Applications of neural networks to the simulation of quantum dynamics of open quantum systems

Sayantan Bandyopadhyay[1], Zhongkai Huang[1], Kewei Sun[2], and Yang Zhao[1*]
[1]*Division of Materials Science, Nanyang Technological University, Singapore 639798, Singapore*
[2]*School of Science, Hangzhou Dianzi University, Hangzhou 310018, China*

(Dated: April 28, 2018)

Despite neural networks' success, their applications to open-system dynamics are few. In this work, non-linear autoregressive neural networks are adopted to generalize time series of expectation values of observables of interest in open quantum systems. Using Dirac-Frenkel time-dependent variation with the multiple Davydov D_2 *Ansatz*, we obtain first stages of dynamical states of both the spin-boson model and the dissipative Landau-Zener model. With calculated data, careful training of the non-linear neural networks is performed. It is shown that training quality of the networks is sufficient to ensure a least mean square error of 1×10^{-11}. Subsequently, the network is cross validated by testing with additional data. Successes of the network training demonstrate that initial data of simulated open-system dynamics contain sufficient knowledge regarding its future propagation. We use the first-stage information and the trained network to predict future values of target observables in the series, and succeed with considerable accuracy.

I. INTRODUCTION

Simulation of quantum dynamics plays an essential role in a variety of fields, such as atomic and molecular physics [1–3], quantum optics [4], solid state physics [5], chemical physics [6], and quantum information science [7]. Since a real system in general cannot be completely isolated from its surroundings, to accurately describe the state of a quantum system, it is necessary to consider the influence of its environment, which has been demonstrated to play a vital role in its dynamical behavior. For instance, the environment may alter the interaction between the two energy levels of a quantum two-state system [10]. Such examples are plentiful in various systems ranging from quantum computers [8] to light-harvesting aggregates [9].

The environment is typically modelled as a bath coupled to the target system. Due to the complexity of system-bath interactions, it is difficult to devise exact analytical solutions for the dynamics of most open quantum systems. A variety of techniques has been developed in the context of open quantum systems, such as the multiconfiguration time-dependent Hartree method [11] and the density matrix renormalization group methodology [12]. In general, methods for the study of open quantum systems can be classified into two types: wavefunction based methods that operate in the Hilbert space [13, 14] and density matrix based approaches in the Liouville space of the system [15, 16]. These methods have proven their validities in fields such as quantum measurement theory [17], quantum statistical mechanics [18], and quantum thermodynamics [19].

*Electronic address: YZhao@ntu.edu.sg

However, methodologies used to investigate dynamics of the open quantum system can differ greatly depending on target systems and desired objectives. For example, approaches used for chemical physics [20] have different approximations from those in quantum optics [21]. But, each method developed for the open quantum system has its limitations. For instance, methods such as quasi-adiabatic propagator path integral (QUAPI) [15] and hierarchical equations of motion [16] are considered exact but computationally expensive. Perturbative methods such as the master equation method are able to provide efficient calculations but can become inaccurate in certain parameter regimes[22].

In recent years, artificial intelligence has attracted considerable attention as a suitable neural network can reduce the computational cost dramatically. Neural networks have been in numerous ways to solve quantum problems. Combined with variational principles it has been used for analysis of the quantum entanglement [23]. It has been shown that deep neural network can efficiently represent most physical states, including the ground states of many-body Hamiltonians [24]. The multiscale entanglement renormalization ansatz along with a deep neural network was used to remove the problem with sampling [25]. Machine learning has also been used along with Bayesian statistics to speed up hybrid Monte Carlo simulations [26]. Recent progress has been made in applying artificial intelligence on dynamics of open quantum system. For example, Cerrillo *et al.* have combined the time evolving density matrix using orthogonal polynomials algorithm with the transfer tensors formalism to the analysis, compression and propagation of non Markovian processes [27–29]. Carleo *et al.* have also demonstrated that a reinforcement-learning scheme is capable of both describing the unitary time evolution of proto-

typical interacting spins models in one and two dimensions [30]. All this provides inspiration for more neural-network investigations into the analysis of the dynamical observables in various open quantum systems.

