A Data-Driven Approach to Quantum Cross-Platform Verification

Ya-Dong Wu,† Yan Zhu,† Ge Bai,† Yuexuan Wang,‡ and Giulio Chiribella\*\,\†\,§

\*QICI Quantum Information and Computation Initiative, Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong
\‡AI Technology Lab, Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong
\§Department of Computer Science, Parks Road, Oxford, OX1 3QD, United Kingdom

The task of testing whether two uncharacterized devices behave in the same way, known as cross-platform verification, is crucial for benchmarking quantum simulators and near-term quantum computers. Cross-platform verification becomes increasingly challenging as the system’s dimensionality increases, and has so far remained intractable for continuous variable quantum systems. In this Letter, we develop a data-driven approach, working with limited noisy data and suitable for continuous variable quantum states. Our approach is based on a convolutional neural network that assesses the similarity of quantum states based on a lower-dimensional state representation built from measurement data. The network can be trained offline with classically simulated data, and is demonstrated here on non-Gaussian quantum states for which cross-platform verification could not be achieved with previous techniques. It can also be applied to cross-platform verification of quantum dynamics and to the problem of experimentally testing whether two quantum states are equivalent up to Gaussian unitary transformations.

Introduction. Quantum cross-platform verification [1–3], the problem of comparing unknown quantum states from experimental data, is crucial for benchmarking quantum simulations and near-term quantum computers [4]. Its role is especially important for intermediate, scale quantum devices that are hard to simulate classically. A natural approach in this context is to choose a trusted device as a reference standard, and to compare other untrusted devices to it. For example, the trusted device could be built and maintained by a quantum computing company, while other devices could be owned by users in distant laboratories. The comparison of two unknown quantum devices can be achieved by estimating their overlap [1,5,10], a fundamental primitive with many applications in quantum information, including quantum state discrimination and classification [11–13]. Recently, Elben et al. proposed an approach [14] for experimentally estimating the overlap between two multi-qubit states from the cross correlations of Pauli measurements. While this approach works well for systems with a small number of qubits, it becomes more difficult to implement as the number of qubits grows, as it requires sampling over an exponentially large set of Pauli measurements, randomly drawn from the uniform distribution [15]. Moreover, this approach requires an estimate of the correlations between the measurements performed on the two unknown states. In scenarios where the states are produced in distant laboratories, achieving these correlations of uniformly random measurements implies an additional communication cost.

An alternative approach to characterize quantum states from limited measurement data is provided by deep neural networks [16–23]. Recently, neural networks for learning similarities among quantum states have been developed for the tasks of state discrimination and classification [24–29]. These tasks, however, are fundamentally different from quantum cross-platform verification, because they generally presuppose that the state of the system is identifiable by a finite set of predetermined labels. In cross-platform verification, this assumption is too restrictive, as it prevents the application to continuous families of quantum states, such as the family of coherent states in quantum optics.

In this Letter, we develop a convolutional neural network achieving cross-platform verification of quantum states from a general, continuously-parametrized state family, working with noisy and incomplete measurement data, without requiring randomization over a exponential sets of measurements, or correlations between the measurements performed on different states. The network can be trained with classically simulated data from a fiducial set of quantum states sharing structural similarities with the quantum states to be compared. After training, it embeds the measurement data into a low-dimensional feature space in a way that reflects the similarity of quantum states and enables cross-platform verification. Our approach is inspired by a classical technique for the recognition of human faces blurred and incomplete images [30], a task that shares similarities with the task of comparing continuous-variable quantum states from finite-statistics approximations of their Wigner function.

We test the performance of our network on a number of non-Gaussian quantum states whose experimental preparation could not be efficiently verified with previous techniques [1,31,32]. After an offline training on simulated data, the trained network is tested with both simulated data and actual experimental data from a resonator coupled to a superconducting transmon qubit. The re-
sults indicate a high success rate for cross-platform quantum verification even using incomplete measurement data that are insufficient for a reliable state estimation with maximum likelihood estimation. Our approach can also be extended to cross-platform verification of quantum dynamics and cross-platform verification of equivalence of quantum states up to a given set of unitary operations, such as the set of Gaussian unitaries.

