Calibration of Quantum Sensors by Neural Networks

Valeria Cimini,[1] Ilaria Gianani,[1,2,*] Nicolò Spagnolo,[2] Fabio Leccese,[1] Fabio Sciarrino,[2,3] and Marco Barbieri[1,4]
[1]Dipartimento di Scienze, Università degli Studi Roma Tre, Via della Vasca Navale 84, 00146, Rome, Italy
[2]Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro, 5, 00185, Rome, Italy
[3]Consiglio Nazionale delle Ricerche, Istituto dei sistemi Complessi (CNR-ISC), Via dei Taurini 19, 00185, Rome, Italy
[4]Istituto Nazionale di Ottica-CNR, Largo Enrico Fermi 6, 50125, Florence, Italy

(Received 30 April 2019; published 4 December 2019)

Introducing quantum sensors as a solution to real world problems demands reliability and controllability outside of laboratory conditions. Producers and operators ought to be assumed to have limited resources readily available for calibration, and yet, they should be able to trust the devices. Neural networks are almost ubiquitous for similar tasks for classical sensors: here we show the applications of this technique to calibrating a quantum photonic sensor. This is based on a set of training data, collected only relying on the available probe states, hence reducing overhead. We found that covering finely the parameter space is key to achieving uncertainties close to their ultimate level. This technique has the potential to become the standard approach to calibrate quantum sensors.

DOI: 10.1103/PhysRevLett.123.230502

Quantum technologies are experiencing a worldwide effort to foster their applications beyond what is achieved in a laboratory. In particular, for quantum sensing, quantum resources are promising to reach an accuracy beyond what is permitted from classical counterparts [1]. This advantage, however, is conditioned on a robust operation in the presence of noise as well as imperfections of the measuring instruments [2,3]. Many methods have been proposed to this end, including error correction [4–7], monitoring of the environment [8], and imperfection-tolerant probe state design [9,10].

Regardless of the method adopted, it is vital to devise analysis methods that grant optimal use of the collected data for estimation, i.e., to obtain the so-called optimal estimator. In simple instances, the maximum likelihood approach or methods based on Bayes' theorem are known to provide such an estimator [11,12]. On the other hand, these are generally computationally intensive, often require a more involved characterization of the system [13–16], and thus pose difficulties in scaling to configurations with increased complexity. Further, characterization is generally based on preparing quantum states with different requirements than those actually used in the estimation routine [17–19]. From the perspective of compact architectures, the resulting requirement of flexibility may come at odds with that of reliability and reduced costs [20]. A method that is self-consistent, resource economic, and versatile is desirable.

Nowadays, the incredible amount of data collected in diverse problems, such as financial market analysis or bioinformatics, requires protocols which are capable of efficiently self-adjusting their operation. The size and complexity of these problems has imposed machine learning (ML) algorithms as the mainstream solution in these situations [21,22]. ML has been recently proposed and applied as a tool for characterization and optimization of quantum systems, as well as for handling quantum physics problems [23]. Notable examples include its adoption for the learnability of quantum measurements, states, and processes [24–31], validation of multiparticle interference [32,33], quantum state engineering [34–36], and as a tool for quantum experiment design and control [37–43]. In the context of quantum metrology, ML has found an application in quantum phase estimation protocols to efficiently extract the information encoded in the probe [44–46]. More specifically, ML represents a powerful toolbox to optimize, via adaptive protocols [47–49], the performance of a sensor operating with a small collection of repetitions [12]. These considerations suggest the viability of this approach for the calibration of quantum sensing apparatuses. Specifically, neural networks can extract an output value of the parameter of interest, following their training on a set of inputs associated with a calibrated set of parameters; this is an efficient algorithm that can be run on ordinary machines [50,51]. No explicit modeling of the imperfections is thus needed, as that information will be taken into account, although in implicit form, in the training itself. This results in a practical relevance of the method, in particular, for sensors whose operation depends on a large number of parameters. On the other hand, the training data can only be collected in finite time, hence with finite statistics: it is important to understand how the associated uncertainty influences the quality of the final estimation and its capability of showing quantum enhancement. In this Letter, we discuss the characterization of a quantum phase sensor based on $N00N$ states by means of the neural