In this work, the dynamical neural network is employed to simulate open-system quantum dynamics. The first stage of the dynamics is obtained by using the Dirac-Frenkel variational principle with the multiple Davydov D_2 trial state. With data obtained, we train and test suitable nonlinear neural nets to determine the validity of our approach as non-linear autoregressive neural network is the simplest and the most convenient method available to us that can be trained to predict a time series from that series past values efficiently. For example, this method has been used to study the nonlinear behaviours such as chaotic optical intensities in laser [31, 32]. Our approach has been applied, with great success, to the dynamics of the time-independent spin-boson Hamiltonian and the time-dependent Landau-Zener Hamiltonian.

The rest of the paper is structured as follows. In Sec. II, we present the neural network approach and our trial wave function, the multiple Davydov D_2 *Ansatz*. Dynamics of the spin-boson model is investigated in Sec. III A. Finally, dynamics in the dissipative Landau-Zener model is described in Sec. III B. Conclusions are drawn in Sec. IV.

II. METHODOLOGY

This methodology section is divided into two subsections based on the objectives. First part of the section contains a brief introduction of the biologically inspired neural network, while the second part covers the intricacies of the multiple Davydov trial states which was used to model the dynamical observables data. As the primary focus of this work is the neural networks we shall introduce it first.

A. Neural network method

Artificial neural networks are designed to mimic the abilities of their biological counterparts when augmented with appropriate mathematical models. It has shown great promise in modeling dynamic nonlinear time series due to its intelligence to identify time series patterns and nonlinear characteristics [31, 32].

As shown in Fig. 1, the elementary neuron is the basic building block of the network. It takes R inputs (P_i) and assigns an appropriate weight (W_{1i}) to the inputs. The weight factors model the synapse between the two neurons. Since a biological neuron has a threshold for activation, the artificial neurons are equipped with a transfer function. Further, a bias b is added to the algorithm as an offset value that helps the signal to exceed the transfer function's threshold. The weighted inputs $W_{1i}P_i$ and the bias are summed to form the net input

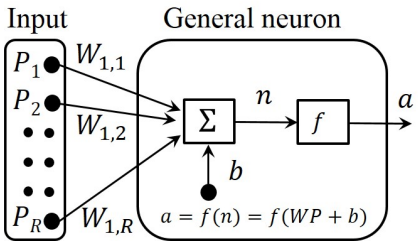

Figure 1: Elementary neuron with R inputs. The sum of weighted inputs and bias b form the net input n to the transfer function f to generate the output.

n to the transfer function. The output can be generated using various transfer functions, such as the linear function ($n = f(n) = a$), the log-sigmoid function ($n = f(n) = \frac{1}{1+e^{-n}}$), and the hyperbolic tangent sigmoid function ($a = f(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}}$). Using mathematical notation, the output a of a neuron can be presented as follows,

$$a = f(b + \sum_i W_{1i}P_i) \tag{1}$$

Artificial neural networks are formed by connecting individual neurons using weighted links. The neural network aims to transform the raw inputs into meaningful outputs. Among various types of neural networks, the multi-layer perceptron network stands out due to its flexibility since it contains an input layer, a variable number of hidden layers and an output layer. In the framework of the multi-layer perceptron, recurrent neural networks can add additional weights to the network to form a directed graph along a sequence, allowing it to use its internal state (memory) to process sequences of inputs and deal with problems with temporal component. The non-linear autoregressive neural network is a type of recurrent neural network that describes a discrete, non-linear, autoregressive system which is given by the following equation

$$y(t) = h(y(t-1), y(t-2), ..., y(t-d)) \tag{2}$$

where d is the number of delays, meaning that the neural network uses d previous values of the series to predict the value of a data series y at time t. $h(\cdot)$ is an unknown function ahead of the training of the neural network. After the training, this function is approximated by utilizing the optimization of the network weights and neuron bias.

As shown in Fig. 2, the non-linear autoregressive neural network is a multilayer, recurrent network with feedback connections and uses previous values of the actual time series to predict next values. The number of neurons and hidden layers are user definable. Decreasing the number of neurons and hidden layers can improve the computation time per iteration at the cost of its accuracy. It also has an added problem of being too specific to the

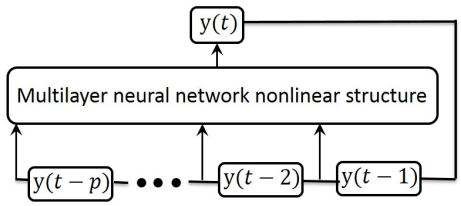

Figure 2: Non-linear autoregressive neural network is a recurrent, dynamic network with feedback connections. The next term in a sequence is predicted from a fixed number of previous terms using delay taps.

input data. On the other hand, a large number of neurons and layers will build a complicated network and lead to a overfitting problem. In this case, the performance on the training data will increase while the performance on unseen data becomes worse. Thus to provide reliable performance, these numbers need to be determined by optimization through a trial-and-error procedure.