**Framework.** Two experimenters, Alice and Bob, own two quantum devices producing copies of two unknown quantum states $\rho$ and $\sigma$, respectively. Alice and Bob want to determine whether their devices prepare the same quantum state, that is, whether $\rho = \sigma$. Alice (Bob) can perform a limited set of quantum measurements, denoted by $\mathcal{M}_A$ ($\mathcal{M}_B$). Each measurement $M \in \mathcal{M}_A$ ($M \in \mathcal{M}_B$) corresponds to a positive operator-valued measure (POVM), that is, a set of positive operators $M = (M_j)_{j=1}^k$ acting on the system’s Hilbert space and satisfying the normalization condition $\sum_{j=1}^k M_j = 1$. To compare their states, Alice and Bob pick two subsets of measurements $S_A \subset \mathcal{M}_A$ and $S_B \subset \mathcal{M}_B$, respectively. The measurements can be chosen independently and randomly, although there is no need for Alice and Bob to sample them from the uniform distribution, or from any specific probability distribution. In general, the sets of performed measurements $S_A$ and $S_B$ do not need to be informationally complete.

By performing a measurement $M \in S_A$ ($M \in S_B$) on multiple copies of $\rho$ ($\sigma$), Alice (Bob) obtains a vector of experimental frequencies $d$. This classical data is then sent to a verifier, Charlie, whose task is to decide whether $\rho$ and $\sigma$ are same state. For each measurement $M \in S_A$ ($S_B$), Alice (Bob) provides Charlie with a pair $(m, d)$, where $m$ is a parametrization of the measurement $M$. Here, the parametrization $m$ could be either a full description of the POVM $M$, or a lower-dimensional parametrization valid only for measurements in the set $\mathcal{M}_A$ ($\mathcal{M}_B$). Ideally, if Charlie had access to the statistics of an informationally complete set of quantum measurements, performed on an unlimited number of copies of the quantum states, then he could obtain complete classical descriptions of both $\rho$ and $\sigma$, and could calculate their fidelity. For higher-dimensional quantum systems, however, this approach is not viable, because only limited data from a limited set of measurements are available. In the following, we develop an alternative approach in which the similarity of quantum states is assessed in terms of the Euclidean distance between low-dimensional state representations constructed by a deep neural network.

**Neural network for cross-platform state verification.** Here we introduce a deep neural network that determines whether two unknown states are same or not. We call our network StateNet, in analogy to FaceNet [30], a popular neural network for the identification of human faces.

StateNet uses a convolutional neural network [33] to produce a compact representation of quantum states. For the training, we choose a set of fiducial states and classically simulate measurement data with finite statistics. For each fiducial state $\tau$, we denote by $\chi^{(i)}_\tau$ the measurement data $\left(\mathbf{m}^{(i)}_j, d^{(i)}_j\right)_{j=1}^K$ of $\tau$ obtained in the $i$-th simulated experiment, where $i \in \{1, 2, \cdots, K\}$ and $K$ is the number of simulated measurement data sets for each state. Each data set $\chi^{(i)}_\tau$ is fed into a deep neural network to produce a state representation $r^{(i)}_\tau \in \mathbb{R}^n$, where $n$ is the dimension of the state representation space. In the training phase, we optimize the parameters of the convolutional neural network with respect to a loss function, called triplet loss [30]. In each round of optimization, we choose a representation $r^{(l)}_\tau$ as a reference with a fixed state $\tau$ and $l$. Then, we compare the Euclidean distances $||r^{(l)}_\tau - r^{(i)}_\tau||_2$ of state representations associated to different measurements on the same state $\tau$, with the Euclidean distances $||r^{(l)}_\tau - r^{(m)}_\xi||_2$ of state representations associated to different states $\tau$ and $\xi$, for every $1 \leq i \leq K$, $i \neq l$, $1 \leq m \leq K$, and for a randomly chosen fiducial state $\xi \neq \tau$. Optimizing triplet loss aims to minimize $||r^{(l)}_\tau - r^{(i)}_\tau||_2$ for multiple rounds of experiments with respect to same fiducial state $\tau$, and simultaneously maximize $||r^{(l)}_\tau - r^{(m)}_\xi||_2$ for each pair of different fiducial states $\tau \neq \xi$.

After the training is concluded, the network maps measurement data from the same quantum state to representation vectors close to each other, whereas it maps measurement data from different states to representation vectors that are further apart. Quantum states are then compared by evaluating the Euclidean distance between the corresponding state representations. To decide whether two state representations correspond to the the same quantum state, the network uses a threshold value that balances both the false rejection rate and the false acceptance rate over a new set of unseen measurement data obtained from the fiducial states. Hereafter, this set will be called the validation data set. The complete

![Figure 1](https://via.placeholder.com/150)