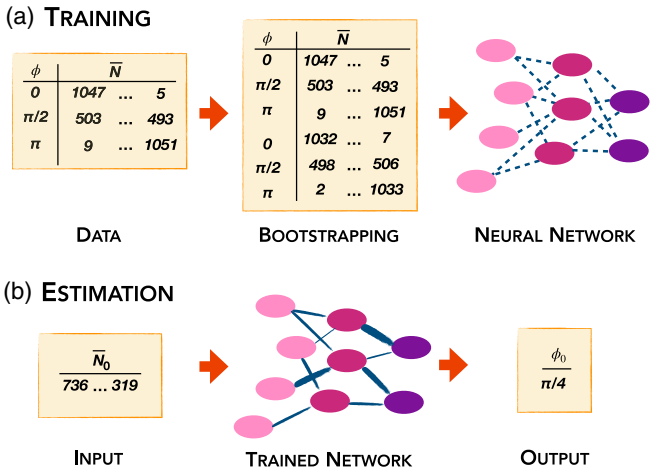

FIG. 1. Schematics of the use of neural networks for parameter estimation. The first step (a) consists of training the network by inputting a set of training data using bootstrapping to account for uncertainties. Upon completion, the actual estimation (b) uses the trained network to extract an estimate of the parameter.

network. We find that the simplicity of the characterization does not impact heavily the metrological capabilities of the device. The algorithm has no strict requirements in its settings to provide near-optimal performance of the estimation, provided that the set of the parameter is explored with adequate resolution. Thanks to its compact implementation and its scalability, this method can find wide applications in future quantum technologies.

The concept behind our investigation is depicted in Fig. 1. The training step [Fig. 1(a)] consists of collecting a set of data $\vec{N}(\phi)$ corresponding to different values of the parameter of interest ϕ; in general $\vec{N}$ will be in the form of a vector, since it contains multiple measurements, as needed to obtain a correct normalization, to remove ambiguities or to account for multiple parameters. This is used as an input to a network, consisting of a set of neurons connected among them, possibly forming subsequent layers. The training procedure establishes the weights associated with the connections between each pair of neurons. Unavoidably, an uncertainty will be associated with these measured data, and the network needs to be trained to account for this variability. If the noise statistics on each measurement in $\vec{N}$ is known, a bootstrapping method can be employed. This consists of generating multiple fictitious runs n_b of the calibration measurements performed and can be achieved by means of a Monte Carlo routine [52]. Bootstrapping does not provide additional information on the mean value of the measured counts, but it is required to supply the training information on the variance of the counts. For a fixed network size, the quality of the training will be influenced by the resolution at which ϕ is sampled, as well as the number of repetitions M_{train} used for each value of ϕ. Once the training is complete, the device can be used for parameter estimation: the network is now operated

