

ARTICLE OPEN

Experimental neural network enhanced quantum tomography

Adriano Macarone Palmieri¹, Egor Kovlakov², Federico Bianchi³, Dmitry Yudin¹, Stanislav Straupe^{1,2*}, Jacob D. Biamonte¹ and Sergei Kulik²

Quantum tomography is currently ubiquitous for testing any implementation of a quantum information processing device. Various sophisticated procedures for state and process reconstruction from measured data are well developed and benefit from precise knowledge of the model describing state-preparation-and-measurement (SPAM) apparatus. However, physical models suffer from intrinsic limitations as actual measurement operators and trial states cannot be known precisely. This scenario inevitably leads to SPAM errors degrading reconstruction performance. Here we develop a framework based on machine learning which generally applies to both the tomography and SPAM mitigation problem. We experimentally implement our method. We trained a supervised neural network to filter the experimental data and hence uncovered salient patterns that characterize the measurement probabilities for the original state and the ideal experimental apparatus free from SPAM errors. We compared the neural network state reconstruction protocol with a protocol treating SPAM errors by process tomography, as well as to an SPAM-agnostic protocol with idealized measurements. The average reconstruction fidelity is shown to be enhanced by 10% and 27%, respectively. The presented methods apply to the vast range of quantum experiments which rely on tomography.

npj Quantum Information (2020)6:20; <https://doi.org/10.1038/s41534-020-0248-6>

INTRODUCTION

Rapid experimental progress realizing quantum enhanced technologies places an increased demand on methods for validation and testing. As such, various approaches to augment state- and process tomography have recently been proposed. A persistent problem faced by these contemporary approaches are systematic errors in state-preparation-and-measurements (SPAM). Such notoriously challenging errors are inevitable in any experimental realization.^{1–11} Here we develop a data-driven, deep-learning based approach to augment state- and detector tomography that successfully minimized SPAM error on quantum optics experimental data.

Several prior approaches have been developed to circumvent the SPAM problem. One line of thought leads to the so-called randomized benchmarking protocols,^{2,12,13} which were designed for quality estimation of quantum gates in the quantum circuit model. The idea is to average the error over a large set of randomly chosen gates, thus effectively minimizing the average influence of SPAM. Randomized benchmarking in its initial form however, only allowed to estimate an average fidelity for the set of gates, so more elaborate and informative procedures were developed.^{3,14} Another example is gate set tomography.^{4,15,16} Therein the experimental apparatus is treated as a black box with external controls allowing for (i) state preparation, (ii) application of gates and, (iii) measurement. These unknown components (i)–(iii) are inferred from measurement statistics. Both approaches require long sequences of gates and are not suited for a simple prepare-and-measure scenario in quantum communication applications. Indeed, in such a scenario the experimenter faces careful calibration of the measurement setup, or in other words quantum detector tomography,^{5,6,17} which works reliably if known probe states can be prepared.^{18–21}

As (imperfect) quantum tomography is a data-driven technique, recent proposals suggest a natural benefit offered by machine-learning methods. Bayesian models were used to optimize the

data collection process by adaptive measurements in state reconstruction,^{7,8,22} process tomography,²³ Hamiltonian learning,²⁴ and other problems in experimental characterization of quantum devices.²⁵ Neural networks were proposed to facilitate quantum tomography in high-dimensions. In such approaches neural networks of different architectures, such as restricted Boltzmann machines,^{9,10,26} variational autoencoders,¹¹ and other architectures²⁷ are used for efficient state reconstruction; interestingly, a model for tackling a more realistic scenario of mixed quantum states has been proposed.²⁸

Our framework differs significantly and is based on supervised learning, specifically tailored to address SPAM errors. Our method hence compensates for measurement errors of the specific experimental apparatus employed, as we demonstrate on real experimental data from high-dimensional quantum states of single photons encoded in spatial modes. The success of our approach bootstraps the well-known noise filtering class of techniques in machine learning.

