Learning quantum dynamics with latent neural ordinary differential equations

Matthew Choi [1,*] Daniel Flam-Shepherd,[1,2,*] Thi Ha Kyaw [1,3,†] and Alán Aspuru-Guzik[1,2,3,4,‡]

[1]*Department of Computer Science, University of Toronto, Toronto, Ontario M5S 2E4, Canada*
[2]*Vector Institute for Artificial Intelligence, Toronto, Ontario M5S 1M1, Canada*
[3]*Department of Chemistry, University of Toronto, Toronto, Ontario M5G 1Z8, Canada*
[4]*Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada*

(Received 22 October 2021; accepted 15 March 2022; published 4 April 2022)

The core objective of machine-assisted scientific discovery is to learn physical laws from experimental data without prior knowledge of the systems in question. In the area of quantum physics, making progress towards these goals is significantly more challenging due to the curse of dimensionality as well as the counterintuitive nature of quantum mechanics. Here we present the QNODE, a latent neural ordinary differential equation (ODE) trained on expectation values of closed and open-quantum-systems dynamics. It can learn to generate such measurement data and extrapolate outside of its training region that satisfies the von Neumann and time-local Lindblad master equations for closed and open quantum systems, respectively, in an unsupervised means. Furthermore, the QNODE rediscovers quantum-mechanical laws such as the Heisenberg's uncertainty principle in a data-driven way, without any constraint or guidance. Additionally, we show that trajectories that are generated from the QNODE that are close in its latent space have similar quantum dynamics while preserving the physics of the training system.

DOI: 10.1103/PhysRevA.105.042403

I. INTRODUCTION

Deep learning and neural networks have recently become the powerhouse in machine learning (ML) and have successfully been used to tackle complex problems in classical [1–3] and quantum mechanics [4–7] (see Refs. [8–12] for reviews). Machine-assisted scientific discovery is still in its infancy, but progress has been made, mostly by building the correct inductive bias, or structure, into the model or loss function. For example, physical conservation laws can be learned [1,2]. Other work has made progress, for example, in a purely data-driven approach learning relationships between quantum experiments and entanglement using generative models [13]. Recently, neural ordinary differential equations (ODEs) were introduced [14,15], a neural network layer defined by differential equations. Neural ODEs provide the perfect model for physics, since many physical laws are governed by ODEs and thus every neural ODE has the correct inductive bias built into the model itself.

Quantum computing is another prominent area of research currently, with the potential to outperform the capabilities of the best classical computers [16–18]. To make advances in this so-called noisy intermediate-scale quantum (NISQ) era [19], where a quantum computer can possess hundreds of qubits, several NISQ algorithms [20–23] have been proposed. One of the fundamental challenges for NISQ devices to scale-up is to understand noises involved in devices. However, accounting for these undesired environmental effects would require exponential classical compute resources [24]—one solution attempt is to employ quantum computer-aided design [25,26], which still cannot fully account for the environmental effects due to limited NISQ hardware availability.

In general, the study of open quantum systems is important for quantum computing as well as many other areas of physics, from many-body phenomenon [27,28] and light-matter interaction [29–31] to nonequilibrium physics [32,33].

Here we demonstrate that latent ODEs can be trained to generate and extrapolate measurement data from dynamical quantum evolution in both closed and open quantum systems using only physical observations without specifying the physics *a priori*. This is in line with treating the quantum system as a black box and the "shut up and calculate" philosophy [34], all the while ignoring ontological interpretation [35] of quantum physics. The QNODE can predict and extrapolate quantum trajectories much longer than trained on without needing to solve the underlying Schrödinger equation of motion or the time-local Lindblad master equation [31,36] (setting $\hbar = 1$ onwards, see Appendix C for detailed derivation): $d\hat{\rho}_S(t)/dt = -i[\hat{H}, \hat{\rho}_S(t)] + \sum_\nu \Gamma_\nu [\mathbf{A}_\nu \hat{\rho}_S(t) \mathbf{A}_\nu^\dagger - \frac{1}{2}\{\mathbf{A}_\nu^\dagger \mathbf{A}_\nu, \hat{\rho}_S(t)\}]$. Here, $\Gamma_\nu \geqslant 0$ are decay rates, $\mathbf{A}_\nu$ are superoperators, depending on the physical noise model considered, and $\{a, b\} = ab + ba$ is an anticommutator, while $\hat{H}$ is the quantum system Hamiltonian and $\hat{\rho}_S$ is the system density matrix. By setting $\Gamma_\nu = 0$, we arrive at the von Neumann equation for a closed quantum system. There are some attempts to learn open quantum systems [37,38] using recurrent neural networks (RNNs), but these models have inaccurate long-term predictions and poor extrapolation [39]. In addition, there exists similar

*These authors contributed equally to this work.
†thihakyaw@cs.toronto.edu
‡alan@aspuru.com