**Figure 1.** Assessing the similarity of characterized vs uncharacterized quantum states. For fully characterized quantum states (left), the similarity can be assessed by computing their fidelity or any other distance measure. In cross-platform verification, instead, only measurement data from uncharacterized quantum states are available (right). The task is then to determine whether two given data sets have been generated from the same state or not. In general, the data sets can refer to different measurements, which may be noisy and incomplete.
Cross-platform verification of continuous-variable quantum states. The structure of StateNet is well adapted to continuous-variable states, which provide a natural testbed for our model. A continuous-variable quantum state \( \rho \) is characterized by its Wigner function \( W_\rho (a) := \frac{1}{\pi} \text{tr} (\rho (\hat{a}^\dagger \hat{a} - |a|^2)) \), which is a quantum superposition of two coherent states. An estimate of the value of \( W_\rho (a) \) at any phase-space point \( \alpha \) can be achieved by applying a displacement operation \( D(\alpha) \) on \( \rho \) followed by a measurement of the photon number parity, corresponding to the operator \((-1)^\hat{n}\). This measurement scheme has been implemented in a number of experiments, including \( e.g. \) \([36,37]\), and is widely used for the characterization of quantum states in circuit quantum electrodynamics \([38,39]\).

Suppose that Alice (Bob) performs measurements to estimate the Wigner function on a finite number of points, chosen at random from a square grid over the phase space. The set of all these measurements composes the sets \( M_A = M_B := M \). Alice (Bob) randomly chooses a subset of points in the grid to perform measurements of the Wigner function, which is a subset \( S_A \) \((-S_B\) of \( M \). In general, the points chosen by Alice and Bob need not be the same. After a finite set has been implemented in measurement runs, Alice (Bob) obtains a two-dimensional data image, where some pixels are missing and the value at each of the existing pixels is an estimate of the Wigner function at the associated phase-space point. As a result of the finite statistics, the image will generally be blurred.

To illustrate our method, we test it on cat states, of the form \([30,41]\)

\[
\frac{1}{\sqrt{2(1 + \exp(-2|\alpha|^2))}} (|\alpha\rangle + |\alpha\rangle \rangle, \tag{1}
\]

which is a quantum superposition of two coherent states \(|\alpha\rangle \) and \(|\alpha\rangle \rangle\) with opposite amplitudes, where \( \alpha \in [1,2] \). We train StateNet using simulated measurement data from ideal cat states in Eq. (1), as well as noisy cat states with a fixed amount of thermal noise. After the training is concluded, we test the performance of StateNet in the cross-verification of pairs of noisy cat states degraded by photon loss. For each pair of noisy cat states \( \rho \) and \( \sigma \), we take \( \rho \) to play the role of the reference state, and make it close to its noiseless counterpart \( \rho_{\text{ideal}} \) with fidelity 99%. On the other hand, we regard \( \sigma \) as the untrusted state that needs to be verified, and allow it to be generally noisier, allowing the fidelity with the trusted state \( \rho \) to range between between 84% and 100%. We plot the rejection rate over test pairs with respect to the fidelity between each pair of noisy cat states for multiple scenarios as shown in Fig. [2]. Note that when fidelity equals one, i.e., \( \rho = \sigma \), the plots show the false rejection rates.

Numerical results show that when we either increase the number of phase-space points for the Wigner-function measurements or raise the number of runs per measurement setting, the trained network achieves higher success rate in distinguishing different noisy states with quantum fidelity between 0.9 and 0.99. Notably, the amount of measurement data that the neural network receives as input is not enough for a reliable quantum state tomography using maximum likelihood estimation \([42]\) or generative adversarial networks \([21]\).

Besides data size, the performance of neural networks is also affected by state complexity, which can be measured by its nonclassicality \([33]\). To test this aspect, we increase state complexity by considering cat-like states that are superposition of four coherent states instead of two. The results, shown in the Supplemental Note 3, show that the success rate drops by an amount the depends on the fidelity between the two tested states.

It is also useful to visualize the state representations obtained by embedding different measurement data in feature space. To this purpose, we feed the state representations of training states with \( \alpha \in \{1,1.1, \ldots, 1.9\} \), into a t-distributed stochastic neighbor embedding (t-SNE) algorithm \([44]\) to project the representation vectors into a two-dimensional plane, according to their similarities in feature space. The results, provided in Fig. [3] show that measurement data associated to the same state effectively correspond to nearby vectors, while measurement data associated to different states typically correspond to distant vectors. It is worth stressing that StateNet was not provided with any state parameter (\( e.g. \) it was not provided the value of \( \alpha \)), nor it was provided any information about the fidelity of quantum states. The notion of distance visualized in the figure was developed by the network as a way to minimize the loss function in the training phase.