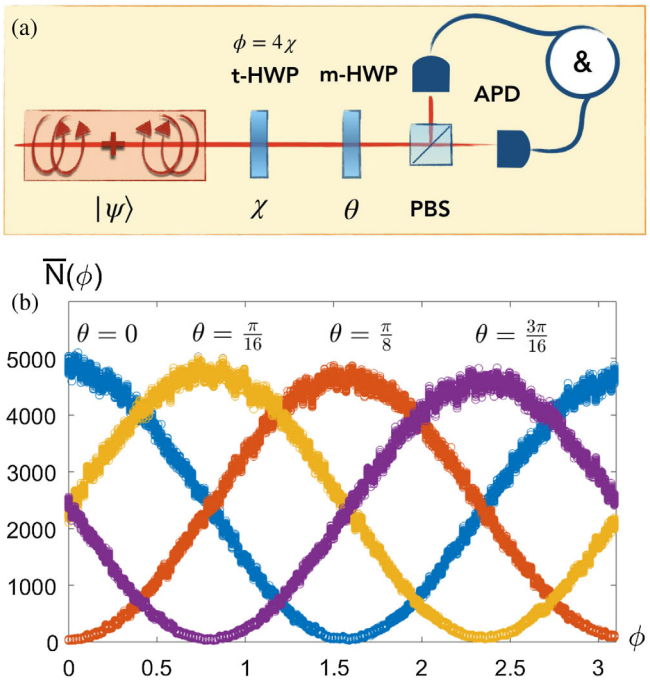

FIG. 2. (a) Experimental setup. The $N00N$ state probe accumulates a phase $\phi = 4\chi$ by means of a test half wave plate (t-HWP). The analysis is carried out by a second measurement plate (m-HWP), a polarizing beam splitter, and coincidence detection of two avalanche photodiodes (APDs). (b) Part of the experimental training set. The registered count rates and those derived from bootstrapping are reported as a function of the calibrated phases ϕ. The labels indicate the angular setting θ of the measurement HWP, corresponding to four different polarization projections.

to accept the collected data $\vec{N}_0$ as an input and to provide an estimation ϕ_0 as the output [Fig. 1(b)]. By using the same bootstrapping method above on $\vec{N}_0$, the uncertainty $\Delta\phi_0$ can also be evaluated.

We test this method in a quantum phase estimation experiment. A two-photon $N00N$ state on the right- and left-circular polarization of a single spatial mode approximating $|\psi\rangle = 2^{-1}[(a_R^\dagger)^2 + (a_L^\dagger)^2]|0\rangle$ is used for this task (see Fig. 2). This is achieved by overlapping on the same spatial mode two otherwise indistinguishable orthogonally polarized photons by means of a polarizing beam splitter (PBS) [16,53]: the photon pairs are generated via a type-I spontaneous parametric down-conversion source using a 3 mm β-barium borate crystal pumped with a 405 nm cw laser; their polarization is then set orthogonal by means of half wave plates (HWPs) and they are sent on a PBS obtaining the circular polarization $N00N$ state. The measurement is then carried out collecting a vector of coincidence counts associated with photon pairs after they have accumulated a relative phase ϕ, resulting from the rotation of their polarization state. For each phase, we look at the coincidence counts relative to four different polarization projections [16].

Even in this simple instance, many imperfections affect the actual experimental apparatus, including those linked to the phase preparation, as well as imperfect splittings on the PBSs and polarization-dependent efficiencies of the detection channels. If a maximum likelihood or Bayesian technique were to be used for phase estimation, expression for the probabilities, encompassing all these parameters, would be needed. In more complex scenarios this becomes even more demanding. We test how the effectiveness of the implicit treatment of imperfections is made possible by neural networks.

In order to perform the training, we collect data for 180 different phase values from 0° to 180° in steps of 1°, simply obtained by rotating a HWP from 0° to 45° in steps of 0.25°—the relation between the phase ϕ and the HWP angle setting χ is $\phi = 4\chi$. The initial training data thus consist of 180 vectors, each containing four counting rates. Each projection accumulates data for 1 s, which, at the observed count rate, corresponds to around $M_{\text{train}} = 10\,000$ total events for each phase, divided among the four projections (Fig. 2).

The coincidence counts of the training set are normalized to the total number of collected events M_{train} to obtain the associated frequencies that are used for the supervised training of the network. This yields a calibration that is independent of the number of events collected during the training, which would otherwise result in an unnecessary limitation for the device. Since the coincidence counts follow a Poissonian distribution, we can test the accuracy by using the bootstrap procedure described above to feed our network with different repetitions associated with each phase. The network is structured as a feed forward network with sigmoid hidden neurons [54]. This includes an input layer, an output layer, and a hidden layer with n_n neurons. The input data are randomly divided in a training set (70%), a validation set (15%), and a test set (15%) (see Supplemental Material [55]). The network is trained with the Levenberg-Marquardt backpropagation algorithm [61,62] designed to optimize for every phase the precision on the validation set using the mean square error metric. The training stops when the validation error stops decreasing.

We have run tests on simulated data to establish the optimal parameters for the training for fixed sampling step and total count rate: for this purpose, we input into the trained network the counting rates corresponding to 30 values of the phase and consider as the error ϵ the mean standard deviation calculated on 100 repetitions and averaged on all phases. Figure 3 shows how this error depends on the number of neurons n_n in the network and on the number of Monte Carlo runs per phase value n_b. The data do not show sharp optimal working points: we have adopted one hidden layer in the network with 30 neurons, trained with 50 Monte Carlo repetitions.

