Letter

Neural quantum propagators for driven-dissipative quantum dynamics

Jiaji Zhang [1], Carlos L. Benavides-Riveros [2,*] and Lipeng Chen [1,†]
[1]Zhejiang Laboratory, Hangzhou 311100, China
[2]Pitaevskii BEC Center, CNR-INO and Dipartimento di Fisica, Università di Trento, I-38123 Trento, Italy

(Received 22 October 2024; accepted 20 December 2024; published 21 January 2025)

Describing the dynamics of strong-laser driven open quantum systems is a very challenging task that requires the solution of highly involved equations of motion. While machine learning techniques are being applied with some success to simulate the time evolution of individual quantum states, their use to approximate time-dependent operators (that can evolve various states) remains largely unexplored. In this work, we develop driven *neural quantum propagators* (NQP), a universal neural network framework that solves driven-dissipative quantum dynamics by approximating propagators rather than wave functions or density matrices. NQP can handle arbitrary initial quantum states, adapt to various external fields, and simulate long-time dynamics, even when trained on far shorter time windows. Furthermore, by appropriately configuring the external fields, our trained NQP can be transferred to systems governed by different Hamiltonians. We demonstrate the effectiveness of our approach by studying the spin-boson and the three-state transition Gamma models.

DOI: 10.1103/PhysRevResearch.7.L012013

Introduction. Thanks to rapid advances in laser technology, researchers can now manipulate quantum pathways by fine-tuning the properties of strong laser fields, which allows for the amplification of specific signals that would otherwise be too weak to detect by using conventional methods [1–3]. These developments in strong-field spectroscopy provide valuable insights into complex molecular systems, significantly improving the signal-to-noise ratio and leading to more precise observations [4–7]. Numerical methods for solving the corresponding equations of motion (EOM) generally fall into two categories: iterative-based approaches, such as Runge-Kutta, and methods based on the time-dependent variational principle, which minimizes the residual of the EOM solution [8–14]. However, despite varying levels of sophistication, these algorithms are often hindered by the high computational cost of iteratively propagating the dynamics.

The rise of machine learning has opened new avenues for theoretical simulations of chemical systems [15–18]. Among these advancements, deep neural networks are now routinely used to approximate force fields in molecular dynamics, neural functionals in density functional theory, or potential energy surfaces in *ab initio* calculations, offering significantly reduced computational costs while maintaining good accuracy [19–27]. Recently, there has been a growing interest in using operator learning, where neural networks act as surrogate maps for the solution operators of partial differential equations, accommodating a wide range of initial and boundary conditions [28,29]. A notable example is the Fourier neural operator (FNO), originally developed to solve the Navier-Stokes equations [30,31]. Although FNO has been successfully applied to various classical continuous systems, its potential to directly solve quantum dynamics equations remains largely unexplored, and to the best of our knowledge, it has only been applied to scattering problems [32,33].

In this Letter, we tackle the challenge of constructing neural operators for quantum dynamics and develop a *neural quantum propagator* (NQP) model for driven-dissipative quantum systems. Originally designed as a solver for Markovian quantum master equations (QME), the NQP has already been applied to model population dynamics and compute various response functions for time-independent (i.e., without external driving forces) Hamiltonians [34]. Here, we aim to overcome this limitation by designing a new architecture for NQP that incorporates external driving fields as additional input. By using the QME as the EOM for the system, our NQP acts as a universal surrogate solver applicable to arbitrary initial states and a wide range of external fields. Unlike conventional FNO, which assumes classical, local interactions, our concept of NQP explicitly focuses on quantum dynamics thus involving nonlocal quantum effects. Notably, due to its rigorous adherence to the composition property of quantum propagators (an aspect rarely explored in the literature of operator learning [35–37]), NQP can predict long-time dynamics, far beyond the training time window. Moreover, we demonstrate that our model can be easily transferred to various systems governed by different Hamiltonians.

The remainder of the paper is organized as follows. First, we present the QME for driven-dissipative quantum systems along with the key technical details of our NQP architecture. We then discuss the training scheme and test the numerical performance of the NQP by using the spin-boson and the

*Contact author: cl.benavidesriveros@unitn.it
†Contact author: chenlp@zhejianglab.com

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

three-state Gamma models. Finally, we offer some conclusions.

Driven-dissipative quantum dynamics. In this work, we are interested in open quantum systems driven by external fields. The Hamiltonian is written as

$$\hat{H}(t) = \hat{H}_0 + \sum_{k=1}^{K} f_k(t)\,\hat{F}_k, \tag{1}$$

where the first term is the system Hamiltonian

$$\hat{H}_0 = \sum_j \varepsilon_j |j\rangle\langle j| + \sum_{j\neq j'} \Delta_{jj'} |j\rangle\langle j'|, \tag{2}$$

with ε_j being the state energies and $\Delta_{jj'}$ the interstate couplings. In the second term, $f_k(t)$ is a scalar function and $\hat{F}_k$, the operator part of the kth external field. The overall system is further interacted with a set of Markovian heat baths and the EOM is described by the QME [38–40]

$$\partial_t \hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H}(t),\ \hat{\rho}(t)] - \sum_j \gamma_j \mathcal{D}_j[\hat{\rho}(t)], \tag{3}$$

where $\hat{\rho}(t)$ is the density operator of the system and

$$\mathcal{D}_j[\hat{\rho}(t)] = (\hat{V}_j^\dagger \hat{V}_j \hat{\rho}(t) + \hat{\rho}(t)\hat{V}_j^\dagger \hat{V}_j - 2\,\hat{V}_j \hat{\rho}(t) \hat{V}_j^\dagger), \tag{4}$$

with γ_j and $\hat{V}_j$ being the coupling strength and the operator function of the jth heat bath, respectively. In the presence of an external field, the time-dependent propagator can be defined through the equation

$$\hat{\rho}(t) = \mathcal{T}_+ \exp\left[\int_{t_0}^{t} \mathrm{d}s\, \mathcal{L}(s)\right]\hat{\rho}(t_0), \tag{5}$$

where $\mathcal{T}_+$ denotes the time-ordered operator and the superoperator $\mathcal{L}(s)$ corresponds to the right-hand side of Eq. (3).