In addition to simulated data, we also use actual experimental data for the testing. For this test, we first train StateNet with simulated measurement data of states

\[
|\phi_n\rangle := \text{SNAP}(\theta) |\alpha\rangle, \tag{2}
\]

where \( \text{SNAP}(\theta) := \sum_n \exp(i \theta_n) |n\rangle \langle n| \) is a selective number-dependent arbitrary phase gate \([39,45]\) with
Figure 3. 2D projections of state representations corresponding to cat states in Eq. 1 with \( \alpha \in \{1, 1.1, \ldots, 1.9\} \) obtained by t-SNE. In the figure, points of the same color correspond to state representations associated to different measurements on the same quantum state.

\[
\theta_n = \begin{cases} 
\pi & n = 0, 1 \\
0 & n \geq 2
\end{cases}
\]

Then we test our trained model on experimental data obtained from a noisy implementation [21, 37] of state \( |\phi_1\rangle \) in a resonator coupled to a superconducting transmon qubit. We randomly choose 5\% of \( 81 \times 81 = 6561 \) noisy experimental data pixels (as shown in Fig. 4(b) in Ref. [21]) as a test data set and check whether our model accepts two resulted data sets as obtained from a same quantum state. The resulting acceptance rate, averaged over multiple trails, is 98\%.

Cross-Platform Verification of Quantum Dynamics. The structure of StateNet can also be adapted to the task of determining whether two quantum devices implement the same quantum dynamics. We illustrate this fact on an important quantum dynamics for continuous variable quantum computing, namely the dynamics generated by Kerr’s nonlinear interaction [40, 46]. Suppose a quantum system, initialized in a coherent state \( |\alpha\rangle \), is subject to the Kerr Hamiltonian \( H_{\text{kerr}} = \pi \hat{a}^2 \hat{a}^\dagger \). Here we consider both an ideal closed-system dynamics, governed by the Schrödinger’s equation with Hamiltonian \( H_{\text{kerr}} \), and a lossy open-system dynamics, governed by the Lindblad master equation \( \dot{\rho} = -i[H_{\text{kerr}}, \rho] + \eta D(\hat{a})(\rho) \), where \( D(\hat{a})(\rho) := \hat{a}\rho - \frac{1}{2}[\hat{a}\rho\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\rho] \) and \( \eta \) is the loss rate. The problem is to determine whether the quantum states generated by these two different quantum evolutions at subsequent snapshots of time are equal or not.

In our numerical experiments, we consider 10 snapshots \( t_i := i/10 \), with \( i = 0, 1, \ldots, 10 \). We train a neural network with the simulated measurement data of quantum states undergoing various unitary evolutions governed by \( H \). Then, we test this trained model for cross-verification of a unitary dynamics against its lossy counterpart (with loss rate \( \eta = 0.5 \)) starting from the same initial states. Fig. 4 shows that the rejection rates increase with time, in agreement with the fact that the deviations between the ideal and lossy dynamics become more visible at larger times. This behaviour is observed in two different measurement scenarios: one in which all the phase-space points in the grid are measured and the other in which only a subset of randomly chosen points is measured.

Cross-Platform Verification of Equivalence up to Gaussian Unitary Operations. As a further application, we provide a variant of StateNet that can decide whether two quantum quantum states are the same up to a unitary transformation in a given set. This functionality can be used to make cross-platform verification robust to unknown systematic unitary errors taking place on individual quantum devices. A common example of such errors are the errors introduced by poorly calibrated displacement pulses, phase rotations and squeezing [47].

In the Wigner function representation, the combination of displacements, rotations, and squeezing corresponds to the following affine transformation in phase space

\[
\begin{pmatrix} x \\ p \end{pmatrix} \rightarrow \begin{pmatrix} \zeta (x \cos \phi + x \sin \phi + \Delta_x) \\ 1/\zeta (p \sin \phi + p \cos \phi + \Delta_p) \end{pmatrix},
\]

where \( \phi \) is a rotation angle, \( \Delta_x \) and \( \Delta_p \) are shifts in position and momentum, respectively, and \( \zeta \) is a squeezing parameter. Each data image, with respect to quantum states in Eq. 2 with \( \theta_0, \theta_1 \in [\pi/2, 3\pi/2] \) and \( \theta_n = 0 \) for \( n \geq 2 \), undergoes an affine transformation with uniformly random parameters \( \phi \in [0, \pi], \Delta_x, \Delta_p \in (-1, 1) \) and \( \zeta \in [5/6, 6/5] \). Then data images with same values of \( \theta \) but different affine transformations are labeled the same, while data images with different values of \( \theta \) are identified with different labels.

