Our experiments mainly include the following three steps: (i) Initialization. The pseudo-pure state for being the input of quantum computation $|0000\rangle$ is prepared. (ii) Evolution. Starting from the state $|0000\rangle$, we create the ground state of the random two-body Hamiltonian by applying the optimized shaped pulses. (iii) Measurement. We measure the two-body reduced density matrices and perform four-qubit quantum state tomography (QST) to estimate the quality of our implementations.

(i) Computational basis $|0\rangle^{\otimes n}$ is usually chosen as the input state for quantum computation. Most of the quantum systems do not start from such an input state, so the special initialization processing is necessary before realizing quantum circuits. In NMR, the sample stays in the Boltzmann distribution at room temperature,

$$\rho_{\text{thermal}} = \mathcal{I}/16 + \epsilon(\sigma_z^1 + \sigma_z^2 + \sigma_z^3 + \sigma_z^4), \qquad (5)$$

where $\mathcal{I}$ is identity and $\epsilon = 10^{-5}$ is the polarization. We can not directly use it as the input state for quantum computation, because such a thermal state is a highly-mixed state [24, 25]. We instead create a so-called pseudo-pure state (PPS) from the thermal state by using the spatial averaging technique [25–27], which consists of local unitary rotations and using z-gradient fields to destroy the undesired coherence. The form



	C1	C2	C3	C4
C1	-1705.5			
C2	41.64	-14558.1		
C3	1.48	69.78	-12330.5	
C4	7.06	1.18	72.36	-16764.1
T2	1.02	0.92	0.87	0.94

Figure 3. Molecular structure and Hamiltonian parameters of ^{13}C-labeled trans-crotonic acid. C_1, C_2, C_3 and C_4 are used as four qubits in the experiment, and M, H_1 and H_2 are decoupled throughout the experiment. In the table, the chemical shifts with respect to the Larmor frequency and J-coupling constants (in Hz) are listed by the diagonal and off-diagonal numbers, respectively. The relaxation timescales T_2 (in seconds) are shown at the bottom.

of 4-qubit PPS is,

$$\rho_{0000} = (1-\epsilon)\mathcal{I}/16 + \epsilon|0000\rangle\langle 0000|, \qquad (6)$$

Here, ρ_{0000} is also a highly-mixed state, but the identity part $\mathcal{I}$ does not evolve under any unitary operations nor influence the experimental signal. It means that we can only focus on the deviated part $|0000\rangle\langle 0000|$ and consider $|0000\rangle\langle 0000|$ as the initial state describing the dynamics of the quantum system. Finally, 4-qubit QST was performed to evaluate the quality of our PPS. We found that the fidelity between the perfect pure state $|0000\rangle$ and the experimentally measured PPS is about 98.7% by the definition $F(\rho,\sigma) = \text{Tr}(\rho\sigma)/\sqrt{\text{Tr}(\rho^2)\text{Tr}(\sigma^2)}$. It sets a solid ground for the subsequent experiments.

(ii) In this step, we prepare the ground state of the given Hamiltonian. The form of the considered Hamiltonian is chosen as,

$$H = \sum_{i=1}^{4}\sum_{k=x,y,z} \omega_l^i\sigma_k^i + \sum_{i,<j=1}^{4}\sum_{k,l=x,y,z} J_{kl}^{ij}\sigma_k^i\sigma_l^j. \qquad (7)$$

Here, the parameters ω_l^i and J_{kl}^{ij} mean the chemical shift and the J-coupling strength, respectively. In experiments, we create the ground states of different Hamiltonians by randomly changing the parameter set $(\omega_l^i, J_{kl}^{ij})$. For the given Hamiltonian, the gradient ascent pulse engineering (GRAPE) algorithm is adopted to optimize a radio-frequency (RF) pulse to realize the dynamical evolution from the initial state $|0000\rangle$ to the target ground state [28–31]. The GRARE pulses are designed to be robust to the static field distributions and RF inhomogeneity, and the simulated fidelity is over 0.99 for each dynamical evolution.

(iii) In principle, we only need to measure the two-body reduced density matrices (2-RDMs) to determine the original 4-qubit Hamiltonian through our trained network. Actually, we also perform 4-qubit QST which naturally includes the 2-RDMs after preparing these states [32–34], to evaluate the performance of our implementations. In NMR system, only single-coherence operators can be directly observed, for instance, only one qubit is in σ_x or σ_y and the rest qubits are in σ_z or I. Hence, we need the additional readout pulses before the signal is acquired when we attempt to measure other operators. Readout pulses will transfer the target operators to the single-coherence operators. The 17 readout pulses can fully measure all the elements of final density matrix:

$$\begin{aligned}\text{Pulse set}:[&XXXX, IIYY, YYXX, IIIY, XYXX,\\ &YXYI, IXYI, IIIX, XIYY, YXII, YYXY,\\ &XYXI, IIYX, IXIY, IIXI, IYIY] \end{aligned} \qquad (8)$$