Performing quantum state estimation implies the reconstruction of the density matrix ρ of an unknown quantum state given the outcomes of known measurements.^{29–31} In general, a measurement is characterized by a set of positive operator valued measures (POVMs) $\{M_a\}$ with index $a \in \mathcal{A}$ the different configurations of the experimental apparatus (set \mathcal{A}). Given the configuration a , the probability of observing an outcome γ is given:

$$\mathbb{P}(\gamma|\alpha, \rho) = \text{Tr}(M_{a\gamma}\rho), \quad (1)$$

where $M_{a\gamma} \in M_a$ are POVM elements, i.e., positive operators satisfying the completeness relation $\sum_{\gamma} M_{a\gamma} = \mathbb{1}$. A statistical estimator maps the set of all observed outcomes $\mathcal{D}_N = \{Y_n\}_{n=1}^N$ onto an estimate of the unknown quantum state $\hat{\rho}$. A more general concept of quantum process tomography stands for a protocol dealing with estimation of an unknown quantum operation acting on quantum states.^{32,33} Process tomography

¹Skolkovo Institute of Science and Technology, Moscow 121205, Russia. ²Quantum Technologies Centre, Faculty of Physics, M. V. Lomonosov Moscow State University, Moscow 119991, Russia. ³Department of Computer Science, University of Milan-Bicocca, Milan 20133, Italy. *email: straups@quantum.msu.ru

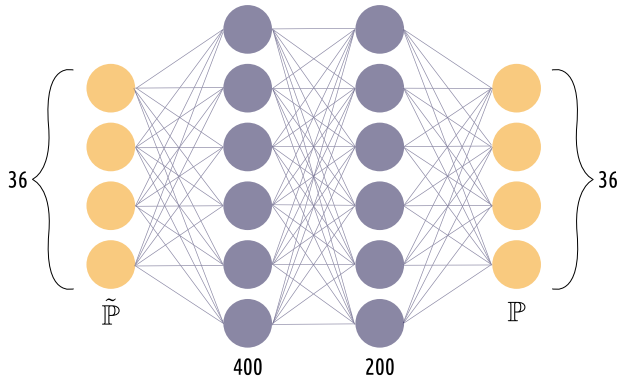


Fig. 1 The DNN architecture employed in our experiments. Input and output layers constitute of 36 neurons each and two hidden layers of 400 and 200 neurons, respectively. The DNN modifies its internal parameters to find a function $\mathcal{F} : \tilde{\mathbb{P}}(\gamma|\alpha, \tilde{\rho}) \rightarrow \mathbb{P}(\gamma|\alpha, \rho)$ which translates between the experimentally estimated probabilities $\tilde{\mathbb{P}}(\gamma|\alpha, \tilde{\rho})$, subjected to SPAM errors, at the input and ideal $\mathbb{P}(\gamma|\alpha, \rho)$ at the output. To achieve this goal the network is forced to reduce the Kullback–Leibler divergence amongst pairs of distributions. An early stopper is applied in order to avoid overfitting during the training phase.

uses measurements on a set of known test states $\{\rho_a\}$ to recover the description of an unknown operation. (See Supplementary Material for the thorough discussion of quantum process tomography and its application for calibration of the measurement setup).

The reconstruction procedure requires knowledge of the measurement operators $\{M_{a\gamma}\}$, as well as the test states $\{\rho_a\}$ in the case of process tomography. However, both tend to deviate from the experimenter’s expectations due to stochastic noise and systematic errors. While stochastic noise may to some extent be circumvented by increasing the sample size, systematic errors are notoriously hard to correct. The only known way to make tomography reliable is to explicitly incorporate these errors in (Eq. 1). Thus, trial states and measurements should be considered as acted upon by some SPAM processes: $\tilde{\rho}_a = \mathcal{R}(\rho_a)$ and $\tilde{M}_{a\gamma} = \mathcal{M}(M_{a\gamma})$, and the models for these processes should be learned independently from a calibration procedure. Such calibration is essentially tomography on its right. For example, the reconstruction of measurement operators is known as detector tomography^{5,6,17,34,35} and requires ideal preparation of calibration states. The most straightforward approach is calibration of the measurement setup with some close-to-ideal and easy to prepare test states, or calibration of the preparation setup with known and close-to-ideal measurements. In this case, one may then infer the processes \mathcal{R} and/or \mathcal{M} explicitly—for example—in the form of the corresponding operator elements, and incorporate this knowledge in the reconstruction procedure. Ideally, this procedure should produce an estimator free from bias caused by systematic SPAM errors. (See Supplementary Material for the detailed description of this procedure applied to our experiment).