We now define a discrete time grid with a time step δ_t and $t_n = n\delta_t$. The kth external field at t_n is denoted by $f_n^k = f_k(t_n)$. All K external fields from t_m to t_n can thus be represented as a matrix

$$\vec{f}_{n,m} = \begin{bmatrix} f_m^1 & f_m^2 & \cdots & f_m^K \\ f_{m+1}^1 & f_{m+1}^2 & \cdots & f_{m+1}^K \\ \vdots & \vdots & \ddots & \vdots \\ f_n^1 & f_n^2 & \cdots & f_n^K \end{bmatrix}, \tag{6}$$

which is further aligned as a vector and used as the input of the neural network.

Next, we introduce the collective index $x = (j, j')$ and align the density operator as a vector $\hat{\rho}(t_n) \sim \vec{\rho}_n = \{\rho(x_0, t_n), \rho(x_1, t_n), \ldots, \}$, with $\rho(x, t) = \langle j|\hat{\rho}(t)|j'\rangle$. Formally, this vectorization is identical to using the twin-space or Liouville space representation [41]. The propagator defined in Eq. (5) can be recast into the matrix-vector form as

$$\vec{\rho}_n = \boldsymbol{G}[\vec{f}_{n,m}; \vec{\rho}_m]. \tag{7}$$

Here the propagator $\boldsymbol{G}$ can be regarded as a functional that projects the given initial state $\vec{\rho}_m$ and the external field $\vec{f}_{n,m}$ to their corresponding final state $\vec{\rho}_n$ satisfying Eq. (3). As a consequence of the linearity of quantum mechanics, the composition property holds and reads

$$\vec{\rho}_{n_2} = \boldsymbol{G}[\vec{f}_{n_2,n_1}; \boldsymbol{G}[\vec{f}_{n_1,n_0}; \vec{\rho}_{n_0}]]. \tag{8}$$

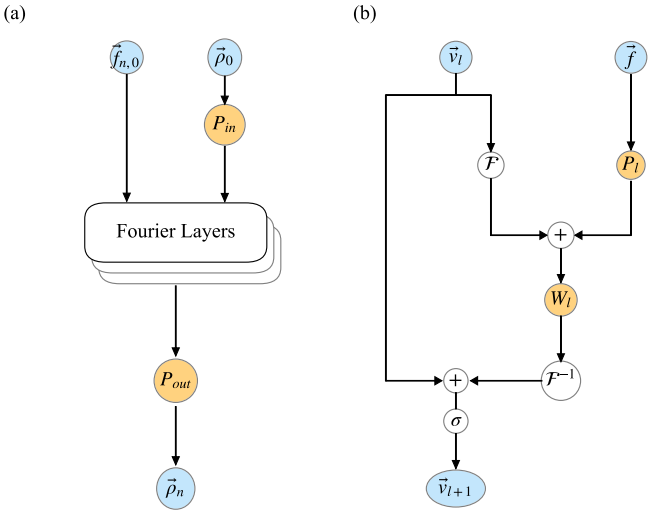


FIG. 1. The architecture of (a) the driven NQP model and (b) the lth Fourier layer, respectively. $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transform and its inverse. $+$ and σ are the element-wise addition and the GeLU activation function.

Although our discussion is limited to the case of Markovian approximation, the same protocol can be easily extended to the non-Markovian case by replacing the QME with other numerically exact methods such as the hierarchical equations of motion [42].

Neural quantum propagators with driven fields. To construct the NQP for driven dynamics, we follow the adaptive FNO transformer [43] and stochastic Fourier differential equation [44] architectures. Our model thus takes an arbitrary initial state $\vec{\rho}_0$ and external field $\vec{f}_{n,0}$ as the input, and outputs the target $\vec{\rho}_n$, following Eq. (7). The upper time limit of the model is chosen as $n \leqslant N_t$ ($N_t = t_{\max}/\delta_t$). For ease of numerical construction, we limit n to be an integer but higher resolution can be obtained by either reducing the time step or adopting the superresolution algorithm. Crucially, because of the composition property of quantum propagators, long-time dynamics outside this time window can be obtained by recursively applying the propagator.

Figure 1 sketches the architecture of our driven NQP model. In Fig. 1(a), P_{in} and P_{out} are linear feed-forward neural networks that serve as the projection between the physical space and the latent Fourier space. As shown in Fig. 1(b), the lth Fourier layer performs the following operation:

$$\vec{v}_{l+1} = \sigma(\vec{v}_l + \mathcal{F}^{-1}[\boldsymbol{W}_l(\mathcal{F}[\vec{v}_l] + \boldsymbol{P}_l \vec{f}_{n,0})]). \tag{9}$$

The model first performs the Fourier transform on v_l, which is the output of the previous layer or P_{in}. It then passes the external field vector $\vec{f}_{n,0}$ to another linear feed-forward neural network P_l. The result of these two routes is then added together and passed to the point-wise convolution W_l, which serves as learnable parameters on the Fourier space. It is then followed by the inverse Fourier transform and the nonlinear Gaussian error linear unit (GeLU) activation function. The leftmost route serves as the residual connection, which is introduced to improve numerical stability. The learnable

parameters are all those contained in P_{in}, P_{out}, W_l, and P_l, respectively.