With the network architecture fixed according to the parameters discussed above, i.e., a 30 neurons single hidden layer, we have proceeded to the final training of the network. The testing performed so far only addresses the self-consistency of the operation of the network. However, the assessment of the calibration for the sensor needs further investigation. In order to perform a phase estimation detached from that performed in the training, we have then collected five further sets for the phases 20.8°, 45°, 90°, 140°, and 168.8° at different accumulation times (0.5, 1, and 4 s) at the same generation rate as for the training and 0.5 s at 30% of the initial generation rate. These result in a different number of collected events M_{exp} for each set. We can thus observe how the uncertainty scales with M_{exp} for fixed training parameters. The uncertainties in the phase $\Delta\phi$ are evaluated with the standard deviation over the 50 bootstraps on the experimental data collected for the estimation. In Fig. 4, we show such uncertainties compared to the associated Cramér-Rao bound (CRB) [1,63]. The uncertainties remain close to the ultimate limit, which also takes into account the reduced contrast of the coincidence oscillations. Remarkably, we

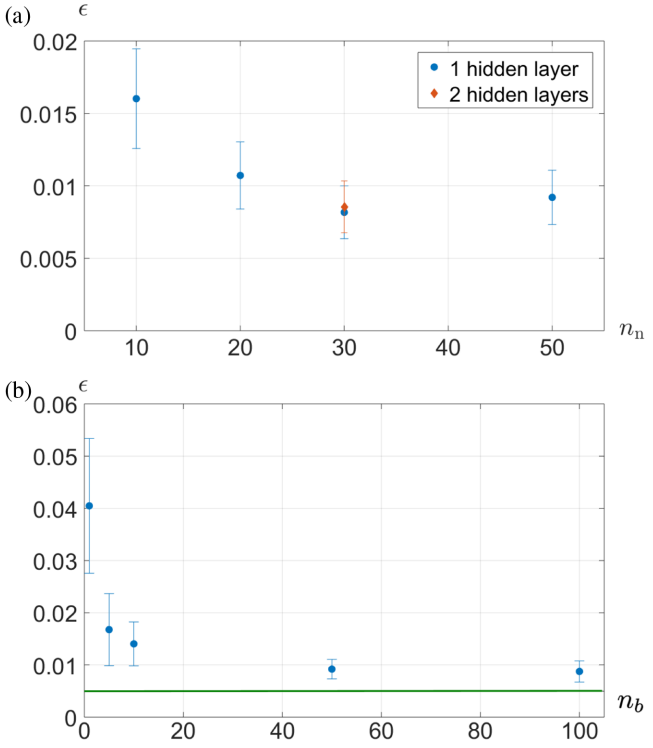

FIG. 3. Exploration of the training parameters, based on $M_{\text{train}} = 10\,000$ simulated events. (a) Estimation error ϵ as a function of the number of neurons n_n in the hidden layers of the network at $n_b = 50$. For $n_n = 30$, we have considered a single layer or two layers with 20 and 10 neurons. (b) Estimation error ϵ for a different number of Monte Carlo repetitions n_b obtained from bootstrapping at $n_n = 30$. The solid line indicates the Cramér-Rao bound. (a),(b) The error bars are calculated from 35 different trainings; this takes into account the variations due to the random initialization of the neural network algorithm at the beginning of the procedure.