RESULTS

Given the estimates of raw probabilities inferred from the experimental dataset $\tilde{\mathbb{P}}(\gamma|\alpha, \tilde{\rho}) = \text{Tr}(\tilde{M}_{a\gamma}\tilde{\rho})$, one wants to establish a one-to-one correspondence $\tilde{\mathbb{P}}(\gamma|\alpha, \tilde{\rho}) \leftrightarrow \mathbb{P}(\gamma|\alpha, \rho)$ with the ideal probabilities for the measurement setup free from SPAM errors. We use a deep neural network (DNN) to approximate the map from $\tilde{\mathbb{P}}$ to \mathbb{P} .

To train and test the DNN we prepare a dataset of N Haar-random pure states $\mathcal{D}_N = \{|\psi_i\rangle\}_{i=1}^N$. For a d -dimensional Hilbert

space, reconstruction of a Hermitian density matrix with unit trace requires at least d^2 different measurements. The network is trained on the dataset, consisting of $d^2 \times N$ frequencies experimentally obtained by performing the same d^2 measurements $\{\tilde{M}_\gamma\}_{\gamma=1}^{d^2}$ for all N states (in our experiments $d = 6$). These frequencies are fed to the input layer of the feed-forward network consisting of $d^2 = 36$ neurons. Training is performed by minimization of the loss function, defined as the sum of Kullback–Leibler divergences between the distributions of predicted probabilities $\{p_\gamma^i\}_{\gamma=1}^{d^2}$ at the output layer of the network and the ideally expected probabilities $\{\mathbb{P}_\gamma^i\}_{\gamma=1}^{d^2}$, which are calculated for the test states as $\mathbb{P}_\gamma^i = \text{Tr}(M_\gamma \rho_i)$ assuming errorless projectors M_γ :

$$L = \sum_{i=1}^N D_{\text{KL}}(\{\mathbb{P}^i\} || \{p^i\}) = \sum_{i=1}^N \sum_{\gamma=1}^{d^2} \mathbb{P}_\gamma^i \log \left(\frac{\mathbb{P}_\gamma^i}{p_\gamma^i} \right). \quad (2)$$

We tested different neural architectures with different configuration parameters; currently, there are few guidelines that explain how to find a suitable neural network to solve a specific problem. In general deeper architectures are more difficult to train due to the increasing number of parameters; the best architecture we have found uses two hidden layers as shown in Fig. 1. The first hidden layer is chosen to consist of 400 neurons, whilst the second contains 200. (See Supplementary Material for DNN architecture and the details of training process). To prevent overfitting we applied dropout between the two hidden layers with drop probability equal to 0.2, i.e., at each iteration we randomly drop 20% neurons of the first hidden layer in such a way that the network becomes more robust to variations. We use a rectified linear unit as an activation function after both hidden layers, while in the final output d^2 -dimensional layer we use a softmax function to transform the predicted values to valid normalized probability distributions. Following the standard paradigm of statistical learning, we divided our dataset of overall $N = 10,500$ states (represented by their density matrix elements) into 7000 states for training, 1500 states for validation, and 2000 for testing. The validation set is an independent set and is used to stop the network training as soon as the error evaluated for this set stops decreasing.

We fix the set of tomographically complete measurements $\{\mathbb{M}_a\} = \mathbb{M}$ to estimate all matrix elements of ρ using (1) and an appropriate estimator. We will assume that our POVM \mathbb{M} consists of d^2 one-dimensional projectors $M_\gamma = |\varphi_\gamma\rangle\langle\varphi_\gamma|$. These projectors are transformed by systematic SPAM errors into some positive operators \tilde{M}_γ . Experimental data consist of frequencies $f_\gamma = n_\gamma/n$, where n_γ is the number of times an outcome γ was observed in a series of n measurements with identically prepared state ρ . For the time being, we assume, that all the SPAM errors can be attributed to the measurement part of the setup, and the state preparation may be performed reliably. This is indeed the case in our experimental implementation (see Supplementary Material).