The key novelty of our model can be further appreciated by the following fact: removing the rightmost route associated with P_l results in the (original) nondriven NQP [34], as sketched in Fig. 1(b). The inclusion of $\vec{f}_{n,0}$ introduces the correction to $\vec{v}_l$ in Fourier space that comes from the effects of the external field. For simplicity, we use a linear layer for P_l and merge with $\mathcal{F}[\vec{v}_l]$ through the element-wise addition. Replacing them with more specific structures, such as convolution, may improve the performance, but will not be discussed in this work.

Physics-informed training algorithm. To train the NQP model, we adopt the physics-informed training algorithm and define the loss function

$$L = \alpha L_{\text{data}} + (1-\alpha)L_{\text{phys}}, \tag{10}$$

where L_{data} and L_{phys} are the data and physics-informed loss functions and α is a hyperparameter that will be dynamically adjusted during the training process.

The first term L_{data} is evaluated from a prepared training dataset, which contains in total N_{data} samples. The pth sample of the dataset is composed of $\vec{f}_{n,0}^{(p)}$, an external field vector, and $\vec{\rho}_n^{(p)}$, the corresponding dynamical states. L_{data} is thus defined as

$$L_{\text{data}} = \frac{1}{N_{\text{data}}N_t}\sum_{p=1}^{N_{\text{data}}}\sum_{n=0}^{N_t}\left|\left|\vec{\mu}_n^{(p)} - \vec{\rho}_n^{(p)}\right|\right|_F, \tag{11}$$

where $||\cdot||_F$ denotes the Frobenius norm, $\vec{\mu}_n^{(p)}$ is the dynamical state for the pth sample predicted by the model via the equation

$$\vec{\mu}_n^{(p)} = \boldsymbol{G}_\theta\left[\vec{f}_{n,0}^{(p)}; \vec{\rho}_0^{(p)}\right], \tag{12}$$

where $\boldsymbol{G}_\theta$ is the aforementioned NQP model (with θ being the entire set of learnable parameters) and $\vec{\rho}_0^{(p)}$ is the initial state for the pth sample. This training dataset is prepared by solving the time evolution using some conventional iterative method.

The second term L_{phys} is defined as the residual of Eq. (3) [45]

$$L_{\text{phys}} = \frac{1}{N_{\text{phys}}N_t}\sum_{p=1}^{N_{\text{phys}}}\sum_{n=0}^{N_t}\left|\left|\partial_t\vec{\mu}_n^{(p)} - \boldsymbol{\mathcal{L}}_n\vec{\mu}_n^{(p)}\right|\right|_F, \tag{13}$$

where $\boldsymbol{\mathcal{L}}_n$ represents the right-hand side (time-dependent) operator of Eq. (3). To evaluate such a cost function, we introduce a physics dataset with N_{phys} samples in total. Interestingly, compared to L_{data}, the evaluation of L_{phys} requires only the initial states. In fact, the explicit expression of each dynamical state $\vec{\mu}_n^{(p)}$ in Eq. (13) can be avoided by using the predicted propagator $\boldsymbol{G}_\theta$ that relates this state with the initial one $\vec{\rho}_0 \sim \hat{\rho}(0)$, as indicated in Eq. (12).

We now present numerical results for the spin-boson system and three-state transition Gamma system. The accuracy of the model is demonstrated by comparing the time evolution of the density operator with that from the fourth-order Runge-Kutta method (RK4).

Numerical experiments. The spin-boson system, relevant for quantum control of pulse reverse engineering and controllable dissipative dynamics [9,46,47], is governed by the Hamiltonian

$$\hat{H}_0 = \frac{\omega_z}{2}(|e\rangle\langle e| - |g\rangle\langle g|) + \omega_x(|e\rangle\langle g| + |g\rangle\langle e|). \tag{14}$$

We assume two heat baths: $\hat{V}_1 = |e\rangle\langle g|$ and $\hat{V}_2 = |g\rangle\langle e|$, describing the absorption and emission processes. For the external field, we assume only one field, with $\hat{F} = |e\rangle\langle e|$, and $f(t) = e^{i\omega_f t}$. In what follows, we choose $\omega_z = \omega_0$ and use it as the unit for all the other parameters. The other parameters are chosen as $\omega_x = 0.5$, $\gamma_1 = 0.1$, and $\gamma_2 = 0.2$, respectively. The range of ω_f is chosen as $\omega_f \in (0.2, 1.0)$. For the time parameters, we let $\delta_t = 0.05$. The time limit $t_{\max}$, associated with time point N_t, plays an important role in the performance of the model and is discussed in the Appendix. We sample the initial states to evaluate L_{phys} (13) from the Gaussian unitary ensemble as explained in more detail in the Appendix.

We choose $t_{\max} = 20$ with $N_t = 400$ and the initial state $\hat{\rho}(0) = |g\rangle\langle g|$. We focus on the time evolution of the density operator up to $t = 80$ by comparing it with the reference result obtained from the RK4 using a time step of 0.05. In Fig. 2, we present the dynamics of populations $p_j(t) = \langle j|\hat{\rho}(t)|j\rangle$ ($j = g, e$) and coherence $\rho_{eg}(t) = \langle e|\hat{\rho}(t)|g\rangle$ at $\omega_f =$ [Fig. 2(a)] 0.2, [Fig. 2(b)] 0.4, [Fig. 2(c)] 0.6, and [Fig. 2(d)] 1.0, respectively, representing the typical slow, intermediate, and fast modulation forces. As shown in Fig. 2, the system exhibits different dynamics within $t \leqslant 20$ at these four typical ω_f cases. Notice that our model cannot only predict the time evolution up to $t_{\max}$ with high accuracy, but also produce the correct long-time coherent dynamics for times far beyond $t_{\max}$.