We train StateNet with data corresponding to \( \theta_0, \theta_1 \in \{\pi/2, 3\pi/4, \ldots, 3\pi/2\} \) and then test it over unseen data images corresponding to \( \theta_0, \theta_1 \in \{5\pi/8, 7\pi/8, \ldots, 11\pi/8\} \) to decide whether two data images correspond to the same quantum state up to certain distortion effect. Numerical results show that both the average rate to reject different labels and the average rate to accept same labels is around 90\%. Fig. 5 shows ex-
data images of quantum states with different values of different affine transformations, and different columns contain a pair of data images of the same quantum state with displacement, rotations, and squeezing. Each column contains 4900 pixels randomly chosen from $81 \times 81 = 6561$ pixels, and the value at each pixel is an estimate of the Wigner function by sampling the true outcome probability distribution 500 times.

Figure 5. Cross-platform verification of equivalence up to displacement, rotations, and squeezing. Each column contains a pair of data images of the same quantum state with different affine transformations, and different columns contain data images of quantum states with different values of $\theta_0$ and $\theta_1$. The number between each pair of data images is the Euclidean distance between their state representations produced by StateNet. Each data image contains 4900 pixels randomly chosen from $81 \times 81 = 6561$ pixels, and the value at each pixel is an estimate of the Wigner function by sampling the true outcome probability distribution 500 times.

Examples of image data with both the same and different labels, where the values between pairs of images are the Euclidean distances between the corresponding state representations produced by StateNet. By balancing both the false rejection rate and the false acceptance rate over a validation set, we obtain a distance threshold 0.4, which makes our model accept all pairs of states that are same up to an affine transformation, and reject those that belong to different classes in Fig. 5.

Conclusions. In this work we introduced a model of neural network for cross-platform verification of unknown quantum states. Our model has two important benefits: (i) applicability to continuous state families (the model does not require the possible quantum states to correspond to a finite set of predefined labels), and (ii) no need of correlated measurements (experimenters can independently choose which set of measurements are performed in their laboratories, without the need of coordinating their measurement settings). The applicability to continuous state families makes our method suited to continuous variable systems, which are difficult to treat with previous approaches. We stress, however, that in principle our framework is not limited to the continuous variable setting, and can also be adapted to cross-platform verification of multi-qubit states, such as Hamiltonian ground states. The adaptation of StateNet to the multi-qubit setting is a promising direction, which we plan to address in a forthcoming study.

Our work opens up the application of metric learning techniques [18] to the characterization of quantum systems produced by noisy intermediate scaled quantum devices [49]. Furthermore, it sheds light on how machines can discover physical concepts, such as the similarity of two quantum states, without any hard-coded information about the corresponding physical theories [50-52].

Acknowledgement. This work was supported by funding from the Hong Kong Research Grant Council through grants no. 17300918 and no. 17307520, through the Senior Research Fellowship Scheme SRFS2021-7502, the Croucher Foundation, and by the John Templeton Foundation through grant 61466, The Quantum Information Structure of Spacetime (qiss.fr). YXW acknowledges funding from the National Natural Science Foundation of China through grants no. 61872318. Research at the Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

References:

[1] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, Cross-platform verification of intermediate scale quantum devices, Phys. Rev. Lett. 124, 010504 (2020).
[2] S. Flammia, Quantum computer crosscheck, Physics 13, 3 (2020).
[3] J. Carrasco, A. Elben, C. Kokail, B. Kraus, and P. Zoller, Theoretical and experimental perspectives of quantum verification, PRX Quantum 2, 010102 (2021).
[4] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, Quantum certification and benchmarking, Nat. Rev. Phys. 2, 382 (2020).
[5] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Quantum fingerprinting, Phys. Rev. Lett. 87, 167902 (2001).
[6] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, Learning the quantum algorithm for state overlap, New J. Phys. 20, 113022 (2018).
[7] U. Chabaud, E. Diamanti, D. Markham, E. Kashefi, and A. Joux, Optimal quantum-programmable projective measurement with linear optics, Phys. Rev. A 98, 062318 (2018).
[8] M. Fanizza, M. Rosati, M. Skotiniotis, J. Calsamiglia, and V. Giovannetti, Beyond the swap test: Optimal estimation of quantum state overlap, Phys. Rev. Lett. 124, 060503 (2020).
[9] L. Guerini, R. Wiersema, J. F. Carrasquilla, and L. Aolita, Quasiprobabilistic state-overlap estimator for nisq devices, arXiv preprint arXiv:2112.11618 (2021).
[10] A. Anshu, Z. Landau, and Y. Liu, Distributed quantum inner product estimation, in Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (2022) pp. 44–51.
[11] C. W. Helstrom, Quantum detection and estimation theory, Journal of Statistical Physics 1, 231 (1969).
[12] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, Vol. 1 (Springer Science & Business Media, 2011).