We reconstruct high-dimensional quantum states encoded in the spatial degrees of freedom of photons. The most prominent example of such encoding uses photonic states with orbital angular momentum (OAM)³⁶ as relevant to numerous experiments in quantum optics and quantum information. However, OAM is only one of two quantum numbers, associated with orthogonal optical modes, and radial degree of freedom of Laguerre–Gaussian beams^{37,38} as well as full set of Hermite–Gaussian (HG) modes³⁹ offer viable alternatives for increasing the accessible Hilbert space dimensionality. One of the troubles with using the full set of orthogonal modes for encoding is the poor quality of projective measurements. Existing methods to remedy the situation⁴⁰ trade reconstruction quality for efficiency, significantly reducing the latter. Complex high-dimensional projectors are especially vulnerable to measurement

errors and fidelities of state reconstruction are typically at most ~ 0.9 in high-dimensional tomographic experiments.⁴¹ That provides a challenging experimental scenario for our machine-learning-enhanced methods.

Our experiment is schematically illustrated in Fig. 2. We use phase holograms displayed on the spatial light modulator as spatial mode transformers. At the preparation stage an initially Gaussian beam is modulated both in phase and in amplitude to an arbitrary superposition of HG modes, which are chosen as the basis in the Hilbert space. At the detection phase the beam passes through a phase-only mode-transforming hologram and is focused to a single mode fiber, filtering out a single Gaussian mode. This sequence corresponds to a projective measurement in mode space, where the projector \hat{M}_ν is determined by the phase hologram. (See Supplementary Material for the details of the experimental setup and state preparation and detection methods). In dimension $d=6$, we are able to prepare an arbitrary superposition expressed in the basis of HG modes as $|\psi\rangle = \sum_{i,j=0}^2 c_{ij} |\text{HG}_{ij}\rangle$. In the measurement phase we used a symmetric informationally complete POVM, which is close to

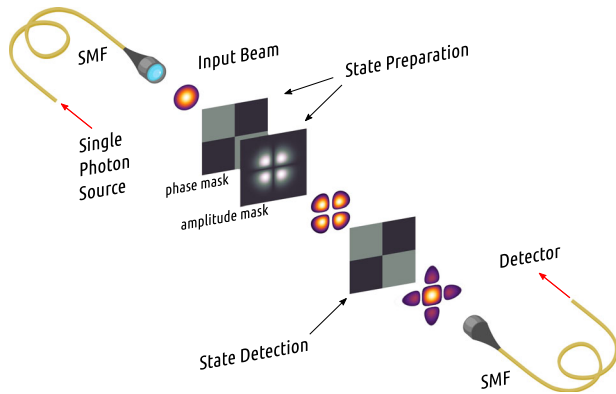


Fig. 2 Experimental setup for preparation and measurement of spatial qudit states. In the generation part, single photons from a heralded source are beam-shaped by a single mode fiber (SMF) and then transformed by a hologram displayed on a spatial light modulator. Analogously, the detection part consists of a hologram corresponding to the chosen detection mode, followed by a single mode fiber and a single photon counter. The hologram in the generation part produces high-quality HG modes with the use of amplitude modulation, while a phase-only hologram at the detection part sacrifices projection quality for efficiency.