Universality. One of the main features of the NQP introduced in this Letter is its universality, which renders it applicable to different Hamiltonians. This is established by assuming the constant force $f_k(t) = c_k$, with $c_k \in (c_{k,\min}, c_{k,\max})$. The effective system Hamiltonian now becomes $\hat{H}_c = \hat{H}_0 + \sum_k^K c_k\hat{F}_k$. Our trained model can be easily employed to predict dynamics for all $\hat{H}_c$. This significantly improves the transferability over previously developed models, which, being limited to a specific system, have to be retrained when applied to a different one [34]. To test the performance of our NQP model when applied to different Hamiltonians, we adopt, in the following, the three-state transition Gamma model.

Three-state Gamma system. This model plays an important role in stimulated Raman adiabatic population transfer [48,49]. The system Hamiltonian is defined as

$$\hat{H}_{c_1,c_3} = \sum_{j=1}^{3}\omega_j|j\rangle\langle j| + c_1(|1\rangle\langle 2| + |2\rangle\langle 1|) + c_3(|2\rangle\langle 3| + |2\rangle\langle 3|). \tag{15}$$

We treat the first term as $\hat{H}_0$ and the rest as constant external fields. The state $|2\rangle$ is the transition state between $|1\rangle$ and $|3\rangle$. The transition between different states can be tuned by varying the interstate couplings c_1 and c_3. The system-bath coupling operator is chosen as $\hat{V}_j = |j\rangle\langle j|$ for $j = 1 \sim 3$, respectively. We set the energies of three states as $\omega_1 = 0.0$, $\omega_2 = 0.1$, and $\omega_3 = 1.0$. For the heat baths, we set $\gamma_1 = \gamma_3 = 0.1$, $\gamma_2 = 0.2$,

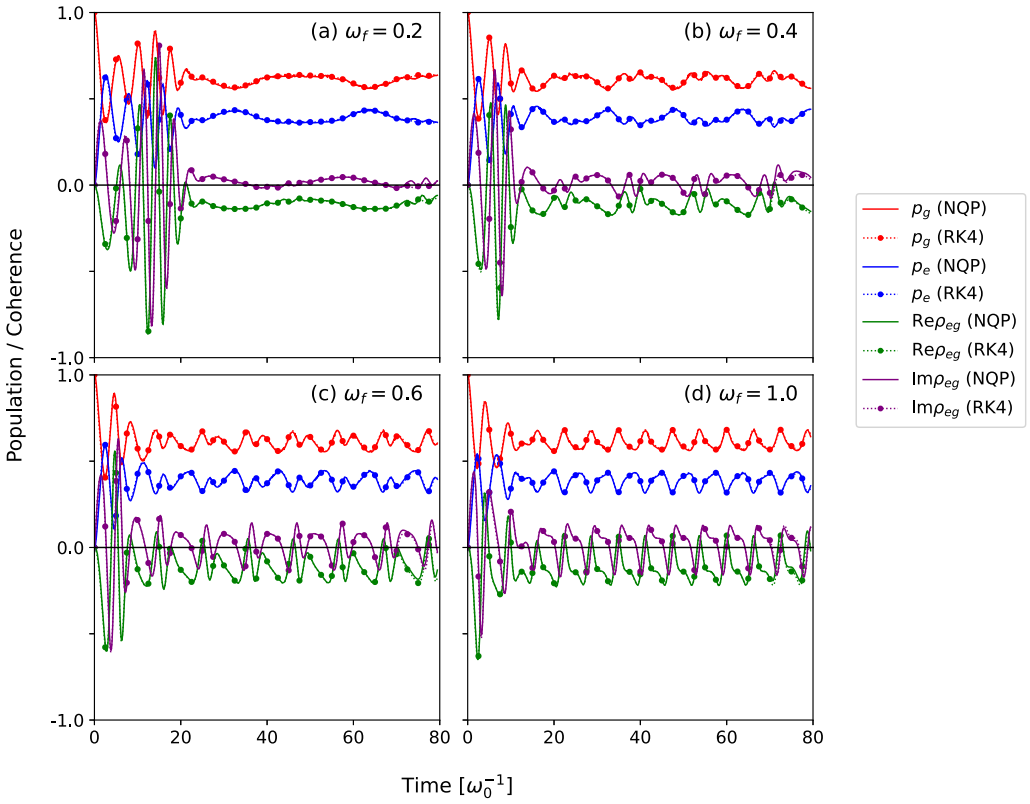


FIG. 2. The population and coherence dynamics of the spin-boson model at $\omega_f =$ (a) 0.2, (b) 0.4, (c) 0.6, and (d) 1.0 as predicted by the NQP model. The red, blue, green, and purple curves represent the populations p_g, p_e and the real and imaginary parts of ρ_{eg}, respectively. The solid and dashed lines represent the results from the NQP model and the reference RK4 (shown with a marker for better illustration), respectively.

assuming a stronger dissipation for the transition state. The range of c_1, c_3 is chosen as $c_1, c_3 \in (0.2,\ \ 0.8)$.

We choose $t_{\max} = 2$, $\delta_t = 0.05$, and the initial state $\hat{\rho}(0) = |1\rangle\langle 1|$. In Fig. 3, we show population dynamics up to $t = 40$ for $c_1 = 0.3$ and $c_3 =$ [Fig. 3(a)] 0.2, [Fig. 3(b)] 0.4, [Fig. 3(c)] 0.6, and [Fig. 3(d)] 0.8, respectively. For all cases, our NQP model predicts accurate dynamics at times far beyond $t_{\max}$, demonstrating the adaptability of the model on different Hamiltonians. In addition, we find that changing $t_{\max}$ in the range of 0.5–10 does not show any apparent differences in the performance of the model. This suggests that different Hamiltonians can be easily handled by our NQP model, even with small training time windows.