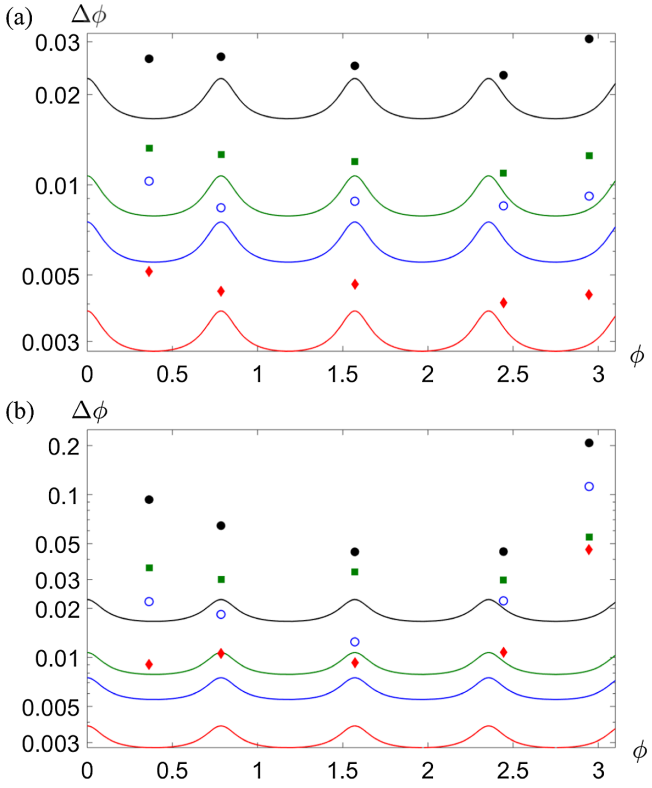

FIG. 4. Uncertainties for parameter estimation. (a) Training set with a sampling step of 1°. (b) Training set with a sampling step of 2°; notice the logarithmic scale. (a),(b) The points refer to total numbers of events $M_{\mathrm{exp}} \simeq 1000$ (black solid dots), $M_{\mathrm{exp}} \simeq 5000$ (green squares), $M_{\mathrm{exp}} \simeq 10\,000$ (blue empty dots), and $M_{\mathrm{exp}} \simeq 40\,000$ (red diamonds); the curves are the Cramér-Rao bounds for the same number of events.

observe that such performances are achieved also for a number of events exceeding those used for the training, making the calibration with the neural network robust to noise. We have repeated the training stage decreasing the step size of the phase to 2°. The distance to the CRB is more pronounced: the network has received insufficient training to produce an output with an accuracy close to optimal. We noticed that, also in this case, a reduction of the uncertainty is observed when increasing the number of events M_{exp}. We put these observations in quantitative terms by inspecting $F_{M_{\mathrm{exp}}} = \Delta^2\phi/\sigma^2$, i.e., the ratio between the measured variance $\Delta^2\phi$ and the one σ^2 at the CRB at a given M_{exp}. We obtain the following values: $F_{1000} = 1.25$, $F_{5000} = 1.39$, $F_{10\,000} = 1.48$, and $F_{40\,000} = 1.53$. We note that the F value increases with the repetition number, which implies that variance is diminishing with M_{exp}, however, not as fast as the CRB is. This is due to the lack of resolution in the training, which is not a fundamental limit, and in our case it is solely constrained by the accuracy of the actuator of m-HWP.

While overall successful, this characterization underlines some issues that need to be addressed. The first one is a boundary effect that is observed close to $\phi = 0°$ and $\phi = 180°$. This border effect prevents us from obtaining a reliable estimate close to this value and can be removed by increasing the region explored for the training, which should span a wider phase interval than the one potentially covered in the measurement. Excess uncertainty is associated with phases close to $\phi = 90°$, a problem we can attribute to the particular symmetries of the count signals (see Fig. 2). A second issue regards a periodicity ambiguity arising due to the form of the system probabilities. We note that using the four count signal scheme allows us to suppress the periodicity ambiguity on the phase from 0 to $\pi/2$ to the range $0 - \pi$, but still leaves an ambiguity in the range $0 - 2\pi$. This latter ambiguity can be approached by integrating adaptive techniques, able to discriminate an unknown phase in the full interval [12], within the phase estimation protocol.

In order to verify that the versatility of the method is not associated with reduced metrological capabilities, we have performed numerical simulations. In particular, we considered an alternative scenario consisting of maximum likelihood estimation associated with a calibration procedure that performs a reconstruction of the system outcome probabilities. Contrary to the adoption of neural networks, such method requires information on a theoretical model describing the system parameters and imperfection. In general, this requirement becomes impractical for large-size sensors. Nevertheless, numerical simulations and additional analysis on experimental data show that, for a fixed number of resources employed in the calibration, the neural networks method is capable of matching or even outperforming maximum likelihood estimation (see Supplemental Material [55]).