It is so-called QST processing [35]. Here $X = \exp(-i\sigma_x\pi/4)$, $Y = \exp(-i\sigma_y\pi/4)$ and I is a 2×2 identity operation. Hence, we can estimate the quality of the experimental implementations by computing the fidelity between the target ground state ρ_{th} and the experimentally reconstructed density matrix. Considering that the ground state of

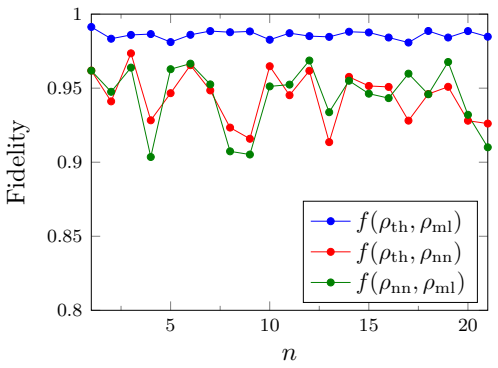


Figure 4. Fidelities of the density matrices $\rho_{\rm ml}$ and $\rho_{\rm nn}$. The blue dots mean the fidelities between the theoretical states $\rho_{\rm th}$ and the experimental states $\rho_{\rm ml}$. The green dots, $f(\rho_{\rm nn}, \rho_{\rm ml})$, are fidelities between our neural network predictions $\rho_{\rm nn}$ and the experimental states $\rho_{\rm ml}$ with the average fidelity over 94.5%.

the target Hamiltonian should be real, the experimental density matrix should also be real. We use a maximum likelihood approach to reconstruct the most likely pure state $\rho_{\rm ml}$ [36]. By reconstructing states $\rho_{\rm NN}$ merely based on the experimental 2-RDMs, the performance of our trained neural network can be evaluated by comparing the states $\rho_{\rm ml}$ with the states $\rho_{\rm nn}$.

In experiments, we created the ground state of the 21 randomized Hamiltonians in Eq. (7) and performed 4-qubit QST for them after the preparations. First, we present the fidelities of all the prepared density matrices in experiments. On average, the fidelity between the states $\rho_{\rm ml}$ and $\rho_{\rm th}$ is about 98.2%. Second, only 2-RDMs of these density matrices are used to reconstruct 4-qubit states by using our trained neural network framework. The fidelities $F(\rho_{\rm ml}, \rho_{\rm nn})$ are over 94.5%. Figure 2 shows the fidelity details of these density matrices. The results indicate that the original 4-qubit state can be efficiently reconstructed by our trained neural network which merely uses 2-RDMs, instead of the traditional QST.

Finally, we attempt to evaluate the confidence of the expected results by analyzing the potential error sources in experiments. The infidelity of the experimental density matrix is mainly caused by some primary aspects in experiments, including decoherence effects, imperfections of the PPS preparation, and imprecision of the optimized pulses. From a theoretical perspective, we numerically simulate the influence of the optimized pulses and the decoherence effect of our qubits and compare the fidelity computed in this manner with the ideal case to evaluate the quality of final density matrix. As a numerical result, an about 0.2% infidelity was created in average and the 1.2% error related to the infidelity of the initial state preparation. Additionally, the other errors can also be attributed to the infidelity such as imperfections in the readout pulses and spectral fitting.

Discussion— As a famous double-edged sword in experimental quantum computing, QST captures full information of quantum states on the one hand, while on the other hand, its implementation consumes a tremendous amount of resources. Unlike traditional QST that requires exponential number of experiments with the growth of system size, the recent approach by measuring RDMs and reconstructing the full state thereafter opens up a new avenue to efficiently realize experimental QST. However, there is still an obstacle in this approach, that it is in general computationally hard to construct the full state from its local information.

This is a typical problem empowered by machine learning. In this work, we apply the neural network model to solve this problem and demonstrate the feasibility of our method with up to 7 qubits in the simulation. It should be noticed that 7-qubit QST in experiments is already a significant challenge in many platforms – the largest QST to date is of 10 qubits in superconducting circuits, where the theoretical state is a GHZ state with rather simple mathematical form. We further demonstrate that our method works well in a 4-qubit NMR experiment, thus validating its usefulness in practice. We anticipate this method to be a powerful tool in future QST tasks of many qubits due to its accuracy and convenience.

Acknowledgments. We thank Yi Shen for helpful discussions. Tao Xin and Gui-lu Long are grateful to the following funding sources: the National Natural Science foundation of China (11175094); National Basic Research Program of China (2015CB921002). Jun Li is supported by the National Natural Science Foundation of China (Grants No. 11605005). Ningping Cao and Bei Zeng acknowledges Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Institute for Advanced Research (CIFAR) and Chinese Ministry of Education (20173080024).