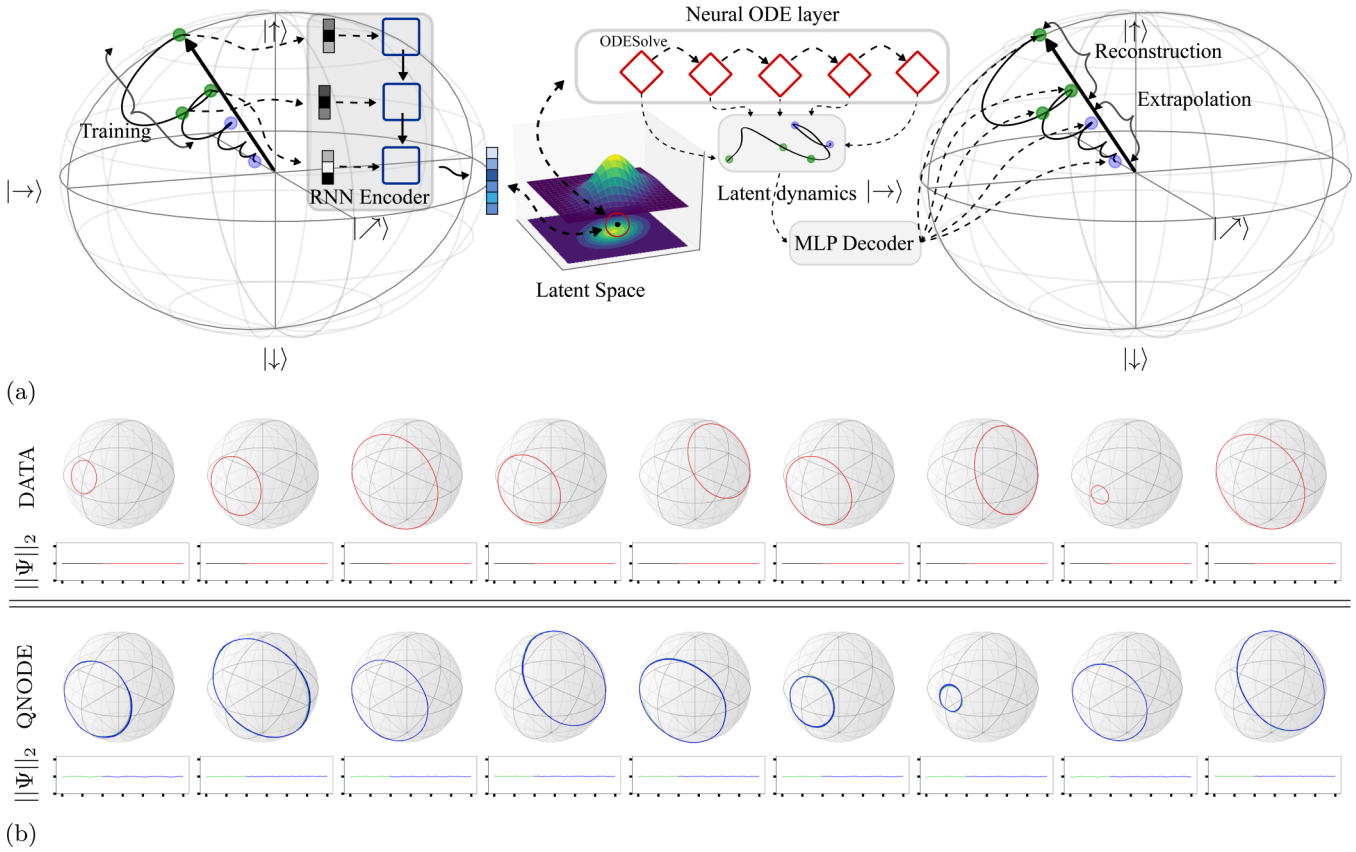


FIG. 1. (a) The model. Visualizing QNODE's components. We first encode the dynamics of a qubit in the form of a time series of its observables. In the case of a simple two-level quantum system, we are looking at the Bloch vector evolution, which is mapped via a recurrent neural network to a latent representation of the dynamics. Then a neural ODE layer outputs a latent representation of subsequent dynamics, which are mapped to quantum dynamics by an MLP decoder. QNODE can reconstruct the training dynamics and extrapolate the dynamics forward in time. (b) Closed-quantum-system dynamics. A table comparing examples of closed-system quantum dynamics from the training DATA (top row) with examples of closed-system quantum dynamics generated by the QNODE (bottom row). Beneath each Bloch sphere is the time-series plot of the norm $||\Psi||_2$ of the dynamics. The green solid line is two arbitrary time units (arb. units, i.e., when $\hbar = \omega = 1$) of trained dynamics and blue is four arb. units of extrapolated dynamics. Black and red are the real quantum dynamics, with black being the actual training region or two arb. units and red being four arb. units of dynamics unseen to the QNODE. Furthermore, we set the limits of the y axis for $||\Psi||_2$ from 0.5 to 1.5.

explorations of studying open quantum systems using kernel methods [40], convolutional neural networks [41,42], and recurrent neural nets [43]. However, the difference between these and ours is that our QNODE is unsupervised learning while all the above references are based on supervised learning methods.

II. PRELIMINARIES

A. QNODE model

We repurpose the latent variable neural ordinary differential equation (ODE) model [14], which is a generative latent function time-series model. The model is trained as a variational autoencoder [44,45] on the quantum dynamics time series, and in this context we refer to the model as the quantum dynamics latent neural ODE (QNODE). We visualize each component of the model in Fig. 1. The encoder is a RNN, which takes in the training quantum dynamics $\mathbf{x}^{\mathcal{O}}(t_1), \ldots, \mathbf{x}^{\mathcal{O}}(t_n)$ sequentially backwards in time and outputs the parameters of the distribution on the latent representation $\mathbf{h}^{\mathcal{O}}(t_0)$ for some initial t_0. Using a standard Gaussian prior on the latents, the dynamics $\mathbf{x}^{\mathcal{O}}(t)$ are modeled with a Gaussian likelihood.

The subsequent learned latent representation $\mathbf{h}^{\mathcal{O}}(t)$ of the quantum dynamics comes from a neural ODE layer [14] which parameterizes the continuous dynamics of the system using an ordinary differential equation specified by a multilayer perceptron (MLP) with parameters θ:

$$\frac{d\mathbf{h}^{\mathcal{O}}(t)}{dt} = \mathrm{MLP}_\theta(\mathbf{h}^{\mathcal{O}}(t), t). \tag{1}$$

Starting from the latent space with $\mathbf{h}^{\mathcal{O}}(t_0)$, we can obtain the subsequent latent dynamics $\mathbf{h}^{\mathcal{O}}(t)$ from the output of this layer, which is the solution to this ODE initial value problem at any time t. A black-box differential equation solver computes this:

$$\mathbf{h}^{\mathcal{O}}(t_{1:N}) = \mathrm{ODESolve}[\mathbf{h}^{\mathcal{O}}(t_0), \mathrm{MLP}_\theta, t_{0:N}], \tag{2}$$

where $\mathbf{h}^{\mathcal{O}}(t_{1:N}) = \mathbf{h}^{\mathcal{O}}(t_1), \ldots, \mathbf{h}^{\mathcal{O}}(t_N)$ is the latent representation of the quantum dynamics $\mathbf{x}^{\mathcal{O}}(t_1), \ldots, \mathbf{x}^{\mathcal{O}}(t_N)$. A single MLP is used to decode the latent dynamics into quantum

TABLE I. Training Parameters. We provide the number of hidden units for the RNN, ODE layer and MLP for all 3 models and their respective hyperparameters.

Model	RNN HU	ODE HU	MLP HU	Learning Rate	Epochs	Total MSE	Average MSE
Closed	48	48	48	4×10^{-3}	7500	2.765×10^{-3}	2.560×10^{-6}
Open	53	53	53	7×10^{-3}	7500	9.964×10^{-4}	9.226×10^{-7}
Two Qubit	170	170	170	2×10^{-3}	7200	2.375×10^{-2}	2.199×10^{-6}

dynamics. For any latent state, the entire latent trajectory is uniquely defined and we can extrapolate this latent trajectory to make predictions arbitrarily far forward in time. The continuously defined dynamics provided by the neural ODE allow us to work over arbitrary times and avoid any discretization of the time intervals. This is akin to treating $\mathbf{h}^{\mathcal{O}}(t)$ as $\hat{\rho}_S(t)$ in the quantum dynamics. The model parameters can be found in Table I.

B. Data generation

To generate time-series data for a quantum system, we first consider a simple two-level quantum system $\hat{H}_1 = (\omega\hat{\sigma}_z + \Delta\hat{\sigma}_x)/2$, where ω is the energy splitting of the two-level system, Δ is the detuning, and $\hat{\sigma}$'s are the usual Pauli matrices. ω, Δ are sampled from a Gaussian distribution (see SM A 1). For each set of $\{\omega, \Delta\}$, we obtain time-series data of the three expectation values that define the Bloch vector, $\langle\hat{\sigma}_x(t)\rangle, \langle\hat{\sigma}_y(t)\rangle, \langle\hat{\sigma}_z(t)\rangle$, for $t \in [0, t_N]$, by numerically solving the corresponding von Neumann equation or Lindblad master equation using QUTIP's [46] numerical solver, given an arbitrary initial quantum state, i.e., $|\psi(0)\rangle = \hat{\mathcal{U}}_R|0\rangle$—we produce two datasets, one each for the open and closed systems. Here $\hat{\mathcal{U}}_R$ refers to the 2×2 Haar random unitary matrix, and t_N is the total training time. In this way one can take the time-series data as a series of projective measurements made across various times $[0, t_N]$ on the statistical ensemble, which is prepared in an initial quantum state $\hat{\rho}_S(0)$.

Similarly, for a two-qubit system we use the following Hamiltonian: $\hat{H}_2 = (\omega_1\hat{\sigma}_z^1 + \Delta_1\hat{\sigma}_x^1)/2 + (\omega_2\hat{\sigma}_z^2 + \Delta_2\hat{\sigma}_x^2)/2 + J\sigma_x^1\sigma_x^2$, where ω's, Δ's, J's are all sampled from the same Gaussian distribution as in the single-qubit case. The initial states are sampled from the Haar random matrix for each qubit space and tensor product afterward.

III. MAIN RESULTS

We conduct three main experiments using the QNODE trained on both closed and open two-level quantum systems—(1) Generating quantum dynamics from random positions in the QNODE's latent space, (2) testing if the QNODE's generated dynamics preserve the Heisenberg uncertainty principle (HUP), and (3) testing if the QNODE learns an interpretable latent space by performing interpolations and assessing its learned physics.

In the following, for each trajectory generated by QNODE the green is two arbitrary time units (arb. units, i.e., when $\hbar = \omega = 1$) of trained dynamics and blue is four arb. units of extrapolated dynamics. Black and red lines are the real quantum dynamics, with black being the actual training region or two arb. units and red being four arb. units of dynamics unseen to the QNODE.