The training process starts with tuning the weights and biases of the network to reach optimization which is defined by the minimization of the least mean square error (MSE). The error between the network output a and the target output t is given by

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \tag{3}$$

To minimise the linear combination of the squared errors, the common minimization algorithms available are the Bayesian regularization [33], Levenberg-Marguardt [34, 35], and the scaled conjugated gradient algorithms. Levenberg-Marguardt algorithm is mostly used due to its efficiency in generating a trained network which has excellent generalization abilities. Backpropagation in used to calculate the Jacobian of the performance index with respect to the weight and bias of the variables.

Since the nonlinear optimisation algorithm depends on the second order properties of the errors surface, the Hessian matrix is a required calculation. But since it is difficult to calculate efficiently, various methods are used to find an approximation. Hessian matrix H is approximated as $J^T J$, where J is the Jacobian for the calculation which is computed using the standard backpropagation method and is given by $J = \nabla \mathbf{E}$ [36]. $\mathbf{E}$ is the error function with respect to the weights. In this work, we assume that the performance is a given by MSE. The gradient is given by $g = J^T \mathbf{E}$.

The bias are optimized using the algorithm given

$$b_{k+1} = b_k - \left[J^T J + \mu I\right]^{-1} J^T \mathbf{E} \tag{4}$$

where μ is a scalar and I is an identity matrix.

The error autocorrelation function gives us the relation between the time and the errors in the prediction. In a perfect prediction model, the non-zero value should only occur at zero lag which is the MSE. If the error autocorrelation has a pattern, then it implies that the neural network is lacking some component which is generating the extra pattern in the errors. It can generally be improved by using a higher number of delays. For a series $y(t)$ the coefficients are given by

$$r_k = \frac{\sum_{t=k+1}^{T}\left(y_t - \overline{y}\right)\left(y_{t-k} - \overline{y}_{t-k}\right)/(T-k)}{\sum_{t=1}^{T}\left(y_t - \overline{y}\right)^2/T} \tag{5}$$

where $\overline{y}$ is the sample mean and r_k describes the auto correlation of y_t and y_{t-k}. For an artificial neuron network, the variable should be predictive and not correlated.

In this work, we have calculated one set of data for a specific case using the multiple Davydov D_2 *Ansatz*. This data set was calculated at a very small-time step to both give us many points to train with and increase the accuracy of the data. We have used this data to train a network with a large number of delays but a smaller number of neurons.

B. Multiple Davydov trial states

In this subsection, we present a brief introduction to time-dependent variation with the multiple Davydov *Ansätze* and define some operators of interest. For this purpose, we consider the spin-boson Hamiltonian as an example. The Hamiltonian is expressed as

$$\hat{H} = \hat{H}_S + \hat{H}_B + \hat{H}_{SB} \tag{6}$$

where

$$\hat{H}_S = \frac{\epsilon}{2}\sigma_z + \frac{\Delta}{2}\sigma_x \tag{7}$$

$$\begin{aligned}\hat{H}_B &= \sum_{q=1}^{N}\hbar\omega_q \hat{b}_q^\dagger \hat{b}_q \\ \hat{H}_{SB} &= \sum_{q=1}^{N}\frac{\lambda}{2}\sigma_z(\hat{b}_q^\dagger + \hat{b}_q)\end{aligned} \tag{8}$$

Here, σ_x and σ_z are the Pauli matrices. ϵ and Δ are the energy bias and the coupling constant between two electronic states, respectively. $\hbar = 1$ is assumed throughout. ω_q indicates the frequency of the q-th mode of the bath with creation (annihilation) operator $\hat{b}_q^\dagger(\hat{b}_q)$. λ_q is the strength of the coupling between the system and the q-th mode. The environment and its coupling to the system are characterized by an Ohmic type spectral density function,

$$J(\omega) = \sum_q \gamma_q^2 \delta(\omega - \omega_q) = 2\alpha\omega_c e^{-\omega/\omega_c} \tag{9}$$

where α is the dimensionless coupling strength, ω_c denotes the cutoff frequency.