Conclusion. In summary, we develop an NQP model—a universal neural network architecture that treats external fields as additional inputs and can handle arbitrary initial quantum states—for simulating driven-dissipative quantum dynamics. It is striking that the trained NQP can predict quite a long-time dynamics for different external fields far beyond the training time window. We also show that, by appropriately configuring the external fields, the model can be transferred to systems described by different Hamiltonians. These results highlight the flexibility and expressibility of our NQP model, rendering it a powerful tool for studying quantum dynamical problems, where an ensemble average has to be taken over a set of initial conditions and external field parameters.

One of the main issues of the NQP model is that the number of learnable parameters scales exponentially with the system size [29]. Similar to other operator-learning frameworks, this scaling problem arises from the fact that the model approximates a functional of inputs with high dimensionality [30]. To reduce computational costs in future applications, one potential solution is to represent the density matrix using tensor network structures, which are known for their polynomial scaling [50–54]. Alternatively, insights can be drawn from exact expressions for steady states [55]. Finally, another promising avenue of research is to reverse the learning protocol by using the learned propagators for quantum control problems, crucial for the development of quantum technologies [56,57].

Acknowledgments. We thank Maxim Gelin for enlightening discussions during the preparation of the manuscript. J.Z. and L.P.C. acknowledge support from the National Natural Science Foundation of China (Grant No. 22473101). C.L.B.-R. gratefully acknowledges financial support from the Royal Society of Chemistry and the European Union's Horizon Europe Research and Innovation program under the Marie Skłodowska-Curie Grant Agreement No. 101065295-RDMFTforbosons. Views and opinions expressed are, however, those of the author only and do not necessarily reflect those of the European Union or the European Research Executive Agency.

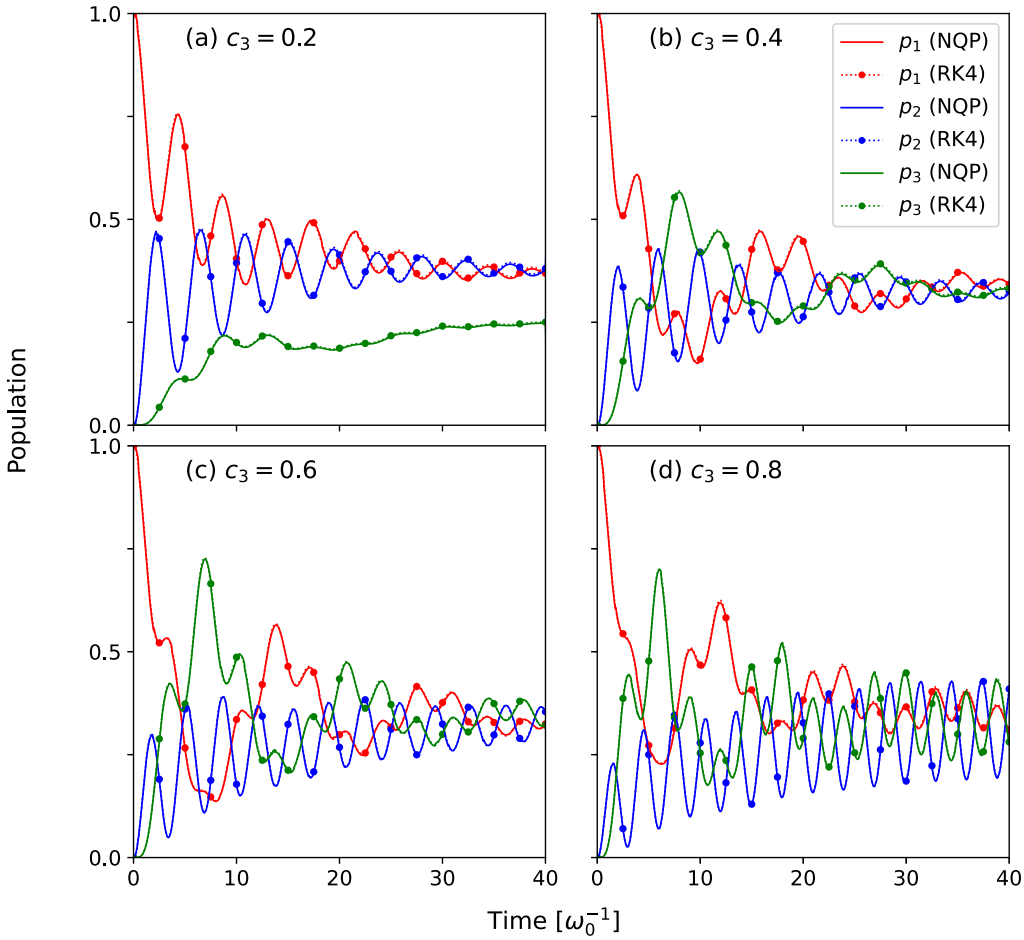


FIG. 3. The population dynamics of the three-state system at $c_1 = 0.3$ and different c_3. In all panels, the solid and dash-dotted lines represent results obtained from our NQP model and the reference RK4, respectively.