A. Generated dynamics

After training the QNODE on trajectories from both closed- and open-quantum-system dynamics for the single-qubit Hamiltonian $\hat{H}_1$, we test if the model will produce dynamics that resemble the training data and satisfy the von Neumann (time-local Lindblad master) equation for the closed (open) quantum system case. To generate quantum dynamics from the QNODE, we initialize the dynamics randomly in the latent space by sampling $\mathbf{h}^{\mathcal{O}}(t_0)$ from a standard Gaussian distribution. Using that random latent point, we generate quantum dynamics and extrapolate forwards in time for an additional four arb. units. We plot the dynamics generated by the QNODE on the Bloch sphere and compare them with its training data in Fig. 1(b) for the QNODE trained on the closed system and Fig. 2(a) for the QNODE trained on the open system. Inspecting the dynamics in both, it is apparent that the QNODE can generate similar dynamics to the training data. Importantly, the generated dynamics obeys quantum mechanics, since in Fig. 1(b) $||\Psi(t)||_2 \approx 1$ for the training and extrapolation time while in Fig. 2(b) $||\Psi(t)||_2$ decays towards 0, as expected in an open quantum system, where we have bit-flip and dephasing noises, i.e., $\mathbf{A}_1 = \hat{\sigma}_-$ and $\mathbf{A}_2 = \hat{\sigma}_z$, with sets of parameters for Γ, which is sampled from a Gaussian distribution (see SM A 1). Here, $||\mathcal{C}||_2$ is the l_2 norm of the vector $\mathcal{C}$.

Quantitatively, we find that the QNODE trains and performs very well for both closed and open quantum trajectories. The average mean-squared error (MSE) between the exact and reconstructed ones are shown in Figs. 5 and 6. They are in the order of 10^{-2} for the worst cases. Furthermore, the latent trajectories generated for the closed and open systems as shown in Fig. 7 match the trajectories we expect from their respective systems. This signifies that the neural ODE layer has indeed learned quantum dynamics and shows that the MLP layer mainly projects these latent trajectories back onto the Bloch Sphere.

B. Heisenberg uncertainty principle

Next we test if the QNODE is capable of learning the Heisenberg uncertainty principle. For this purpose we generate 50 trajectories from the QNODE exactly as before and compute the variance in $\hat{\sigma}_x$ and $\hat{\sigma}_z$ denoted as var(x) and var(z) from 0 to 6 arb. units. We plot the variances over time in Figs. 2(b) and 2(c) for the closed and open quantum system, respectively, with the same coloring as before (green for training and blue for extrapolation). We see that over time, in-

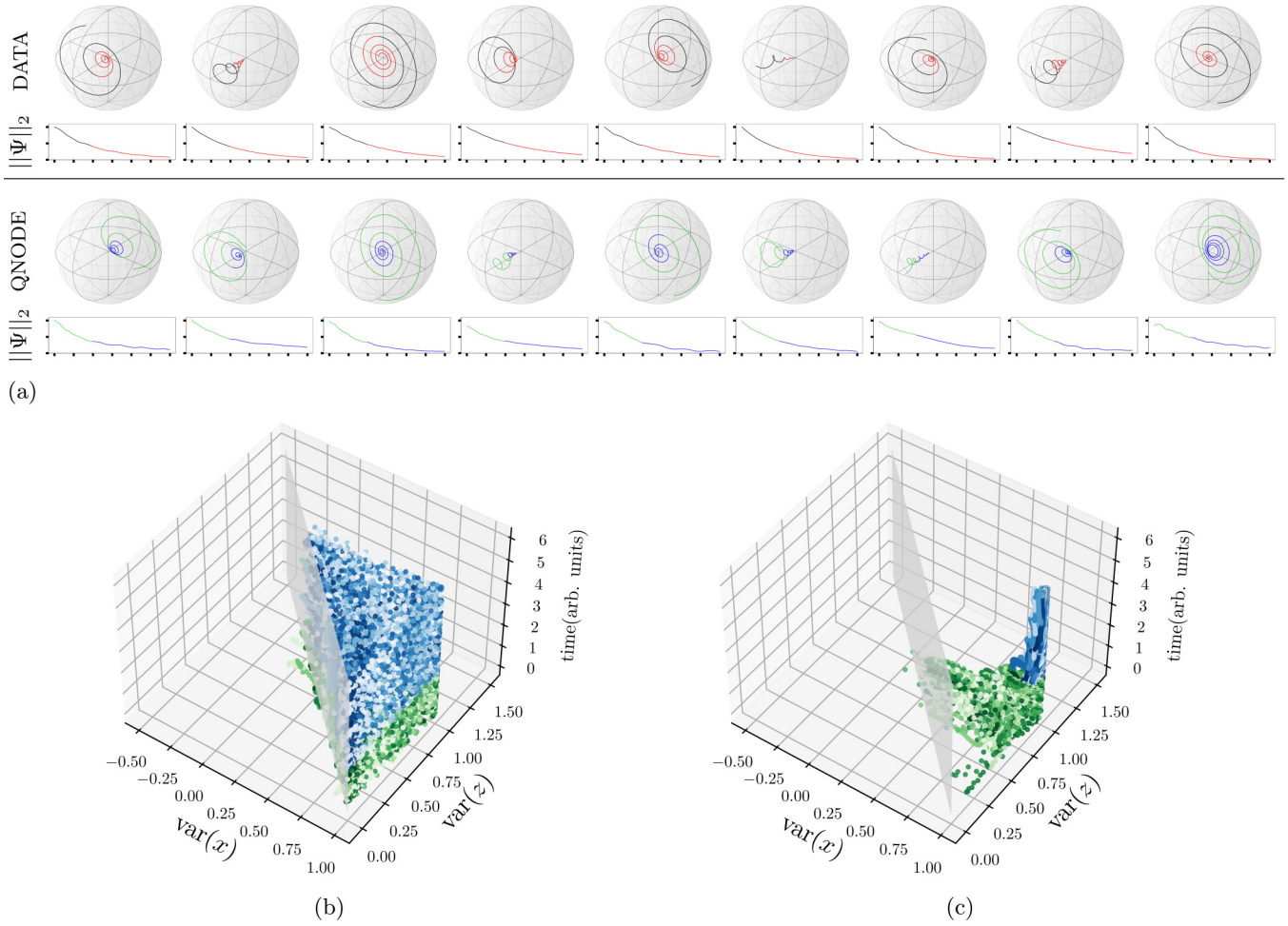


FIG. 2. Open system dynamics. (a) A table comparing examples of open-system quantum dynamics that the QNODE is trained on (top row) with examples of closed-system quantum dynamics generated by the QNODE (bottom row). Beneath each Bloch sphere is the time-series plot of the norm $||\Psi||_2$ of the dynamics. The green solid line is two arb. units of trained dynamics and blue is four arb. units of extrapolated dynamics. Black and red lines are the real quantum dynamics, with black being the actual training region or two arb. units and red being four arb. units of dynamics unseen to the QNODE. Furthermore, we set the limits of the y axis for $||\Psi||_2$ from 0.0 to 1.0. Discovering the uncertainty principle. [(b), (c)] The variance plotted through time of generated quantum dynamics produced by the QNODE as well as plane-separating dynamics that satisfy the uncertainty principle from those that do not.

cluding during which the model is extrapolating, the variances are almost entirely bounded by the plane $\mathrm{var}(z) + \mathrm{var}(x) = 1$, signaling the model does learn to produce dynamics that satisfy the uncertainty principle.

C. Latent space interpolation

We then analyze the smoothness of the latent space to see if it learns a representation that is interpretable while satisfying quantum mechanics. Specifically, we test if trajectories that are close in the latent space have similar quantum dynamics and also preserve physics. This can be done by interpolating in the latent space from one latent point $\mathbf{h}_i^{\mathcal{O}}(t)$ to another $\mathbf{h}_j^{\mathcal{O}}(t)$, and decoding the quantum dynamics on the path in the latent space between the two endpoints points $\mathbf{h}_1^{\mathcal{O}}(t_0)$ and $\mathbf{h}_8^{\mathcal{O}}(t_0)$ in Fig. 3. There we plot one latent space interpolation from the QNODE trained on the open system (the top three rows) and one from the closed system (the bottom three rows). We use spherical linear interpolation on the path and decode at six equally spaced steps along it, corresponding to $\mathbf{h}_2^{\mathcal{O}}(t_0), \ldots, \mathbf{h}_7^{\mathcal{O}}(t_0)$ in the figure. For each interpolation, in the three rows, at the top, we first plot the latent dynamics at each interpolation point $\mathbf{h}_i^{\mathcal{O}}(t)$, in the second row we plot the decoded quantum dynamics, and in the third row is the quantum state's norm $||\Psi(t)||_2$. We use the same coloring as before for training time and extrapolation time. To see if the model has learned a notion of similarity between trajectories corresponding to their quantum dynamics, we perform interpolations from very different training quantum dynamics. Both interpolations show a smooth transition between the different dynamics that preserves physics, given the behavior of the time series of $||\Psi(t)||_2$ in each case.