In conclusion, we have applied a neural network algorithm to the calibration of a quantum phase sensor. This method compares favorably to previous investigations that require a complete reconstruction of the functioning of the device [13,14] or to the data fitting pattern technique [19]. The practical relevance of our technique lies foremost in lacking the requirement of a detailed theoretical model encompassing all sensor parameters and noise sources for regularization of the data, otherwise crucial to perform an accurate interpolation procedure. The uncertainties on the calibration data are easily taken into account by means of repeated training with bootstrapped data. Furthermore, calibration of the sensor can be performed by having access only to the same set of input states and computational resources that are employed in the estimation process. This is particularly relevant in the perspective of large-scale fabrication of such devices, for which an analysis based on off-line characterization states would be impractical. This practical relevance is kept also with respect to cases in which an effective characterization can be carried out in terms of extra parameters—as long these remain fixed. Indeed, this would rely on a modeling of the imperfections, which might only be captured in part.

Further perspectives of this work can be found in extending the application of neural networks for the calibration of quantum sensors operating in the multiparameter regime, where multiple phases [64–67] and relevant system physical quantities [16,68], including noise, have to be measured simultaneously. Indeed, in those scenarios, reliable calibration methods become particularly crucial due to the increasing complexity of characterizing all the system parameters, as well as the computational overhead in handling a large amount of data.

We acknowledge support from the Amaldi Research Center funded by the Ministero dell'istruzione dell'università e della ricerca (Ministry of Education, University and Research) program "Dipartimento di Eccellenza" (CUP:B81I18001170001).

Note added.—Recently, we became aware of two works [69,70] in which neural networks have been applied to quantum state estimation and quantum simulation.