[13] M. Sasaki, A. Carlini, and R. Jozsa, Quantum template matching, Phys. Rev. A 64, 022317 (2001)

[14] V. Dunjko and H. J. Briegel, Machine learning & artificial intelligence in the quantum domain: a review of recent progress, Reports on Progress in Physics 81, 074001 (2018).

[15] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, The randomized measurement toolbox, arXiv preprint arXiv:2203.11374 (2022).

[16] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, Nat. Phys. 14, 447 (2018).

[17] G. Torlai and R. G. Melko, Latent space purification via neural density operators, Phys. Rev. Lett. 120, 240503 (2018).

[18] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, Nat. Mach. Intell. 1, 155 (2019).

[19] 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, Integrating neural networks with a quantum simulator for state reconstruction, Phys. Rev. Lett. 123, 230504 (2019).

[20] E. S. Tishkov, V. Tisnová, A. E. Ulamov, A. Lovsovsky, and A. K. Fedorov, Experimental quantum homodyne tomography via machine learning, Optica 7, 448 (2020).

[21] S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Quantum state tomography with conditional generative adversarial networks, Phys. Rev. Lett. 127, 140502 (2021).

[22] A. W. R. Smith, J. Gray, and M. S. Kim, Efficient quantum state sample tomography with basis-dependent neural networks, PRX Quantum 2, 020348 (2021).

[23] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, Dynamically encoding quantum information using 100-photon schrödinger cat states, Science 342, 607 (2013).

[24] V. Vuletić, M. D. Lukin, R. G. Melko, and M. Endres, Quantum template matching, Phys. Rev. A 86, 062318 (2012).

[25] A. Kenfack and K. Życzkowski, Negativity of the wigner function as an indicator of non-classicality, J. Opt. B: Quantum Semiclass. Opt. 9, 396 (2004).

[26] T. Fösel, S. Krastanov, F. Marquardt, and L. Jiang, Efficient cavity control with re-inforcement learning, Phys. Rev. X 12, 011059 (2022).

[27] Y. V. Sivak, A. Eckbusch, H. Liu, B. Royer, I. Tsiontsou, and M. H. Devoret, Model-free quantum control with reinforcement learning, Phys. Rev. X 12, 011059 (2022).

[28] B. Vlastakis, G. Kirchmair, Z. Leghtas, S. E. Nigg, L. Frunzio, S. M. Girvin, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Deterministically encoding quantum information using 100-photon schrödinger cat states, Science 342, 607 (2013).

[29] M. Kidra, M. Kervinen, I. Strandberg, S. Ahmed, M. Scigliuzzo, A. Osman, D. P. Lozano, M. O. Tholén, R. Borgani, D. B. Haviland, G. Ferrini, J. Bylander, A. F. Kockum, F. Quijandria, P. Delsing, and S. Gasparinetti, Robust preparation of wigner-negative states with optimized snap-displacement sequences, PRX Quantum 3, 030301 (2022).

[30] A. I. Lvovsky and M. G. Raymer, Continuous-variable optical quantum-state tomography, Rev. Mod. Phys. 81, 299 (2009).

[31] L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert, Reliable quantum certification of photonic state preparations, Nat. Commun. 6, 8498 (2015).

[32] Y.-D. Wu, G. Bai, G. Chiribella, and N. Liu, Efficient verification of continuous-variable quantum states and devices without assuming identical and independent operations, Phys. Rev. Lett. 126, 240503 (2021).

[33] V. V. Albert, L. Frunzio, L. Jiang, and R. J. Schoelkopf, Cavity state manipulation using photon-number selective phase gates, Phys. Rev. Lett. 115, 130503 (2015).

[34] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, Dynamically protected cat-qubits: a new paradigm for universal quantum computation, New J. Phys. 16, 045014 (2014).

[35] S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Quantum state tomography with conditional generative adversarial networks, Phys. Rev. Lett. 127, 140502 (2021).

[36] A. W. R. Smith, J. Gray, and M. S. Kim, Efficient quantum state sample tomography with basis-dependent neural networks, PRX Quantum 2, 020348 (2021).

[37] Y. Zhu, Y.-D. Wu, G. Bai, D.-S. Wang, Y. Wang, and G. Chiribella, Flexible learning of quantum states with generative query neural networks, arXiv preprint arXiv:2202.06804 (2022).

[38] E. Magesan, J. M. Gambetta, A. D. Corcoles, and J. M. Chow, Machine learning for discriminating quantum measurement trajectories and improving readout, Phys. Rev. Lett. 114, 200501 (2015).

[39] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, Nat. Phys. 13, 431 (2017).