IV. DISCUSSION

In this work we propose the latent neural ODE for quantum dynamics: QNODE, which is capable of learning and generating quantum trajectories of closed and open quantum systems without any prior knowledge of quantum physics.

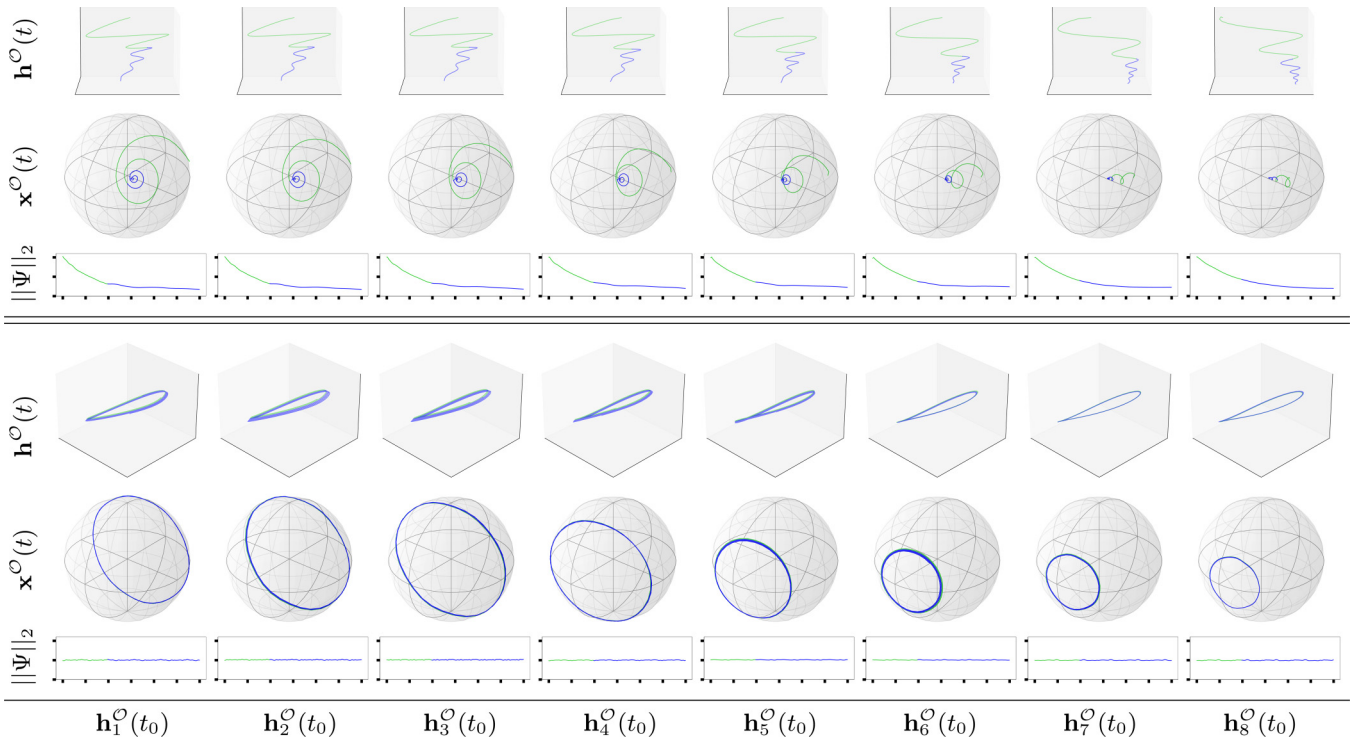


FIG. 3. Latent space interpolations. A table of two interpolations in the latent space of the QNODE with the two different training dynamics: the top super-row is from the open system and bottom super-row is from the closed system. In each super-row there are three subrows showing, for each interpolation point: (1) the plotting latent dynamics, (2) decoded quantum dynamics, and (3) time series of the quantum state's norm. Each column is from a different interpolation point in the latent space. The green solid line is two arb. units of trained dynamics and blue is four arb. units of extrapolated dynamics. Furthermore, we set the limits of the y axis for $||\Psi||_2$ for the open and closed system from 0.5 to 1.5 and 0.0 to 1.0, respectively.

Based on the projective measurement data along the evolving time series, we find that QNODE can learn, reconstruct, and extrapolate the quantum trajectories with high accuracy. Furthermore, the Heisenberg uncertainty principle of a qubit is recovered from the QNODE's generated dynamics. The evidence from our numerical experiments demonstrates that the QNODE is capable of generating quantum dynamics that preserve physical laws, even when they extrapolate outside of the training time.

The QNODE is capable of learning systems beyond the simple two-level quantum system presented so far. The difficulty lies in generating enough data from a larger dimensional

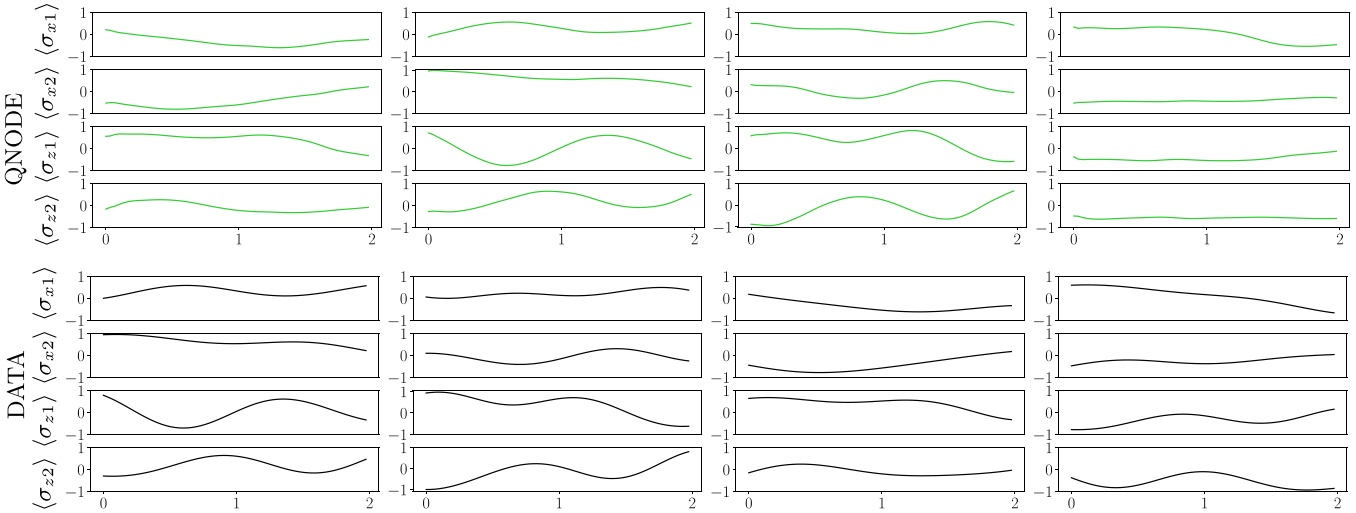


FIG. 4. Two-qubit samples. The green and black trajectories represent expectations values of generated random samples and the training data, respectively. Our model creates random samples similar to the training data, showing that it learns valid dynamics. The x axis traverses from zero to two arb. units.