*ilaria.gianani@uniroma3.it

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[2] S. Dooley, M. Hanks, S. Nakayama, W. J. Munro, and K. Nemoto, npj Quantum Inf. **4**, 24 (2018).
[3] R. Nichols, T. R. Bromley, L. A. Correa, and G. Adesso, Phys. Rev. A **94**, 042101 (2016).
[4] T. Unden, P. Balasubramanian, D. Louzon, Y. Vinkler, M. B. Plenio, M. Markham, D. Twitchen, A. Stacey, I. Lovchinsky, A. O. Sushkov, M. D. Lukin, A. Retzker, B. Naydenov, L. P. McGuinness, and F. Jelezko, Phys. Rev. Lett. **116**, 230502 (2016).
[5] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Nat. Commun. **9**, 78 (2018).
[6] F. Reiter, A. S. Sørensen, P. Zoller, and C. A. Muschik, Nat. Commun. **8**, 1822 (2017).
[7] L. Cohen, Y. Pilnyak, D. Istrati, A. Retzker, and H. S. Eisenberg, Phys. Rev. A **94**, 012324 (2016).
[8] F. Albarelli, M. A. C. Rossi, D. Tamascelli, and M. G. Genoni, Quantum **2**, 110 (2018).
[9] M. Kacprowicz, R. Demkowicz-Dobrzański, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Nat. Photonics **4**, 357 (2010).
[10] N. Thomas-Peter, B. J. Smith, A. Datta, L. Zhang, U. Dorner, and I. A. Walmsley, Phys. Rev. Lett. **107**, 113603 (2011).
[11] R. Blandino, M. G. Genoni, J. Etesse, M. Barbieri, M. G. A. Paris, P. Grangier, and R. Tualle-Brouri, Phys. Rev. Lett. **109**, 180402 (2012).
[12] A. Lumino, E. Polino, A. S. Rab, G. Milani, N. Spagnolo, N. Wiebe, and F. Sciarrino, Phys. Rev. Applied **10**, 044033 (2018).
[13] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, Nat. Phys. **5**, 27 (2009).
[14] L. Zhang, H. B. Coldenstrodt-Ronge, A. Datta, G. Puentes, J. S. Lundeen, X.-M. Jin, B. J. Smith, M. B. Plenio, and I. A. Walmsley, Nat. Photonics **6**, 364 (2012).
[15] M. Altorio, M. G. Genoni, M. D. Vidrighin, F. Somma, and M. Barbieri, Phys. Rev. A **92**, 032114 (2015).
[16] E. Roccia, V. Cimini, M. Sbroscia, I. Gianani, L. Ruggiero, L. Mancino, M. G. Genoni, M. A. Ricci, and M. Barbieri, Optica **5**, 1171 (2018).
[17] G. Brida, L. Ciavarella, I. P. Degiovanni, M. Genovese, A. Migdall, M. G. Mingolla, M. G. A. Paris, F. Piacentini, and S. V. Polyakov, Phys. Rev. Lett. **108**, 253601 (2012).
[18] L. A. Rozema, D. H. Mahler, R. Blume-Kohout, and A. M. Steinberg, Phys. Rev. X **4**, 041025 (2014).
[19] M. Altorio, M. G. Genoni, F. Somma, and M. Barbieri, Phys. Rev. Lett. **116**, 100802 (2016).
[20] K. Rudinger, S. Kimmel, D. Lobser, and P. Maunz, Phys. Rev. Lett. **118**, 190502 (2017).
[21] K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press, Cambridge, MA, 2012).
[22] P. Simon, *Too Big to Ignore: The Business Case for Big Data* (Wiley, New York, 2013).
[23] V. Dunjko and H.-J. Briegel, Rep. Prog. Phys. **81**, 074001 (2018).
[24] E. Magesan, J. M. Gambetta, A. D. Córcoles, and J. M. Chow, Phys. Rev. Lett. **114**, 200501 (2015).
[25] S. Aaronson, Proc. R. Soc. A **463**, 3089 (2007).
[26] A. Rocchetto, S. Aaronson, S. Severini, G. Carvacho, D. Poderini, I. Agresti, M. Bentivegna, and F. Sciarrino, Sci. Adv. **5**, eaau1946 (2019).
[27] S. Yu, F. Albarran-Arriagada, J. C. Retamal, Y.-T. Wang, W. Liu, Z.-J. Ke, Y. Meng, Z.-P. Li, J.-S. Tang, E. Solano, L. Lamata, C.-F. Li, and G.-C. Guo, Adv. Quantum Technol. **2**, 1800074 (2019).
[28] N. Spagnolo, E. Maiorino, C. Vitelli, M. Bentivegna, A. Crespi, R. Ramponi, P. Mataloni, R. Osellame, and F. Sciarrino, Sci. Rep. **7**, 14316 (2017).
[29] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, New J. Phys. **14**, 103013 (2012).
[30] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. O'Brien, J. G. Rarity, A. Laing, and M. G. Thompson, Nat. Phys. **13**, 551 (2017).
[31] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).
[32] F. Flamini, N. Spagnolo, and F. Sciarrino, Quantum Sci. Technol. **4**, 024008 (2019).
[33] I. Agresti, N. Viggianiello, F. Flamini, N. Spagnolo, A. Crespi, R. Osellame, N. Wiebe, and F. Sciarrino, Phys. Rev. X **9**, 011013 (2019).
[34] L. Innocenti, H. Majury, T. Giordani, N. Spagnolo, F. Sciarrino, M. Paternostro, and A. Ferraro, Phys. Rev. A **96**, 062326 (2017).
[35] T. Giordani, E. Polino, S. Emiliani, A. Suprano, L. Innocenti, H. Majury, L. Marrucci, M. Paternostro, A. Ferraro, N. Spagnolo, and F. Sciarrino, Phys. Rev. Lett. **122**, 020503 (2019).
[36] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Phys. Rev. X **8**, 031086 (2018).
[37] H.-J. Briegel and G. De las Cuevas, Sci. Rep. **2**, 400 (2012).