* These authors contributed equally to this work.
† ludw@sustc.edu.cn
‡ lij3@sustc.edu.cn
§ zengb@uoguelph.ca

[1] G. M. D'Ariano, M. De Laurentis, M. G. Paris, A. Porzio, and S. Solimeno, Journal of Optics B: Quantum and Semiclassical Optics **4**, S127 (2002).
[2] H. Häffner, W. Hänsel, C. Roos, J. Benhelm, M. Chwalla, T. Körber, U. Rapol, M. Riebe, P. Schmidt, C. Becher, *et al.*, Nature **438**, 643 (2005).
[3] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad, J. Chiaverini, D. B. Hume, W. M. Itano, J. D. Jost, C. Langer, *et al.*, Nature **438**, 639 (2005).
[4] A. I. Lvovsky and M. G. Raymer, Reviews of Modern Physics **81**, 299 (2009).
[5] M. Baur, A. Fedorov, L. Steffen, S. Filipp, M. Da Silva, and A. Wallraff, Physical review letters **108**, 040502 (2012).
[6] A. Kalev, C. H. Baldwin, and I. H. Deutsch, arXiv preprint arXiv:1511.01433 (2015).
[7] N. Linden, S. Popescu, and W. Wootters, Physical review letters **89**, 207901 (2002).
[8] N. Linden and W. Wootters, Physical review letters **89**, 277906 (2002).
[9] L. Diósi, Physical Review A **70**, 010302 (2004).
[10] J. Chen, Z. Ji, M. B. Ruskai, B. Zeng, and D.-L. Zhou, Journal of Mathematical Physics **53**, 072203 (2012).

[11] J. Chen, Z. Ji, B. Zeng, and D. Zhou, Physical Review A **86**, 022339 (2012).

[12] J. Chen, H. Dawkins, Z. Ji, N. Johnston, D. Kribs, F. Shultz, and B. Zeng, Physical Review A **88**, 012109 (2013).

[13] B. Qi, Z. Hou, L. Li, D. Dong, G. Xiang, and G. Guo, Scientific reports **3**, 3496 (2013).

[14] B. Zeng, X. Chen, D.-L. Zhou, and X.-G. Wen, arXiv preprint arXiv:1508.02595 (2015).

[15] M. Kieferová and N. Wiebe, Physical Review A **96**, 062327 (2017).

[16] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nature Physics **14**, 447 (2018).

[17] N. Le Roux and Y. Bengio, Neural computation **20**, 1631 (2008).

[18] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.*, in *OSDI*, Vol. 16 (2016) pp. 265–283.

[19] D. P. Kingma and J. Ba, arXiv preprint arXiv:1412.6980 (2014).

[20] T. Xin, B.-X. Wang, K.-R. Li, X.-Y. Kong, S.-J. Wei, T. Wang, D. Ruan, and G.-L. Long, Chinese Physics B **27**, 020308 (2018).

[21] L. M. Vandersypen and I. L. Chuang, Reviews of modern physics **76**, 1037 (2005).

[22] J. A. Jones, V. Vedral, A. Ekert, and G. Castagnoli, Nature **403**, 869 (2000).

[23] T. Xin, S. Huang, S. Lu, K. Li, Z. Luo, Z. Yin, J. Li, D. Lu, G. Long, and B. Zeng, Science Bulletin **63**, 17 (2018).

[24] N. A. Gershenfeld and I. L. Chuang, science **275**, 350 (1997).

[25] D. G. Cory, A. F. Fahmy, and T. F. Havel, Proceedings of the National Academy of Sciences **94**, 1634 (1997).

[26] A. F. Fahmy and T. F. Havel, Quantum Computation and Quantum Information Theory: Reprint Volume with Introductory Notes for ISI TMR Network School, 12-23 July 1999, Villa Gualino, Torino, Italy , 471 (2000).

[27] E. Knill, I. Chuang, and R. Laflamme, Physical Review A **57**, 3348 (1998).

[28] N. Boulant, K. Edmonds, J. Yang, M. Pravia, and D. Cory, Physical Review A **68**, 032305 (2003).

[29] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Journal of magnetic resonance **172**, 296 (2005).

[30] C. Ryan, C. Negrevergne, M. Laforest, E. Knill, and R. Laflamme, Physical Review A **78**, 012328 (2008).

[31] D. Lu, K. Li, J. Li, H. Katiyar, A. J. Park, G. Feng, T. Xin, H. Li, G. Long, A. Brodutch, *et al.*, npj Quantum Information **3**, 45 (2017).

[32] G. M. Leskowitz and L. J. Mueller, Physical Review A **69**, 052302 (2004).

[33] J.-S. Lee, Physics Letters A **305**, 349 (2002).

[34] J. Li, S. Huang, Z. Luo, K. Li, D. Lu, and B. Zeng, Phys. Rev. A **96**, 032307 (2017).

[35] T. Xin, J. S. Pedernales, E. Solano, and G.-L. Long, Physical Review A **97**, 022322 (2018).

[36] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, Advances in Atomic, Molecular, and Optical Physics **52**, 105 (2005).