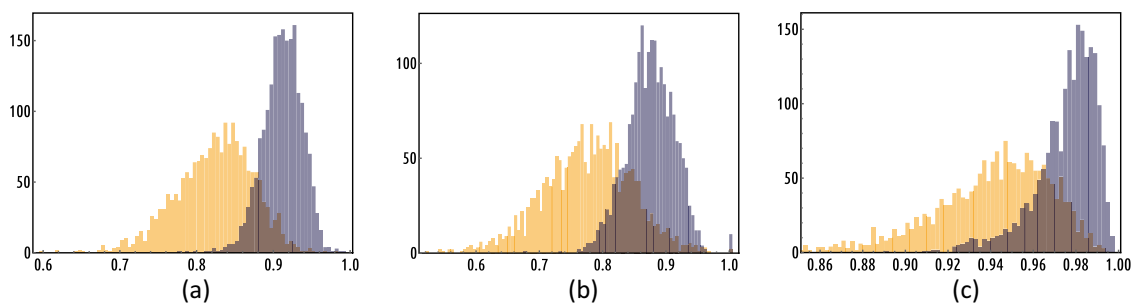


Fig. 3 Results of experimental state reconstruction with phase-only holograms. **a** Fidelity of the experimentally reconstructed states with ideal $F = \langle \psi^j | \hat{\rho}_{(raw/nn)}^j | \psi^j \rangle$ for 2000 test states reconstructed from raw data (orange bars) and reconstructed after neural network processing of the data (blue bars). **b** A similar diagram for purity of the reconstructed states, $\pi = \text{Tr} \hat{\rho}^2$. **c** Fidelity histogram for the case, when the state is reconstructed to be pure. The results of the filtering process are clearly witnessed by the modification of data histogram shapes. Besides the shifting towards higher values that shows average gain over our experimental data, the reduction of FWHM indicates filtering task by the neural network.

optimal for state reconstruction and may be relatively easily realized for spatial modes.⁴¹

We performed state reconstruction using maximum likelihood estimation⁴² for both raw experimental data and DNN-processed data. (See also Supplementary Material for extra information on spatial probability distribution of reconstructed states). In the former case, the log-likelihood function to be maximized with respect to ρ has been chosen as $\mathcal{L}(f_\nu^i | \rho) \propto \sum_{\nu=1}^{36} f_\nu^i \log [\text{Tr}(M_\nu \rho)]$, with frequencies $f_\nu = n_\nu/n$ and i numbering the test set states. Whereas in the latter case, these frequencies have been replaced with predicted probabilities p_ν . The results for $\hat{\rho}_{(raw)}^j = \text{argmax} \mathcal{L}(f_\nu^j | \rho)$ and $\hat{\rho}_{(nn)}^j = \text{argmax} \mathcal{L}(p_\nu^j | \rho)$ with the prepared states $|\psi^j\rangle$ are shown in Fig. 3. Interestingly, the average reconstruction fidelity increases from $F_{(raw)} = (0.82 \pm 0.05)$ to $F_{(nn)} = (0.91 \pm 0.03)$ and this increase is uniform over the entire test set. Similar behavior is observed for the purity—since we did not force the state to be pure in the reconstruction, the average purity of the estimate is less than unity: $\pi_{(raw)} = (0.78 \pm 0.07)$, whereas $\pi_{(nn)} = (0.88 \pm 0.04)$. If the restriction to pure states is explicitly imposed in the reconstruction procedure, the fidelity increase is even more significant, as shown in Fig. 3(c). In this case the initially relatively high fidelity of $F_{(raw)} = (0.94 \pm 0.03)$ increases to $F_{(nn)} = (0.98 \pm 0.02)$ —a very high value, given the states dimensionality.

DISCUSSION

Our results were obtained with analytical correction for some known SPAM errors already performed. In particular, we have explicitly taken into account the Gouy phase-shifts acquired by the modes of different order during propagation (see Supplementary Material). This correction is however unnecessary for neural network post processing. The DNN has been trained without any need of data preprocessing over the experimental dataset, as to say without introducing any phase correction in our initial data, wherein considering the effect of a channel process \mathcal{E} . However, we have achieved average estimation fidelities of $F_{(nn)} = (0.81 \pm 0.19)$ as compared with $F_{(raw)} = (0.54 \pm 0.12)$ for this completely agnostic scenario, showing a dramatic improvement by straightforward application of a learning approach.

In our experiment the prepared input states were close to pure and we did not have a controllable way to systematically change their purity. However, it is exactly the case of pure input states, which is most seriously affected by SPAM errors, since in this case the outcome probabilities are the most sensitive to perturbations in the measurement operators. So the presented experimental results illustrate the worst-case performance of our method. We have

examined the case of mixed states by numerical simulations (details and results of the simulation are provided in Supplementary Material), and conclude that DNN post processing enhances the reconstruction quality for any purity of the input state.