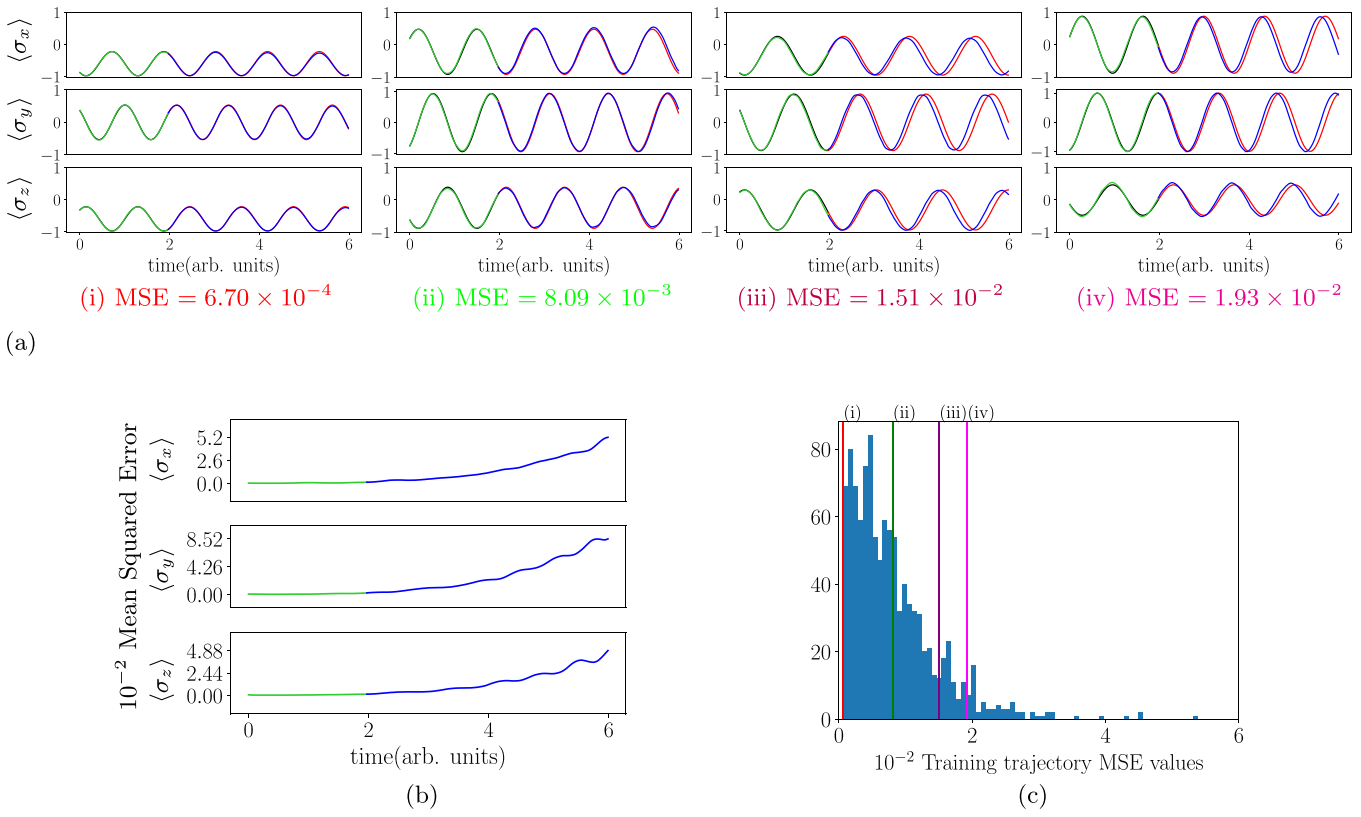


FIG. 5. Closed-system reconstructions: (a) Reconstructions with MSEs ranging from best to worst performing (left to right). (b) The average MSE over time. The plot is bounded by the maximum MSE over six arb. units. (c) A histogram of all MSE values over training data. The colored lines refer to the MSE of the reconstructions of the same color plotted in (a).

Hilbert space. To demonstrate such a point, we have carried out two-qubit numerical experiments with the system Hamiltonian $\hat{H}_2$. The results are shown in Fig. 4, where we show that samples generated by the QNODE are valid and physical. The quantitative analysis regarding the two-qubit case can be seen in Fig. 8. Potentially, multiqubit data could be gathered from recent superconducting qubit experiments [16–18] with randomized quantum gates to produce some physical observables in time. We believe that the QNODE is one step closer to machine-assisted discovery of scientific principles such as quantum phenomena modeled by many interacting classical worlds [47], or could be used to implement dynamical decoupling schemes [48] on the fly during quantum dynamical evolutions as a kind of inverse design [49], thereby paving the way towards large-scale NISQ devices.

The main code used to obtain data presented here can be found in a public repository [50].

ACKNOWLEDGMENTS

We acknowledge fruitful discussions with Areeya Chantasri, Howard Wiseman, and Rodrigo Vargas. T.H.K. and A.A.-G. acknowledge funding from Dr. Anders G. Frøseth. A.A.-G. also acknowledges support from the Canada 150 Research Chairs Program, the Canada Industrial Research Chair Program, and Google, Inc., in the form of a Google Focused Award. Computations and ML training of open, closed, and two-qubit quantum systems were performed on the Niagara supercomputer at the SciNet HPC Consortium [51,52]. SciNet is funded by the Canada Foundation for Innovation, the Government of Ontario, Ontario Research Fund–Research Excellence, and the University of Toronto.

APPENDIX A: METHODS

1. Creating the Training Data

As mentioned in the manuscript, both datasets come from a similar type of Hamiltonian, more specifically, $\hat{H} = (\omega\hat{\sigma}_z + \Delta\hat{\sigma}_x)/2$, where ω is energy splitting of the two-level system, Δ is the detuning, and $\hat{\sigma}$'s are the usual Pauli matrices. ω and Δ are sampled from a Gaussian distribution with a range of 1.5–2.5. Additionally, for the open system, Γ is sampled from a Gaussian distribution with a range of 0.1–0.3. We create 30 of these Hamiltonians and evolve them with 36 sampled initial states from the surface of the Bloch sphere for 60 time steps from 0 to 2 arb. units. We then take the expectation values of the evolution with respect to $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$. This creates our dataset with the size of $(30 * 36, 60, 3)$.

2. Training the Models

The results in the paper are the best models based on the lowest evidence lower bound objective (ELBO) and mean-squared error (MSE). The Table I highlights the hyper-