Appendix: Preparation of the training set. We briefly introduce the preparation of data and physics set used in our experiments. The initial state $\vec{\rho}_0 \sim \hat{\rho}(0)$ is randomly sampled from the Gaussian unitary ensemble as a hermitian matrix. A sigmoid scaling is conducted for the diagonal entries to have trace-norm

$$\langle j|\hat{\rho}|j\rangle \to \frac{\langle j|\hat{\rho}|j\rangle}{\left(\sum_j \langle j|\hat{\rho}|j\rangle\right)}. \tag{A1}$$

The external field vector $\vec{f}_{n,0}$ is prepared by first selecting a specific function form for $f(t)$, and then randomly choosing the related parameters within chosen intervals. For better understanding, we use periodic forces as $f(t) = \exp(i\omega_f t)$. The random sampling is conducted by first choosing ω_f as a uniform random number within the range $\omega_f \in (\omega_{\min}, \omega_{\max})$. Finally, the vector $\vec{f}_{n,0}$ is obtained by embedding $f(t)$ to the time grid t_n.

Computational details. The hyperparameters and training setup of the NQP model are chosen as follows. P_{in}, P_{out}, and P_l are parameterized as two-layer neural networks with a hidden channel of size 256. For W_l, we use four Fourier layers, each of which has a hidden channel of size 128. All the Fourier modes are explicitly included in the model. The total number of trainable parameters is around two million. We prepare the training dataset by randomly sampling $N_{\text{data}} = 2000$ initial states and external field parameters, and integrating the EOM with the RK4. The physics dataset has a size of $N_{\text{phys}} = 200$, and is regenerated at each epoch. The model is trained by optimizing Eq. (10) (see main text) with $\alpha = 0.5$ using the Adam optimizer and a learning rate of 10^{-4}. The training is conducted for 10^4 epochs on a single Nvidia 4090 GPU card within 5 hours until the loss function reaches $\sim 10^{-4}$. The trained model can then be applied to any initial states and any $\omega_f \in (0.2, 1.0)$.

Performance outside the training window. We note that the performance of the model strongly depends on the choice of the maximum time $t_{\max}$, which should be, in principle, large enough. To test the effect of $t_{\max}$ on the model's performance, we retrain the model using smaller $t_{\max}$. We show the population dynamics $p_g(t)$ up to $t = 40$ evaluated from the NQP model with $t_{\max} = 5$ and 10 in Fig. 4. The other settings are the same as the case of $t_{\max} = 20$ in Fig. 2 (see main text). As shown in Fig. 4, the trained models can always predict the correct dynamics within the training time window $t_{\max}$. For $t_{\max} = 5$, the model incorrectly predict dynamics at $t > t_{\max}$ for all ω_f. We find that employing a larger model by increasing the number of layers to 15 and channels to 512 merely leads to longer training time, but cannot improve long-time dynamics outside the training window. Increasing $t_{\max}$ to 10, the performance of the model improves

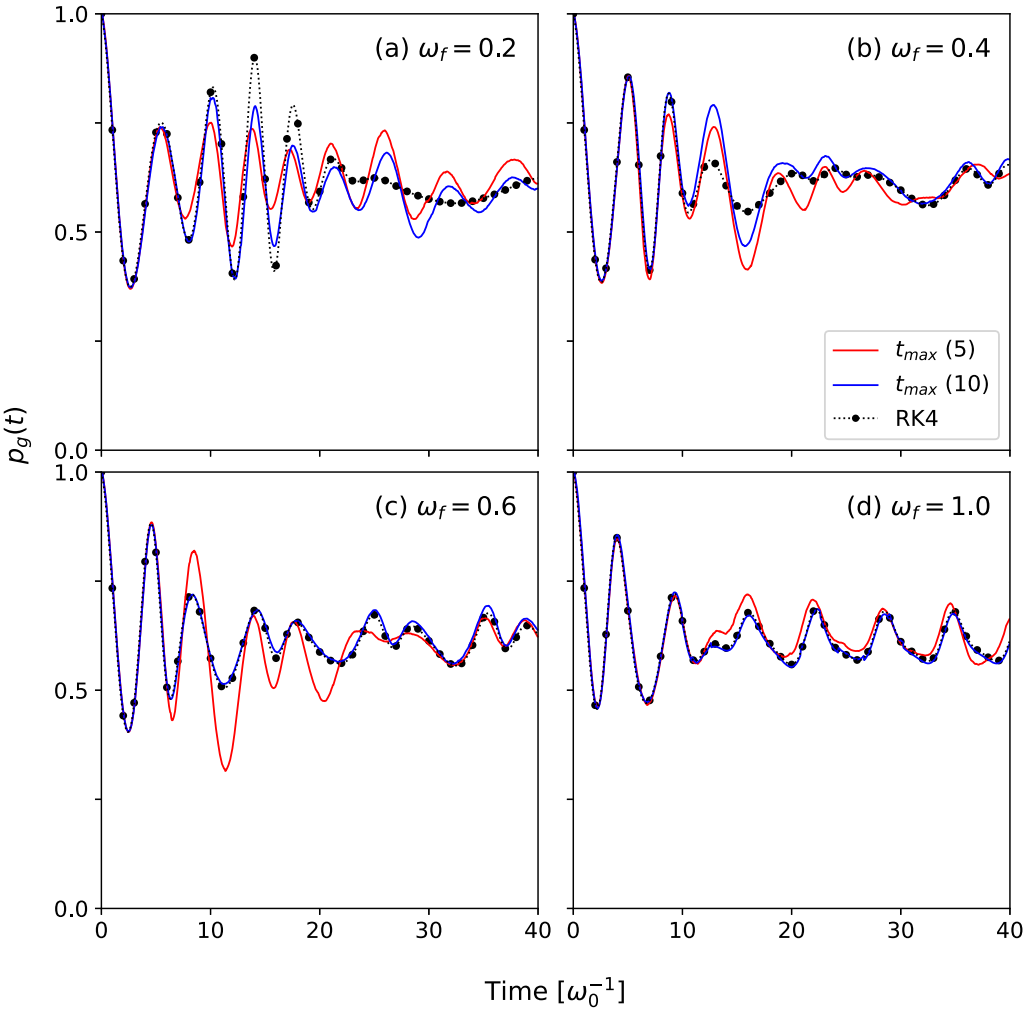


FIG. 4. The population dynamics $p_g(t)$ for different ω_f evaluated from the NQP model with $t_{\max} = 5$ (red) and 10 (blue). The reference RK4 results are shown in black.