DNN processing may be straightforwardly generalized to process tomography, since the latter may always be formulated as a state tomography in higher dimensional space due to Choi–Jamiołkowski isomorphism. It may present a valuable tool in the case where randomized benchmarking and similar protocols cannot be applied, such as in channel testing for quantum communication.

Although the number of neurons in the hidden layers is quite large in the current realization, the training procedure is still fast and we do not believe it will be an issue for reasonable applications. In the end, our method is specifically designed for full state tomography, which itself is limited to rather small Hilbert space dimensionalities due to fast growth of the number of measurements required. Optimization of DNN architecture and applications to protocols providing partial information about the state are interesting directions for future work.

To conclude, our results unambiguously demonstrate that a use of neural-network-architecture on experimental data can provide a reliable tool for quantum state-and-detector tomography.

DATA AVAILABILITY

The data that support this study are available at <https://github.com/Quantum-Machine-Learning-Initiative/dnnquantumtomography>.

CODE AVAILABILITY

The code that supports this study is available at <https://github.com/Quantum-Machine-Learning-Initiative/dnnquantumtomography>.

Received: 27 June 2019; Accepted: 13 January 2020;

Published online: 07 February 2020

REFERENCES

- Rosset, D., Ferretti-Schöbitz, R., Bancal, J.-D., Gisin, N. & Liang, Y.-C. Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses. *Phys. Rev. A* **86**, 062325 (2012).
- Knill, E. et al. Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
- Merkel, S. T. et al. Self-consistent quantum process tomography. *Phys. Rev. A* **87**, 062119 (2013).
- Blume-Kohout, R. et al. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nat. Commun.* **8**, 14485 (2017).
- Lundeen, J. et al. Tomography of quantum detectors. *Nat. Phys.* **5**, 27 (2009).
- Brida, G. et al. Ancilla-assisted calibration of a measuring apparatus. *Phys. Rev. Lett.* **108**, 253601 (2012).
- Huszár, F. & Houlisby, N. M. T. Adaptive Bayesian quantum tomography. *Phys. Rev. A* **85**, 052120 (2012).
- Kravtsov, K. S. et al. Experimental adaptive Bayesian tomography. *Phys. Rev. A* **87**, 062122 (2013).
- Torlai, G. et al. Neural-network quantum state tomography. *Nat. Phys.* **14**, 447 (2018).
- Carleo, G., Nomura, Y. & Imada, M. Constructing exact representations of quantum many body systems with deep neural network. *Nat. Commun.* **9**, 5322 (2018).
- Rocchetto, A., Grant, E., Strelchuk, S., Carleo, G. & Severini, S. Learning hard quantum distributions with variational autoencoders. *npj Quantum Inf.* **4**, 28 (2018).
- Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.* **106**, 180504 (2011).
- Wallman, J. J. & Flammia, S. T. Randomized benchmarking with confidence. *New J. Phys.* **16**, 103032 (2014).
- Roth, I. et al. Recovering quantum gates from few average gate fidelities. *Phys. Rev. Lett.* **121**, 170502 (2018).
- Blume-Kohout, R. et al. Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit (2013). <https://arxiv.org/pdf/1310.4492.pdf>.
- Dehollain, J. P. et al. Optimization of a solid-state electron spin qubit using gate set tomography. *New J. Phys.* **18**, 103018 (2016).
- Bobrov, I. B., Kovlakov, E. V., Markov, A. A., Straupe, S. S. & Kulik, S. P. Tomography of spatial mode detectors. *Optics Express* **23**, 649–654 (2015).
- Mogilevtsev, D., Rehacek, J. & Hradil, Z. Self-calibration for self-consistent tomography. *New J. Phys.* **14**, 095001 (2012).
- Brańczyk, A. M. et al. Self-calibrating quantum state tomography. *New J. Phys.* **14**, 085003 (2012).
- Straupe, S. S. et al. Self-calibrating tomography for angular schmidt modes in spontaneous parametric down-conversion. *Phys. Rev. A* **87**, 042109 (2013).