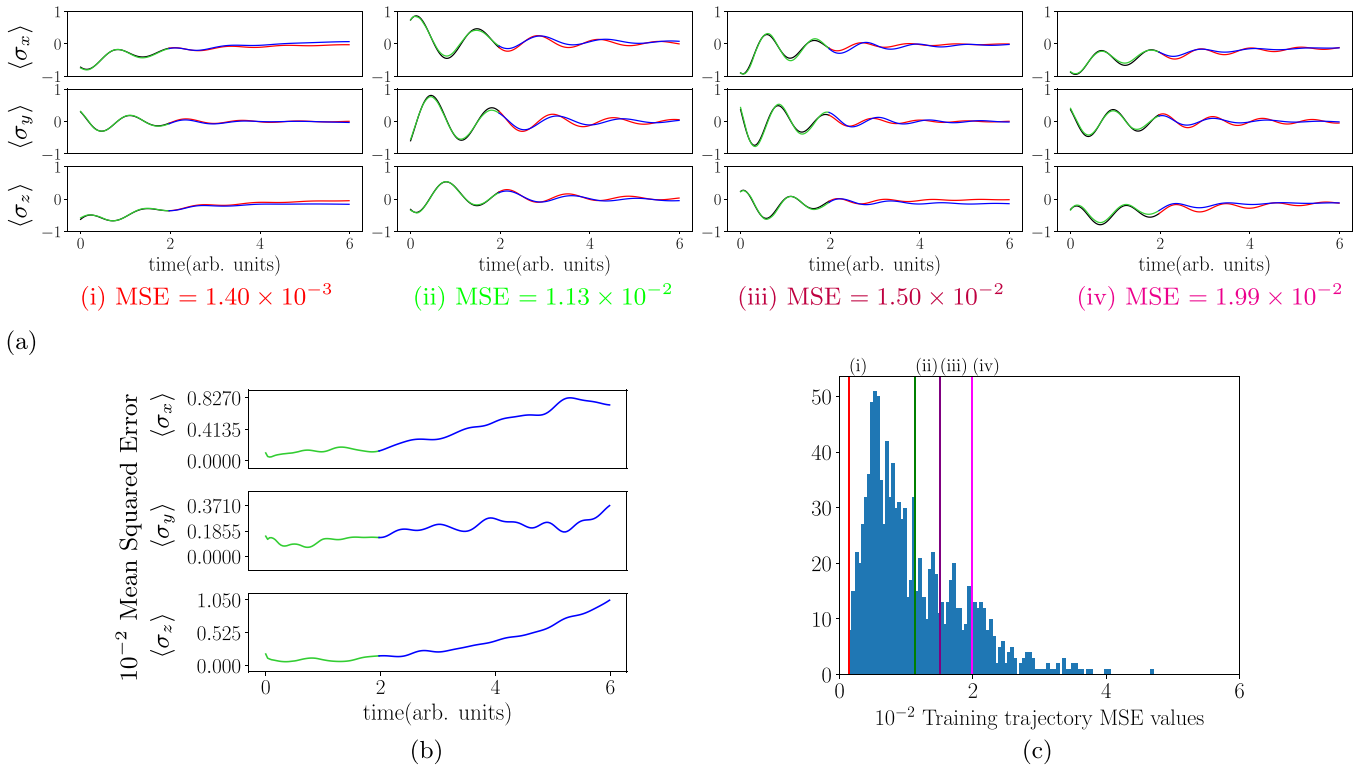


FIG. 6. Open-system reconstructions. (a) Reconstructions with MSEs ranging from best to worst performing (left to right). (b) The average MSE over time. The plot is bounded by the maximum MSE over six arb. units. (c) A histogram of all MSE values over training data. The colored lines refer to the MSE of the reconstructions of the same color plotted in (a).

parameters, e.g., RNN hidden units (HU) and metrics for the best closed, open, and two-qubit models trained.

APPENDIX B: QUANTITATIVE ANALYSES

While the focus of unsupervised learning models is to create new data samples shown in Figs. 1, 2, and 4. We add the reconstruction performance of our models for more quantitative analysis in Figs. 5, 6, and 8. Our models reconstruct the training data well with the bulk of mean squared error losses within 4×10^{-2}.

APPENDIX C: MICROSCOPIC DERIVATION OF THE OPEN-QUANTUM-SYSTEM MASTER EQUATION

In this section, to remind our readers who are not familiar with open-quantum-system treatment, we present a microscopic derivation formalism to arrive at a proper time-local Lindblad master equation for a general open quantum system. Suppose we consider a scenario where a quantum system S weakly interacts with a bath environment B.

In the interaction picture, the system evolution can be written as

$$\frac{d\hat{\rho}_S(t)}{dt} = -\int_0^t d\tau \;\; \mathrm{Tr_B}[\hat{\mathrm{H}}_\mathrm{I}(\mathrm{t}), [\hat{\mathrm{H}}_\mathrm{I}(\tau), \hat{\rho}(\tau)]], \tag{C1}$$

where $\hat{\rho}_S$ represents the system density operator, the subscript B represents the bath, and $\hat{\rho}$ is the combined system and bath density operator. In application of the Born approximation, we assume the interaction between the system and the bath is so small that the bath degrees of freedom $\hat{\rho}_B$ are negligibly affected by the system-bath interaction. Thus the total system at time t can be approximated by $\hat{\rho}(t) \approx \hat{\rho}_S(t) \otimes \hat{\rho}_B$. Therefore we arrive at

$$\frac{d\hat{\rho}_S(t)}{dt} = -\int_0^t d\tau \;\; \mathrm{Tr_B}[\hat{\mathrm{H}}_\mathrm{I}(\mathrm{t}), [\hat{\mathrm{H}}_\mathrm{I}(\tau), \hat{\rho}_\mathrm{S}(\tau) \otimes \hat{\rho}_\mathrm{B}]]. \tag{C2}$$

Furthermore, we assume environmental excitations decay over time and cannot be resolved. With the Markovian approximation, we arrive at the Redfield equation:

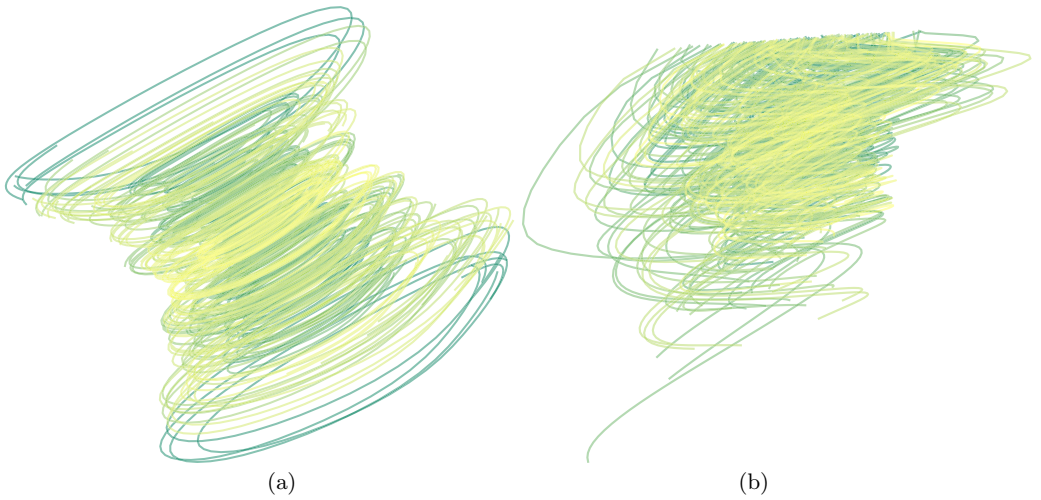
$$\frac{d\hat{\rho}_S(t)}{dt} = -\int_0^t d\tau \;\; \mathrm{Tr_B}[\hat{\mathrm{H}}_\mathrm{I}(\mathrm{t}), [\hat{\mathrm{H}}_\mathrm{I}(\tau), \hat{\rho}_\mathrm{S}(\mathrm{t}) \otimes \hat{\rho}_\mathrm{B}]]. \tag{C3}$$

We then substitute τ by $t - \tau$ and change the upper limit of the integral to ∞. This is allowable provided the integrand vanishes sufficiently fast for $\tau \gg \tau_B$ and the timescale over which the state of the system varies is large compared to the timescale over which the bath correlation functions decay. Thus we arrive at the Markovian quantum master equation,

$$\frac{d\hat{\rho}_S(t)}{dt} = -\int_0^\infty d\tau \;\; \mathrm{Tr_B}[\hat{\mathrm{H}}_\mathrm{I}(\mathrm{t}), [\hat{\mathrm{H}}_\mathrm{I}(\mathrm{t}-\tau), \hat{\rho}_\mathrm{S}(\mathrm{t}) \otimes \hat{\rho}_\mathrm{B}]], \tag{C4}$$

where the time evolution is given by the present state $\hat{\rho}_S(t)$ and is not dependent on the system state in the past. Thus there is no memory effect.

The above procedure is termed the Born-Markov approximation. In general, it does not guarantee Eq. (C4) provides



(a) (b)

FIG. 7. QNODE's latent dynamics of a simple two-level quantum system. Latent trajectories learned by the QNODE for all training data in the closed quantum system (a) and open quantum system (b).