with the increase of ω_f, yielding almost correct dynamics for the cases of $\omega_f = 0.6$ and 1.0. It is thus expected that the optimal choice of $t_{\max}$ is determined by the range of ω_f, which corresponds to the modulation of the external fields. This hyperparameter has to be carefully chosen to get better performance.

[1] D. Jonas, Two-dimensional femtosecond spectroscopy, Annu. Rev. Phys. Chem. **54**, 425 (2003).
[2] C. Consani, S. Ruetzel, P. Nuernberger, and T. Brixner, Quantum control spectroscopy of competing reaction pathways in a molecular switch, J. Phys. Chem. A **118**, 11364 (2014).
[3] D. Keefer and S. Mukamel, Selective enhancement of spectroscopic features by quantum optimal control, Phys. Rev. Lett. **126**, 163202 (2021).
[4] M. F. Gelin, D. Egorova, and W. Domcke, Strong and long makes short: Strong-pump strong-probe spectroscopy, J. Phys. Chem. Lett. **2**, 114 (2011).
[5] M. Woerner, A. Ghalgaoui, K. Reimann, and T. Elsaesser, Two-color two-dimensional terahertz spectroscopy: A new approach for exploring even-order nonlinearities in the nonperturbative regime, J. Chem. Phys. **154**, 154203 (2021).
[6] D. Egorova, M. F. Gelin, M. Thoss, H. Wang, and W. Domcke, Effects of intense femtosecond pumping on ultrafast electronic-vibrational dynamics in molecular systems with relaxation, J. Chem. Phys. **129**, 214303 (2008).
[7] S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, New York, 1995).
[8] P. Hoff, S. Thallmair, M. Kowalewski, R. Siemering, and R. Vivie-Riedle, Optimal control theory closing the gap between theory and experiment, Phys. Chem. Chem. Phys. **14**, 14460 (2012).
[9] A. Norambuena, M. Mattheakis, F. J. González, and R. Coto, Physics-informed neural networks for quantum control, Phys. Rev. Lett. **132**, 010801 (2024).
[10] S. Gupta, H. K. Yadalam, M. Kulkarni, and C. Aron, Quantum jumps in driven-dissipative disordered many-body systems, Phys. Rev. A **109**, L050201 (2024).

[11] A. Ullah, Y. Huang, M. Yang, and P. O. Dral, Physics-informed neural networks and beyond: Enforcing physical constraints in quantum dissipative dynamics, Digit. Discov. **3**, 2052 (2024).

[12] M. Reh, M. Schmitt, and M. Gärttner, Time-dependent variational principle for open quantum systems with artificial neural networks, Phys. Rev. Lett. **127**, 230501 (2021).

[13] P. Kramer and M. Saraceno, *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics*, Lecture Notes in Physics (Springer, Berlin, 1981).

[14] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Variational neural-network ansatz for steady states in open quantum systems, Phys. Rev. Lett. **122**, 250503 (2019).

[15] J. A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K.-R. Müller, and A. Tkatchenko, Combining machine learning and computational chemistry for predictive insights into chemical systems, Chem. Rev. **121**, 9816 (2021).

[16] M. Ceriotti, C. Clementi, and O. Anatole von Lilienfeld, Machine learning meets chemical physics, J. Chem. Phys. **154**, 160401 (2021).

[17] H. Lange, A. V. de Walle, A. Abedinnia, and A. Bohrdt, From architectures to applications: A review of neural quantum states, Quantum Sci. Technol. **9**, 040501 (2024).

[18] J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold, Nature (London) **596**, 583 (2021).

[19] F. Noé, A. Tkatchenko, K. Müller, and C. Clementi, Machine learning for molecular simulation, Annu. Rev. Phys. Chem. **71**, 361 (2020).

[20] X. Du, W. Shao, C. Bao, L. Zhang, J. Cheng, and F. Tang, Revealing the molecular structures of α-Al_2O_3(0001)water interface by machine learning based computational vibrational spectroscopy, J. Chem. Phys. **161**, 124702 (2024).

[21] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, Machine learning force fields, Chem. Rev. **121**, 10142 (2021).

[22] V. Martinetto, K. Shah, A. Cangi, and A. Pribram-Jones, Inverting the KohnSham equations with physics-informed machine learning, Mach. Learn.: Sci. Technol. **5**, 015050 (2024).

[23] J. Gedeon, J. Schmidt, M. J. P. Hodgson, J. Wetherell, C. L. Benavides-Riveros, and M. A. L. Marques, Machine learning the derivative discontinuity of density-functional theory, Mach. Learn.: Sci. Technol. **3**, 015011 (2022).

[24] R. Nagai, R. Akashi, and O. Sugino, Completing density functional theory by machine learning hidden messages from molecules, npj Comput. Mater. **6**, 43 (2020).

[25] J. Kirkpatrick *et al.*, Pushing the frontiers of density functionals by solving the fractional electron problem, Science **374**, 1385 (2021).