- Jackson, C. & van Enk, S. J. Detecting correlated errors in state-preparation-and-measurement tomography. *Phys. Rev. A* **92**, 042312 (2015).
- Granade, C., Ferrie, C. & Flammia, S. T. Practical adaptive quantum tomography. *New J. Phys.* **19**, 113017 (2017).
- Pogorelov, I. A. et al. Experimental adaptive process tomography. *Phys. Rev. A* **95**, 012302 (2017).
- Granade, C. E., Ferrie, C., Wiebe, N. & Cory, D. G. Robust online Hamiltonian learning. *New J. Phys.* **14**, 103013 (2012).
- Lennon, D. et al. Efficiently measuring a quantum device using machine learning (2018). <https://arxiv.org/abs/1810.10042>.
- Carrasquilla, J., Torlai, G., Melko, R. G. & Aolita, L. Reconstructing quantum states with generative models. *Nat. Mach. Intell.* **1**, 155 (2019).
- Xin, T. et al. Local-measurement-based quantum state tomography via neural networks (2018). <https://arxiv.org/pdf/1807.07445.pdf>.
- Torlai, G. & Melko, R. G. Latent space purification via neural density operator. *Phys. Rev. Lett.* **120**, 240503 (2018).
- Banaszek, K., D'Ariano, G. M., Paris, M. G. A. & Sacchi, M. F. Maximum-likelihood estimation of the density matrix. *Phys. Rev. A* **61**, 010304 (1999).
- James, D. F. V., Kwiat, P. G., Munro, W. J. & White, A. G. Measurement of qubits. *Phys. Rev. A* **64**, 052312 (2001).
- Paris, M. & Řeháček, J. (eds) *Quantum State Estimation*, vol. 649 of *Lecture Notes in Physics* (Springer-Verlag, 2004). <http://www.springer.com/gp/book/9783540223290>.
- Chuang, I. L. & Nielsen, M. A. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Optics* **44**, 2455–2467 (1997).
- Poyatos, J. F., Cirac, J. I. & Zoller, P. Complete characterization of a quantum process: the two-bit quantum gate. *Phys. Rev. Lett.* **78**, 390–393 (1997).
- Fiurášek, J. Maximum-likelihood estimation of quantum measurement. *Phys. Rev. A* **64**, 024102 (2001).
- D'Ariano, G. M., Maccone, L. & Presti, P. L. Quantum calibration of measurement instrumentation. *Phys. Rev. Lett.* **93**, 250407 (2004).
- Molina-Terriza, G., Torres, J. P. & Torner, L. Twisted photons. *Nat. Phys.* **3**, 305 (2007).
- Salakhutdinov, V. D., Eliel, E. R. & Löffler, W. Full-field quantum correlations of spatially entangled photons. *Phys. Rev. Lett.* **108**, 173604 (2012).
- Krenn, M. et al. Generation and confirmation of a (100 × 100)-dimensional entangled quantum system. *Proc. Natl Acad. Sci.* **111**, 6243–6247 (2014).
- Kovlakov, E. V., Bobrov, I. B., Straupe, S. S. & Kulik, S. P. Spatial bell-state generation without transverse mode subspace postselection. *Phys. Rev. Lett.* **118**, 030503 (2017).
- Bouchard, F. et al. Measuring azimuthal and radial modes of photons. *Optics Express* **26**, 31925–31941 (2018).
- Bent, N. et al. Experimental realization of quantum tomography of photonic qudits via symmetric informationally complete positive operator-valued measures. *Phys. Rev. X* **5**, 041006 (2015).
- Hradil, Z. Quantum-state estimation. *Phys. Rev. A* **55**, R1561–R1564 (1997).

ACKNOWLEDGEMENTS

The authors acknowledge financial support under the Russian National Technological Initiative via MSU Quantum Technologies Centre and RFBR grant #19-52-80034, and thank Timur Tlyachev and Dmitry Dylov for helpful suggestions on an early version of this study. E.K. acknowledges support from the BASIS Foundation.

AUTHOR CONTRIBUTIONS

A.M.P., F.B. and D.Y. developed the neural network architecture and the training procedure. E.K. and S.S. devised the experiment, E.K. gathered the data. J.B. and S.K. supervised the project. All authors discussed the results and contributed to writing the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at <https://doi.org/10.1038/s41534-020-0248-6>.

Correspondence and requests for materials should be addressed to S.S.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020