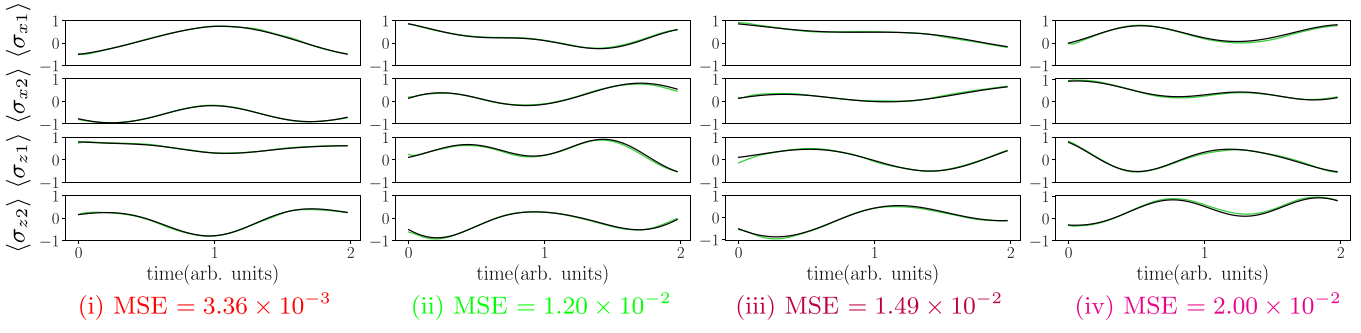
the generator of a dynamical semigroup. Thus, further secular approximation is needed [36]. We proceed by decomposing the interaction Hamiltonian into two parts:

$$\hat{H}_I = \sum_\alpha \hat{A}_\alpha \otimes \hat{B}_\alpha, \tag{C5}$$

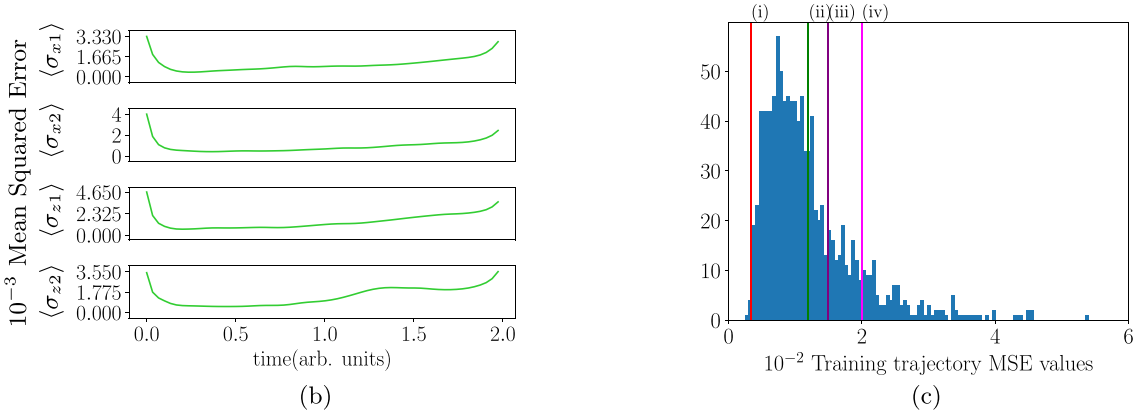
with $\hat{A}_\alpha^\dagger = \hat{A}_\alpha$ and $\hat{B}_\alpha^\dagger = \hat{B}_\alpha$. The secular approximation is achieved if the interaction Hamiltonian is decomposed in terms of the eigenoperators of the system Hamiltonian $\hat{H}_S$. Let us denote the projection onto the eigenspace belonging to the eigenvalue ε in $\hat{H}_S$ as $\hat{\Pi}(\varepsilon)$. Then,

$$\hat{A}_\alpha(\omega) = \sum_{\varepsilon'-\varepsilon=\omega} \hat{\Pi}(\varepsilon)\hat{A}_\alpha\hat{\Pi}(\varepsilon'). \tag{C6}$$

The sum is extended over all energy eigenvalues ε' and ε of H_S with a fixed energy difference ω. As a consequence,



(a)



(b) (c)

FIG. 8. Two-qubit reconstructions. (a) Reconstructions with MSEs ranging from best to worst performing (left to right). (b) The average MSE over time. The plot is bounded by the maximum MSE over two arb. units. (c) A histogram of all MSE values over training data. The colored lines refer to the MSE of the reconstructions of the same color plotted in (a).

we have $[\hat{H}_S, \hat{A}_\alpha(\omega)] = -\omega\hat{A}_\alpha(\omega)$, $[\hat{H}_S, \hat{A}^\dagger_\alpha(\omega)] = +\omega\hat{A}^\dagger_\alpha(\omega)$. The corresponding interaction picture operators take the form

$$e^{i\hat{H}_S t}\hat{A}_\alpha(\omega)e^{-i\hat{H}_S t} = e^{-i\omega t}\hat{A}_\alpha(\omega), \tag{C7}$$

$$e^{i\hat{H}_S t}\hat{A}_\alpha(\omega)e^{-i\hat{H}_S t} = e^{-i\omega t}\hat{A}_\alpha(\omega), \tag{C8}$$

with $[\hat{H}_S, \hat{A}^\dagger_\alpha(\omega)\hat{A}_b eta a(\omega)] = 0$ and $\hat{A}^\dagger_\alpha(\omega) = \hat{A}_\alpha(-\omega)$. We note that $\hat{A}$'s satisfy the completeness relationship: $\sum_\omega \hat{A}_\alpha(\omega) = \sum_\omega \hat{A}^\dagger_\alpha(\omega) = \hat{A}_\alpha$. Eventually, the interaction Hamiltonian in the interaction picture is then

$$\hat{H}_I(t) = \sum_{\alpha,\omega} e^{-i\omega t}\hat{A}_\alpha(\omega)\otimes\hat{B}_\alpha(t) = \sum_{\alpha,\omega} e^{i\omega t}\hat{A}^\dagger_\alpha(\omega)\otimes\hat{B}^\dagger_\alpha(t), \tag{C9}$$

where $\hat{B}_\alpha(t) = e^{i\hat{H}_B t}\hat{B}_\alpha e^{-i\hat{H}_B t}$. By substituting $\hat{H}_I$ back to Eq. (C4) we arrive at

$$\frac{d\hat{\rho}_S(t)}{dt} = \int_0^\infty d\tau \mathrm{Tr}_B[\hat{\mathrm{H}}_\mathrm{I}(\mathrm{t}-\tau)\hat{\rho}_\mathrm{S}(\mathrm{t})\hat{\rho}_\mathrm{B}\hat{\mathrm{H}}_\mathrm{I}(\mathrm{t}) - \hat{H}_I(t)\hat{H}_I(t-\tau)\hat{\rho}_S(t)\hat{\rho}_B] + H.c.$$

$$= \sum_{\omega,\omega'}\sum_{\alpha,\beta} e^{i(\omega'-\omega)t}\Gamma_{\alpha\beta}(\omega)[\hat{A}_\beta(\omega)\hat{\rho}_S(t)\hat{A}^\dagger_\alpha(\omega') - \hat{A}^\dagger_\alpha(\omega')\hat{A}_\beta(\omega)\hat{\rho}_S(t)] + \mathrm{H.c.}, \tag{C10}$$

with a bath correlation function

$$\Gamma_{\alpha\beta}(\omega) = \int_0^\infty d\tau e^{i\omega\tau}\langle\hat{B}^\dagger_\alpha(\tau)\hat{B}_\beta(0)\rangle. \tag{C11}$$

The typical timescale τ_S for which the system evolves is defined as $|\omega' - \omega|^{-1}$, where $\omega' \neq \omega$. By neglecting the rapidly evolving term $\omega' \neq \omega$ during which ρ_S varies appreciably, we have

$$\frac{d\hat{\rho}_S(t)}{dt} = \sum_\omega\sum_{\alpha,\beta}\Gamma_{\alpha\beta}(\omega)(\hat{A}_\beta(\omega)\hat{\rho}_S(t)\hat{A}^\dagger_\alpha(\omega) - \hat{A}^\dagger_\alpha(\omega)\hat{A}_\beta(\omega)\hat{\rho}_S(t)) + \mathrm{H.c.} \tag{C12}$$

Our QNODE model uses quantum data generated from such time-local Lindblad master equations for open-quantum-system dynamics.

[1] S. Greydanus, M. Dzamba, and J. Yosinski, Hamiltonian neural networks, Adv. Neural Inf. Process. Syst. **32**, 15379 (2019).

[2] R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, Discovering Physical Concepts with Neural Networks, Phys. Rev. Lett. **124**, 010508 (2020).

[3] A. D. Tranter, H. J. Slatyer, M. R. Hush, A. C. Leung, J. L. Everett, K. V. Paul, P. Vernaz-Gris, P. K. Lam, B. C. Buchler, and G. T. Campbell, Multiparameter optimisation of a magneto-optical trap using deep learning, Nat. Commun. **9**, 4360 (2018).