[38] L. Innocenti, L. Banchi, A. Ferraro, S. Bose, and M. Paternostro, arXiv:1803.07119.
[39] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, Phys. Rev. X **8**, 031084 (2018).
[40] M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger, Phys. Rev. Lett. **116**, 090405 (2016).
[41] A. A. Melnikov, H. Poulsen Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, Proc. Natl. Acad. Sci. U.S.A. **115**, 1221 (2018).
[42] U. L. Heras, U. Alvarez-Rodriguez, E. Solano, and M. Sanz, Phys. Rev. Lett. **116**, 230504 (2016).
[43] M. August and X. Ni, Phys. Rev. A **95**, 012335 (2017).
[44] N. Wiebe, C. Granade, A. Kapoor, and K. M. Svore, arXiv:1511.06458.
[45] N. Wiebe and C. Granade, Phys. Rev. Lett. **117**, 010503 (2016).
[46] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, Phys. Rev. Lett. **118**, 100503 (2017).
[47] A. Hentschel and B. C. Sanders, Phys. Rev. Lett. **104**, 063603 (2010).
[48] A. Hentschel and B. C. Sanders, Phys. Rev. Lett. **107**, 233601 (2011).
[49] N. B. Lovett, C. Crosnier, M. Perarnau-Llobet, and B. C. Sanders, Phys. Rev. Lett. **110**, 220501 (2013).
[50] M. Caciotta, S. Giarnetti, and F. Leccese, *19th IMEKO World Congress* (Curran Associates, Inc, Red Hook, 2009), Vol. 1, p. 586.
[51] M. Caciotta, S. Giarnetti, F. Leccese, B. Orioni, M. Oreggia, C. Pucci, and S. Rametta, Measurement **78**, 366 (2016).
[52] C. Robert and G. Casella, *Monte Carlo Statistical Method* (Springer, New York, 2004).
[53] S. Slussarenko, M. M. Weston, H. M. Chrzanowski, L. K. Shalm, V. B. Verma, S. W. Nam, and G. J. Pryde, Nat. Photonics **11**, 700 (2017).
[54] MathWorks, MATLAB Deep Learning Toolbox documentation (The MathWorks, Inc., Natick, 2019).
[55] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.123.230502 for more details on the comparison with different techniques and on the network parameters, which includes Refs. [56–60].
[56] M. G. A. Paris, Int. J. Quantum. Inform. **07**, 125 (2009).
[57] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).
[58] Y. Li, L. Pezz, M. Gessner, Z. Ren, W. Li, and A. Smerzi, Entropy **20**, 628 (2018).
[59] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Phys. Rev. A **77**, 032322 (2008).
[60] E. Roccia, V. Cimini, M. Sbroscia, I. Gianani, L. Ruggiero, L. Mancino, M. G. Genoni, M. A. Ricci, and M. Barbieri, Optica **5**, 1171 (2018).
[61] M. T. Hagan and M. B. Menhaj, IEEE Trans. Neural Networks **5**, 989 (1994).
[62] D. Marquardt, J. Soc. Ind. Appl. Math. **11**, 431 (1963).
[63] M. G. A. Paris, Int. J. Quantum. Inform. **07**, 125 (2009).
[64] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Phys. Rev. Lett. **111**, 070403 (2013).
[65] M. A. Ciampini, N. Spagnolo, C. Vitelli, L. Pezzé, A. Smerzi, and F. Sciarrino, Sci. Rep. **6**, 28881 (2016).
[66] L. Pezzè, M. A. Ciampini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and A. Smerzi, Phys. Rev. Lett. **119**, 130504 (2017).
[67] E. Polino, M. Riva, M. Valeri, R. Silvestri, G. Corrielli, A. Crespi, N. Spagnolo, R. Osellame, and F. Sciarrino, Optica **6**, 288 (2019).
[68] V. Cimini, I. Gianani, L. Ruggiero, T. Gasperi, M. Sbroscia, E. Roccia, D. Tofani, F. Bruni, M. A. Ricci, and M. Barbieri, Phys. Rev. A **99**, 053817 (2019).
[69] A. Macarone-Palmier, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J. Biamonte, and S. Kulik, arXiv:1904.05902.
[70] G. Torlai, B. Timar, E. P. L. van Nieuwenburg, H. Levine, A. Omran, A. Keesling, H. Bernien, M. Greiner, V. Vuletić, M. D. Lukin, R. G. Melko, and M. Endres, arXiv:1904.08441.