Finite temperature effects can be simulated by adopting a Monte Carlo importance sampling method, in which initial conditions for variational parameters are chosen with the Monte Carlo importance sampling in accordance with the equilibrium distribution [37].

The multiple Davydov trial states with multiplicity M are essentially M copies of the corresponding single Davydov *Ansatz* [38, 39]. They were developed to investigate the polaron model [13, 14, 40] and the spin-boson model [41] following the Dirac-Frenkel variational principle. In the two-level system, one of the multiple Davydov trial states, the multiple D_2 *Ansatz*, also known as the multi-D_2 *Ansatz* with multiplicity M, can be constructed as

$$\begin{aligned}\left|D_2^M\right\rangle = &\sum_{i=1}^{M}\left\{A_i(t)\left|\uparrow\right\rangle \exp\left[\sum_{q=1}^{N} f_{iq}(t)\hat{b}_q^\dagger - H.c.\right]\left|0\right\rangle\right\}\\ &+\sum_{i=1}^{M}\left\{B_i(t)\left|\downarrow\right\rangle \exp\left[\sum_{q=1}^{N} f_{iq}(t)\hat{b}_q^\dagger - H.c.\right]\left|0\right\rangle\right\}, \quad (10)\end{aligned}$$

where $H.c.$ denotes the Hermitian conjugate, and $|0\rangle$ is the vacuum state of the bosonic bath. A_i and B_i are time-dependent variational parameter for the amplitudes in states $|\uparrow\rangle$ and $|\downarrow\rangle$, respectively, and $f_{iq}(t)$ are the bosonic displacements, where i and q label the i-th coherent superposition state and q-th effective bath mode, respectively. If $M = 1$, the multi-D_2 *Ansatz* is reduced to the usual Davydov D_2 trial state.

Equations of motion of the variational parameters $u_i = A_i, B_i$ and f_{iq} are then derived by adopting the Dirac-Frenkel variational principle,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{u}_i^*}\right) - \frac{\partial L}{\partial u_i^*} = 0. \quad (11)$$

For the multi-D_2 *Ansatz*, the Lagrangian L_2 is given by

$$\begin{aligned}L_2 &= \langle \mathrm{D}_2^M(t)|\frac{i\hbar}{2}\frac{\overleftrightarrow{\partial}}{\partial t} - \hat{H}|\mathrm{D}_2^M(t)\rangle\\ &= \frac{i\hbar}{2}\left[\langle \mathrm{D}_2^M(t)|\frac{\overrightarrow{\partial}}{\partial t}|\mathrm{D}_2^M(t)\rangle - \langle \mathrm{D}_2^M(t)|\frac{\overleftarrow{\partial}}{\partial t}|\mathrm{D}_2^M(t)\rangle\right]\\ &- \langle \mathrm{D}_2^M(t)|\hat{H}|\mathrm{D}_2^M(t)\rangle. \quad (12)\end{aligned}$$

Detailed derivations of the equations of motion for the variational parameters are given in Refs.[10, 42]. In the numerical calculations, the fourth-order Runge-Kutta method is used to integrate the equations of motion.

The expectation value of the observable of interest at zero temperature can be calculated as

$$\left\langle \hat{O}(t)\right\rangle = \left\langle D_2^M(t)\right| \hat{O} \left|D_2^M(t)\right\rangle \quad (13)$$

The expectation value of the observable at a finite temperature can be obtained numerically by the technique of Monte Carlo importance sampling as

$$\left\langle \hat{O}(t)\right\rangle = \frac{1}{N_s}\sum_{i}^{N_s}\left\langle D_2^M(t;\alpha_i)\right| \hat{O} \left|D_2^M(t;\alpha_i)\right\rangle \quad (14)$$

where N_s is the sampling number. The configuration α_i for the bath is numerically generated according to $p(\alpha;\beta)$ by importance sampling, where $p(\alpha;\beta)$ is the Boltzmann distribution used as the weighting function in the importance sampling procedure [37].