[40] J. Gao, L.-F. Qiao, Z.-Q. Jiao, Y.-C. Ma, C.-Q. Hu, R.-J. Ren, A.-L. Yang, H. Tang, M.-H. Yung, and X.-M. Jin, Experimental machine learning of quantum states, Phys. Rev. Lett. 120, 240501 (2018).

[41] V. Cimini, M. Barbieri, N. Tresp, M. Walschaers, and V. Parigi, Neural networks for detecting multimode wigner negativity, Phys. Rev. Lett. 125, 106004 (2020).

[42] S. J. Wetzel, R. G. Melko, J. Scott, M. Panju, and V. Ganesh, Discovering symmetry invariants and conserved quantities by interpreting siamese neural networks, Phys. Rev. Research 2, 033499 (2020).

[43] X. Zhang, M. Luo, Z. Wen, Q. Feng, S. Pang, W. Luo, and X. Zhou, Direct fidelity estimation of quantum states using machine learning, Phys. Rev. Lett. 127, 130503 (2021).

[44] F. Schroff, D. Kalenichenko, and J. Philbin, Facenet: A unified embedding for face recognition and clustering, in CVPR (2015) pp. 815–823.

[45] M. Schultz and T. Joachims, Learning a distance metric for relative comparisons, Advances in neural information processing systems 16 (2003).
SUPPLEMENTARY INFORMATION

Supplementary Note 1  Implementation details of StateNet

A. Structure of StateNet

As shown in Supplementary Fig. 6, our proposed StateNet for learning quantum state similarity consists of a representation network \( f_\xi \) and a \( L_2 \) normalization layer.

\[
\left( m^{(k)}_i, d^{(k)}_i \right)_{i=1}^s \rightarrow f_\xi \rightarrow L_2 \rightarrow r_p^{(k)}
\]

Figure 6. Structure of StateNet.

The representation network \( f_\xi \) is composed of multiple convolutional layers, a max pooling layer, a dropout layer and a last full-connected layer \([53]\) and we depict its structure in Supplementary Fig. 7. We denote all trainable parameters in the representation network as \( \xi \). The input of the representation network is a set of \( s \) measurement results of a quantum state and we denote them as \( \left( m^{(k)}_j, d^{(k)}_j \right)_{j=1}^s \), where \( m_j \) is parameterization of the measurement and \( d_j \) is a vector of experimental measurement outcome frequencies on the state. A \( L_2 \) normalization layer following the representation network is used for rescaling the output of the representation network to have Euclidean norm 1.

Figure 7. Structure of the representation network.

B. Training Details

a. Loss function. In the training, we adopt a loss function called Triplet Loss used in FaceNet. The Triplet Loss minimizes the distance between an anchor and a positive, both of which belongs to the same quantum state, and maximizes the distance between the anchor and a negative which belongs to different quantum states. The detailed steps of training is presented in Algorithm 1.
Algorithm 1: Training of StateNet for learning quantum state similarity.

Data: number of states in training set $N$, state measurement results $\{\{\chi^{(k)}_{\rho_i}\}_{i=1}^{K}\}_{k=1}^{N}$, maximum number of epochs $E$, learning rate $\delta$.

Initialize parameters $\xi$ and randomly, $e = 0$;

while $e < E$ do
  Calculate the state representation $r^{(k)}_{\rho_i}$ for each state measurement result $\chi^{(k)}_{\rho_i}$;
  $L = 0$;
  for $i_1 = 1$ to $N$ do
    Randomly select $K$ state representations $r^{(k)}_{\rho_j}$ from all state representations, where $j \neq i_1$, and denote them as $\{r^{(k)}_{\rho_{neg}}\}_{k=1}^{K}$;
    for $i_2 = 1$ to $K$ do
      Select $r^{(i_2)}_{\rho_1}$ as a reference state representation;
      Calculate the Euclidean distance $\text{dist}_k := ||r^{(i_2)}_{\rho_1} - r^{(k)}_{\rho_1}||_2$ between $r^{(i_2)}_{\rho_1}$ and each $r^{(k)}_{\rho_1}$, where $k \in \{1, 2, \cdots, K\}$;
      Calculate the Euclidean distance $||r^{(i_2)}_{\rho_1} - r^{(k)}_{\rho_{neg}}||_2$ between $r^{(i_2)}_{\rho_1}$ and each element in $\{r^{(k)}_{\rho_{neg}}\}_{k=1}^{K}$ and find the minimum $\text{dist}_{\text{neg}} := \min_{k} ||r^{(i_2)}_{\rho_1} - r^{(k)}_{\rho_{neg}}||_2$;
      $L = L + \sum_{k=1}^{K} \text{dist}_k - K \cdot \text{dist}_{\text{neg}}$;
    end for
  end for
  Update $\xi$ as $\xi = \xi - \delta \nabla_{\xi} L$;
  $e = e + 1$;
end while

b. Training set and validation set. Both the training and the validation sets are simulation data in our experiments. We usually generate all of the simulated data and apportion the data into training and validation sets, with an 80-20 split. The model is first trained over the training set and the threshold is then chosen over the validation set by minimizing both the false rejection rate and the false acceptance rate.