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

[5] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature (London) **549**, 195 (2017).

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

[7] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock, and C. Weitenberg, Identifying quantum phase transitions using artificial neural networks on experimental data, Nat. Phys. **15**, 917 (2019).

[8] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. **91**, 045002 (2019).

[9] S. D. Sarma, D.-L. Deng, and L.-M. Duan, Machine learning meets quantum physics, Phys. Today **72**, 48 (2019).

[10] I. Alhousseini, W. Chemissany, F. Kleit, and A. Nasrallah, Physicist's journeys through the AI world—A topical review. There is no royal road to unsupervised learning, arXiv:1905.01023.

[11] K. Bharti, T. Haug, V. Vedral, and L.-C. Kwek, Machine learning meets quantum foundations: A brief survey, AVS Quantum Sci. **2**, 034101 (2020).

[12] R. Roscher, B. Bohn, M. F. Duarte, and J. Garcke, Explainable machine learning for scientific insights and discoveries, IEEE Access **8**, 42200 (2020).

[13] D. Flam-Shepherd, T. Wu, X. Gu, A. Cervera-Lierta, M. Krenn, and A. Aspuru-Guzik, Learning interpretable representations of entanglement in quantum optics experiments using deep generative models, arXiv:2109.02490.

[14] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, Neural ordinary differential equations, Adv. Neural Inf. Process. Syst. **31**, 6571 (2018).

[15] S. Wiewel, M. Becher, and N. Thuerey, Latent space physics: Towards learning the temporal evolution of fluid flow, Comput. Graphics Forum **38**, 71 (2019).

[16] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[17] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[18] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li *et al.*, Strong Quantum Computational Advantage Using a Superconducting Quantum Processor, Phys. Rev. Lett. **127**, 180501 (2021).

[19] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[20] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum (NISQ) algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[21] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio

et al., Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).

[22] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth *et al.*, The variational quantum eigensolver: A review of methods and best practices, arXiv:2111.05176.

[23] J. Zhang, T. H. Kyaw, S. Filipp, L.-C. Kwek, E. Sjöqvist, and D. Tong, Geometric and holonomic quantum computation, arXiv:2110.03602.

[24] P. Iyer and D. Poulin, A small quantum computer is needed to optimize fault-tolerant protocols, Quantum Sci. Technol. **3**, 030504 (2018).

[25] T. H. Kyaw, T. Menke, S. Sim, N. P. Sawaya, W. D. Oliver, G. G. Guerreschi, and A. Aspuru-Guzik, Quantum Computer-Aided Design: Digital Quantum Simulation of Quantum Processors, Phys. Rev. Appl. **16**, 044042 (2021).

[26] J. Kottmann, M. Krenn, T. H. Kyaw, S. Alperin-Lea, and A. Aspuru-Guzik, Quantum computer-aided design of quantum optics hardware, Quantum Sci. Technol. **6**, 035010 (2021).

[27] T. H. Kyaw, V. M. Bastidas, J. Tangpanitanon, G. Romero, and L.-C. Kwek, Dynamical quantum phase transitions and non-Markovian dynamics, Phys. Rev. A **101**, 012111 (2020).

[28] Y. Ashida, *Quantum many-body physics in open systems: Measurement and strong correlations* (Springer, New York, 2020).

[29] T. H. Kyaw, S. Allende, L.-C. Kwek, and G. Romero, Parity-preserving light-matter system mediates effective two-body interactions, Quantum Sci. Technol. **2**, 025007 (2017).

[30] T. H. Kyaw, D. A. Herrera-Marti, E. Solano, G. Romero, and L.-C. Kwek, Creation of quantum error correcting codes in the ultrastrong coupling regime, Phys. Rev. B **91**, 064503 (2015).

[31] T. H. Kyaw, *Towards a scalable quantum computing platform in the ultrastrong coupling regime* (Springer, New York, 2019).

[32] V. M. Bastidas, T. H. Kyaw, J. Tangpanitanon, G. Romero, L.-C. Kwek, and D. G. Angelakis, Floquet stroboscopic divisibility in non-Markovian dynamics, New J. Phys. **20**, 093004 (2018).

[33] M. Heyl, Dynamical quantum phase transitions: A review, Rep. Prog. Phys. **81**, 054001 (2018).

[34] N. D. Mermin, Could Feynman have said this? Phys. Today **57**, 10 (2004).

[35] D. Bohm and B. J. Hiley, *The Undivided Universe: An Ontological Interpretation of Quantum Theory* (Routledge, Milton Park, Abingdon, Oxfordshire, 1993).

[36] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, England, 2002).

[37] E. Flurin, L. S. Martin, S. Hacohen-Gourgy, and I. Siddiqi, Using a Recurrent Neural Network to Reconstruct Quantum Dynamics of a Superconducting Qubit from Physical Observations, Phys. Rev. X **10**, 011006 (2020).

[38] S. Krastanov, K. Head-Marsden, S. Zhou, S. T. Flammia, L. Jiang, and P. Narang, Unboxing quantum black box models: Learning non-Markovian dynamics, arXiv:2009.03902.

[39] R. T. Chen, B. Amos, and M. Nickel, Learning neural event functions for ordinary differential equations, *International Conference on Learning Representations*, (2021), https://openreview.net/forum?id=kW_zpEmMLdP.

[40] A. Ullah and P. O. Dral, Speeding up quantum dissipative dynamics of open systems with kernel methods, New J. Phys. **23**, 113019 (2021).

[41] D. Wu, Z. Hu, J. Li, and X. Sun, Forecasting nonadiabatic dynamics using hybrid convolutional neural network/long short-term memory network, J. Chem. Phys. **155**, 224104 (2021).

[42] L. E. Herrera Rodriguez and A. A. Kananenka, Convolutional neural networks for long time dissipative quantum dynamics, J. Phys. Chem. Lett. **12**, 2476 (2021).

[43] K. Lin, J. Peng, F. L. Gu, and Z. Lan, Simulation of open quantum dynamics with bootstrap-based long short-term memory recurrent neural network, J. Phys. Chem. Lett. **12**, 10225 (2021).

[44] D. P. Kingma and M. Welling, Auto-encoding variational Bayes, arXiv:1312.6114.

[45] D. J. Rezende, S. Mohamed, and D. Wierstra, in *Proceedings of the 31st International Conference on Machine Learning*, edited by E. P. Xing and T. Jebara, (PMLR, Bejing, China, 2014), Vol. 32, pp. 1278–1286.

[46] J. R. Johansson, P. D. Nation, and F. Nori, Qutip: An open-source Python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **183**, 1760 (2012).

[47] M. J. W. Hall, D.-A. Deckert, and H. M. Wiseman, Quantum Phenomena Modeled by Interactions between Many Classical Worlds, Phys. Rev. X **4**, 041013 (2014).

[48] L. Viola, E. Knill, and S. Lloyd, Dynamical Decoupling of Open Quantum Systems, Phys. Rev. Lett. **82**, 2417 (1999).

[49] R. A. Vargas-Hernández, R. T. Chen, K. A. Jung, and P. Brumer, Inverse design of dissipative quantum steady-states with implicit differentiation, arXiv:2011.12808.

[50] M. Choi, D. Flam-Shepherd, T. H. Kyaw, and A. Aspuru-Guzik, Codes for learning quantum dynamics with latent neural ODEs, https://github.com/aspuru-guzik-group/QNODE (2021).

[51] C. Loken, D. Gruner, L. Groer, R. Peltier, N. Bunn, M. Craig, T. Henriques, J. Dempsey, C.-H. Yu, J. Chen *et al.*, SciNet: Lessons learned from building a power-efficient top-20 system and data centre, J. Phys.: Conf. Ser. **256**, 012026 (2010).

[52] M. Ponce, R. van Zon, S. Northrup, D. Gruner, J. Chen, F. Ertinaz, A. Fedoseev, L. Groer, F. Mao, B. C. Mundim *et al.*, Deploying a top-100 supercomputer for large parallel workloads: The Niagara supercomputer, in *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning)*, PEARC '19 (Association for Computing Machinery, Chicago, 2010), pp. 18.