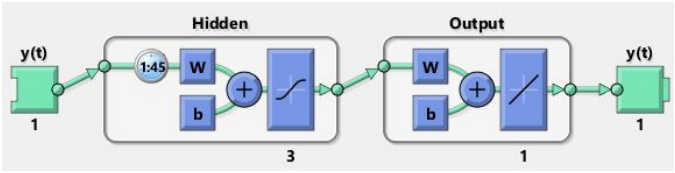

Figure 3: Structure of the neural network used to estimate time series of observables in open quantum systems discussed in this work. The hyperbolic tangent sigmoid transfer function and the linear transfer function are used in the hidden layer and output layer, respectively.

III. RESULTS

In this work, the trajectories of the correlated reduced open system state are computed by means of the time-dependent variational method. These trajectories contain key information about the environmental influence and the effect of initial correlations onto the open system dynamics. The construction of the dynamical neural net depends on the complexity of the observable trajectories. The created neural net needs careful testing and validation for further usage.

A. Dynamics of the spin-boson model

In this subsection, we apply the dynamical neural network method to the dynamics of the spin-boson model. In order to demonstrate the applicability of our approach to unraveling many-body dynamics, a principal observable of interest to look into is the population difference between the electronic states, $P_z(t)$, calculated as

$$P_z(t) = \langle\sigma_z(t)\rangle = \frac{1}{N_s}\sum_{i}^{N_s}\left\langle D_2^M(t;\alpha_i)\right| \sigma_z \left|D_2^M(t;\alpha_i)\right\rangle \quad (15)$$

Firstly, we study the dynamics of the spin-boson model with a single boson mode. The Hamiltonian is expressed by

$$\hat{H} = \frac{\epsilon}{2}\sigma_z + \frac{\Delta}{2}\sigma_x + \hbar\omega\hat{b}^\dagger\hat{b} + \frac{\gamma}{2}\sigma_z(\hat{b}^\dagger + \hat{b}), \quad (16)$$

which can be obtained from Hamiltonian (6) if only one boson mode is present. To study the dynamics of Hamiltonian (16), we have created a network using 3 hidden

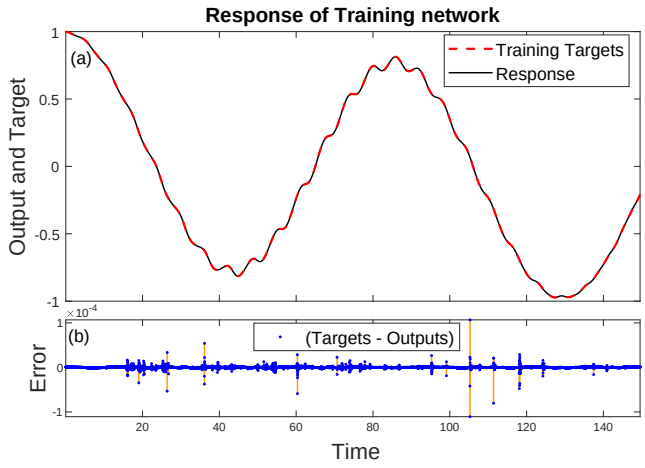

Figure 4: (a) Response of the neural net output to the training data. The training data is on $P_z(t)$ of the spin-boson model with a single boson mode. (b) Difference between the training data and their outputs. The parameter used by the multi-D_2 *Anstaz* for $P_z(t)$ is $M = 5$, $\epsilon = 0$, $\Delta = -0.1\omega$, $k_BT = 1\omega$, $\gamma = 0.8$ and $\omega = 1$.

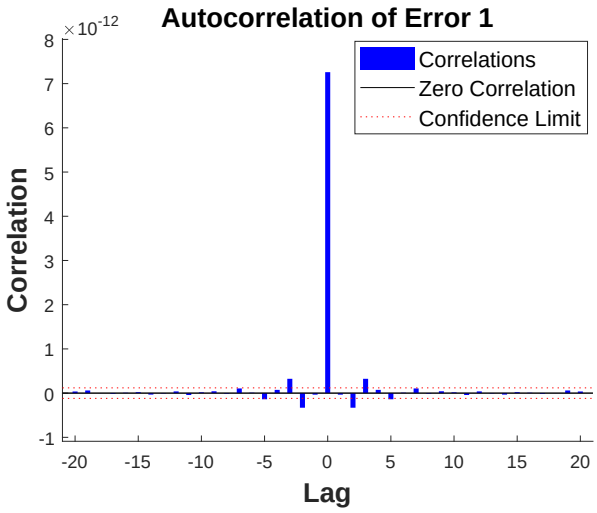

Figure 5: Autocorrelation of error in the prediction for the trained neural net. Autocorrelation coefficients measure the degree of correlation between neighboring data observations in the series of $P_z(t)$ for the coupling strength $\gamma = 0.8$.