c. Initialization and learning rate. We randomly initialize the parameters of the StateNet in each experiment and set the initial learning rate as 0.01. We decreases the learning rate as the number of iterations increases.

d. Number of epochs and training time. The maximum number of epochs $E$ is usually set as 500 in the training. The training time depends on the specific dataset and tasks we consider but it is less than one hour in all of the experiments shown in this paper.

e. Dimension of state representations. The dimension of the state representation space is set to be $n = 32$ for all the numerical experiments.

Supplementary Note 2 Dataset

In this section, we will introduce the types and the generation method of the fiducial states used in the settings we consider.

a. Cross-Platform Verification of Quantum States The training set is composed of the simulated measurement data of both ideal cat states in Eq. (1) with $\alpha \in \{1, 1.1, \cdots, 2\}$ and noisy cat states with certain amount of thermal noises. The test set is composed of noisy cat states with photon loss errors. In order to test our model for noisy experimental data, we also use simulated measurement data of states in Eq. (2) with $\alpha \in \{\frac{\pi}{2}, \frac{5\pi}{8}, \cdots, \frac{3\pi}{2}\}$ to train our proposed model. All of these states are generated by qutip [54].

b. Cross-Platform Verification of Quantum Dynamics In this experiment, fiducial states in the training set is composed of all the states at different time step $t_i$ when the states are initialized in coherent states with $\alpha \in \{1, 1.1, \cdots, 2\}$. All of these states are generated by qutip [54]. The test set consists of quantum states under the noiseless quantum evolution and the noisy quantum evolution.

c. Cross-Platform Verification of Equivalence up to Unitary Operations In this setting, our proposed StateNet is trained by the measurement data corresponding to the states in Eq. (2) with $\theta_0, \theta_1 \in \{\pi/2, 3\pi/4, \cdots, 3\pi/2\}$ and tested by the measurement data corresponding to the states with $\theta_0, \theta_1 \in \{5\pi/8, 7\pi/8, \cdots, 11\pi/8\}$. The states are generated by qutip [54] and the affine transformations are simulated by tools provided in Keras [55].
Figure 8. Rejection rates as functions of quantum fidelity. Solid lines are for four-component cat states and dashed lines are for two-component cat states. Three different scenarios are considered: 50%, 62.5% and 75% of 32 × 32 pixels are randomly selected as input to the neural network respectively. For all the scenarios, each pixel is an estimate of the Wigner function obtained by sampling the true outcome probability distribution 300 times.

**Supplementary Note 3  Performance vs State Complexity**

The prediction performance of StateNet depends on the complexity of quantum states, which can be quantified by their nonclassicality. Here we compare the performances of StateNet with respect to two-component cat states $|\Psi_\alpha\rangle_2 \propto |\alpha\rangle + |\alpha\rangle$ and four-component cat states $|\Psi_\alpha\rangle_4 \propto |\alpha\rangle + |\alpha\rangle + |\alpha\rangle + |\alpha\rangle$, where $\alpha \in [1, 2)$. $|\Psi_\alpha\rangle_4$ has higher nonclassicality than $|\Psi_\alpha\rangle_2$, implying that $|\Psi_\alpha\rangle_4$ has a more complex Wigner function on the phase space. Supplementary Fig. 8 compares the rejection rates in comparing pairs of $|\Psi_\alpha\rangle_4$ and pairs of $|\Psi_\alpha\rangle_2$. Given same amount of measurement data, StateNet has higher success rates in cross-verification of pairs of $|\Psi_\alpha\rangle_2$ than pairs of $|\Psi_\alpha\rangle_4$, indicating that less complex quantum states yields better prediction performance.

**Supplementary Note 4  More on Cross-Platform Verification of Quantum Dynamics**

Supplementary Fig. 9 compares the data images obtained from noiseless state and noisy state at each time step. It is shown that the initial states are exactly same and the data images corresponding to the noisy state become blurrier due to photon error loss.
Figure 9. The figure shows a typical example of a series of data images obtained at $t = 0, 0.1, \ldots, 1$ from noiseless dynamical state versus from noisy dynamical state with photon loss error.