[26] F. Sammüller, S. Robitschko, S. Hermann, and M. Schmidt, Hyperdensity functional theory of soft matter, Phys. Rev. Lett. **133**, 098201 (2024).

[27] X. Gong, S. G. Louie, W. Duan, and Y. Xu, Generalizing deep learning electronic structure calculation to the plane-wave basis, Nat. Comput. Sci. **4**, 752 (2024).

[28] R. Vinuesa and S. L. Brunton, Enhancing computational fluid dynamics with machine learning, Nat. Comput. Sci. **2**, 358 (2022).

[29] S. L. Brunton and J. N. Kutz, Promising directions of machine learning for partial differential equations, Nat. Comput. Sci. **4**, 483 (2024).

[30] N. Kovachki, S. Lanthaler, and S. Mishra, On universal approximation and error bounds for fourier neural operators, J. Mach. Learn. Res. **22**, 1 (2021).

[31] Z. Li, N. B. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, Fourier neural operator for parametric partial differential equations, arXiv:2010.08895.

[32] S. Mizera, Scattering with neural operators, Phys. Rev. D **108**, L101701 (2023).

[33] V. Niarchos and C. Papageorgakis, Learning *s*-matrix phases with neural operators, Phys. Rev. D **110**, 045020 (2024).

[34] J. Zhang, C. L. Benavides-Riveros, and L. Chen, Artificial-intelligence-based surrogate solution of dissipative quantum dynamics: Physics-informed reconstruction of the universal propagator, J. Phys. Chem. Lett. **15**, 3603 (2024).

[35] N. McGreivy and A. Hakim, Weak baselines and reporting biases lead to overoptimism in machine learning for fluid-related partial differential equations, Nat. Mach. Intell. **6**, 1256 (2024).

[36] B. Alkin, A. Fürst, S. L. Schmid, L. Gruber, M. Holzleitner, and J. Brandstetter, Universal physics transformers: A framework for efficiently scaling neural operators, in *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

[37] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar, Neural operator: Learning maps between function spaces with applications to PDEs, J. Mach. Learn. Res. **24**, 1 (2023).

[38] U. Weiss, *Quantum Dissipative Systems*, 4th ed. (World Scientific, Singapore, 2012).

[39] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, New York, 2007).

[40] F. Campaioli, J. H. Cole, and H. Hapuarachchi, Quantum master equations: tips and tricks for quantum optics, quantum computing, and beyond, PRX Quantum **5**, 020202 (2024).

[41] D. Manzano, A short introduction to the Lindblad master equation, AIP Adv. **10**, 025106 (2020).

[42] J. Zhang and L. Chen, A non-Markovian neural quantum propagator and its application to the simulation of ultrafast nonlinear spectra, Phys. Chem. Chem. Phys. **27**, 182 (2025).

[43] J. Guibas, M. Mardani, Z. Li, A. Tao, A. Anandkumar, and B. Catanzaro, Efficient token mixing for transformers via adaptive Fourier neural operators, in *International Conference on Learning Representations* (2022).

[44] C. Salvi, M. Lemercier, and A. Gerasimovics, Neural stochastic PDEs: Resolution-invariant learning of continuous spatiotemporal dynamics, in *Advances in Neural Information Processing Systems*, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Curran Associates, Red Hook, New York, 2022), Vol. 35, pp. 1333–1344.

[45] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-informed machine learning, Nat. Rev. Phys. **3**, 422 (2021).

[46] D. Ran, W.-J. Shan, Z.-C. Shi, Z.-B. Yang, J. Song, and Y. Xia, Pulse reverse engineering for controlling two-level quantum systems, Phys. Rev. A **101**, 023822 (2020).

[47] W. Wu and Z.-Z. Zhang, Controllable dynamics of a dissipative two-level system, Sci. Rep. **11**, 7188 (2021).

[48] J. R. Kuklinski, U. Gaubatz, F. T. Hioe, and K. Bergmann, Adiabatic population transfer in a three-level system driven by delayed laser pulses, Phys. Rev. A **40**, 6741 (1989).

[49] K. Bergmann, H. Theuer, and B. W. Shore, Coherent population transfer among quantum states of atoms and molecules, Rev. Mod. Phys. **70**, 1003 (1998).

[50] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, Rev. Mod. Phys. **93**, 045003 (2021).

[51] J. Ren, W. Li, T. Jiang, Y. Wang, and Z. Shuai, Time-dependent density matrix renormalization group method for quantum dynamics in complex systems, WIREs Comput. Mol. Sci. **12**, e1614 (2022).

[52] W. Li, J. Ren, and Z. Shuai, Numerical assessment for accuracy and GPU acceleration of TD-DMRG time evolution schemes, J. Chem. Phys. **152**, 024127 (2020).

[53] R. Borrelli and M. F. Gelin, Finite temperature quantum dynamics of complex systems: Integrating thermo-field theories and tensor-train methods, WIREs Comput. Mol. Sci. **11**, e1539 (2021).

[54] A. Novikov, D. Podoprikhin, A. Osokin, and D. P. Vetrov, Tensorizing neural networks, in *Advances in Neural Information Processing Systems*, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (Curran Associates, Red Hook, New York, 2015), Vol. 28.

[55] D. Roberts and A. A. Clerk, Driven-dissipative quantum kerr resonators: New exact solutions, photon blockade and quantum bistability, Phys. Rev. X **10**, 021022 (2020).

[56] C. Brif, R. Chakrabarti, and H. Rabitz, Control of quantum phenomena: Past, present and future, New J. Phys. **12**, 075008 (2010).

[57] A. Castro, A. García Carrizo, S. Roca, D. Zueco, and F. Luis, Optimal control of molecular spin qudits, Phys. Rev. Appl. **17**, 064028 (2022).