neurons and 45 delays. This network is displayed as a graphical diagram in Fig. 3. The first block in the diagram shows the delay and the number of neurons used. It also indicates that the transfer function used to calculate the weight and biases is a hyperbolic tangent sigmoid transfer function.

The network created is populated by the data obtained from the multi-D_2 *Ansatz* with $\gamma = 0.8$. The response of the network to the training data can be seen in Fig. 4 (a). The accuracy of this result gives us the first indication in the training quality. In Fig. 4 (b) we have plotted the deviation of the variational outcome from the result obtained by the network. The numerical calculation in the variational method breaks down under very specific conditions. This manifests as small but sharp jumps (singularity points) in the simulation for the observables like $P_z(t)$. To mitigate this error a smaller time step can be utilised. But due to computational constrains, we have optimised between efficiency and accuracy while insuring that inaccuracies remain negligible. The network training is however generalized enough to not reflect the sharp jumps in its output. The number of delays implemented in the network is based on the error autocorrelation function results which are shown in Fig. 5. As stated earlier, for a good training the network should have a lag as close to 0 as possible to have a predictive network.

After the neural network is created and trained, it is tested with additional data to cross validate, which is a process to analyze the training of the network and its ability to predict. The cross validation is implemented for the trained network but using a different data set. This data was obtained by doing a separate numerical calculation for another set of parameters for the spin-boson model with a single boson mode. Validation results corresponding to the target is shown in Fig. 6. Validating the network is an important step as all machine learning algorithms must be checked for their ability to generalize. We can then use the trained network for testing a new set of data. This process also helps us understand the accuracy of the neural network. The neural network is validated using the data obtained from a different coupling strength of $\alpha = 0.5$. We can see that the output and the original data agree with each other. The MSE is less than 10^{-11}. This means that the testing output has a similar or better accuracy than that of the original calculation using the multi-D_2 *Ansatz*, as we allow for a maximum error of 5×10^{-4}. The network is validated using two data sets: $P_z(t)$ for a weak coupling strength and for an intermediate coupling strength.

Secondly, we study the dynamics of the spin-boson model with multiple boson modes, which is represented by Hamiltonian (6). A neural network is carefully trained further following the same procedure described above for the single mode scenario. We demonstrate that a well trained neural net is able to avoid the artificial recurrence problem induced by the limited number of boson modes in the variational method and accurately predict the dissipative behavior of $P_z(t)$ at long times.

As an open quantum system and its thermal bath with a finite number of modes together form an isolated system. According to the Poincaré recurrence theorem, the quantum system will eventually return to a state almost identical to its initial state [43]. To circumvent this recurrence, the bath has to be expanded theoretically to contain an infinite number of boson modes. In numerical calculations, an enormous number of boson modes are forced to be adopted to ensure that the Poincaré recurrence time $T_p = 2\pi/\Delta\omega$ exceeds any time scale of interest, where $\Delta\omega$ is the boson frequency spacing as the

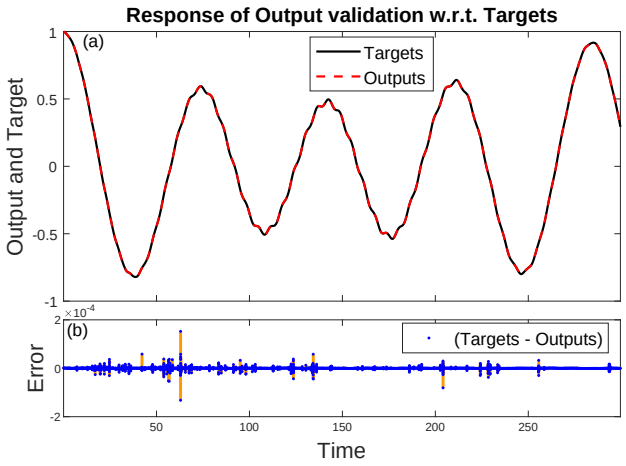

Figure 6: (a) Response of Output validation with respect to Targets. (b) Difference between the targets and the outputs. The target data is on $P_z(t)$ for $M = 5$, $\epsilon = 0$, $\Delta = -0.1\omega$, $k_BT = 1\omega$, $\gamma = 0.5$ and $\omega = 1$.

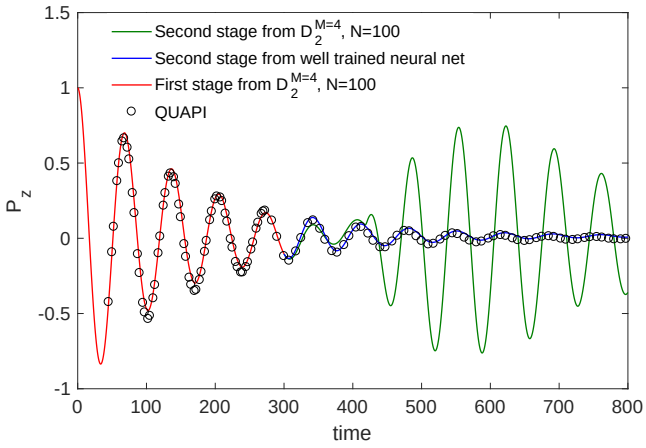

Figure 7: Dynamics of spin-boson model with multiple boson modes. The parameters used are $\epsilon = 0$, $\Delta = 0.1\omega_c$, $\alpha = 0.05$, $k_BT/\omega_c = 0.01$. In the variational calculation, the number of modes is $N = 100$ and multiplicity of multi-D_2 *Ansatz* is $M = 4$.

discretization of the spectral density function [44].

As shown in Fig. 7, a safe distance from the Poincaré recurrence is a timescale less than $t = 400/\omega_c$ if number of modes is set to $N = 100$. In contrast, the dissipated $P_z(t)$ is predicted exactly using the well trained dynamical neural net. After we combine the first stage calculation from the $D_2^{M=4}$ and the second stage from the neural net method, the result is in agreement with the benchmark result from the numerically exact iterative QUAPI technique.

B. Dynamics of the dissipative Landau-Zener model

In this subsection, the dynamical neural network is adopted to study a more complex scenario, the dissipative Landau-Zener Hamiltonian,

$$\hat{H} = \frac{vt}{2}\sigma_z + \frac{\Delta}{2}\sigma_x + \hbar\omega\hat{b}^\dagger\hat{b} + \frac{\gamma}{2}\sigma_x(\hat{b}^\dagger + \hat{b}), \qquad (17)$$

which can be obtained from the Hamiltonian (6) if the first term in Eq. (6) is replaced by a time-independent energy bias $vt\sigma_z/2$. When the energy difference between two diabatic states is swept through an avoided level crossing, Landau-Zener transition comes into play [45, 46]. The Landau-Zener transition is one of the most fundamental phenomena in quantum physics. It is applied to a growing list of physical systems, such as a superconducting flux qubit coupled to a quantum interference device [47] and a nitrogen-vacancy center spin in isotopically purified diamond [48]. The dissipative Landau-Zener transitions have been studied by various methods, such as time-dependent perturbation theory, [49], the master equation method [22], QUAPI and the non-equilibrium Bloch equations [15], and variational method with multi-D_2 *Ansatz* [10]. The observable of interest is the time evolution of the transition probability, $P_{\uparrow\to\downarrow}(t)$, calculated as

$$P_{\uparrow\to\downarrow}(t) = \left\langle D_2^M(t;\alpha_i)\right| \left|\downarrow\right\rangle \left\langle\downarrow\right| \left|D_2^M(t;\alpha_i)\right\rangle \qquad (18)$$

To apply the dynamical neural network method to the dissipative Landau-Zener dynamics, the procedure described in Sec. III A is used. Since Hamiltonian (17) is time-dependent and models a non-linear transition problem, we train the neural network with different parameters to generalize the greater complexity of the training data set. As shown in Appendix A, the neural network created is populated by the transition probability obtained from the multi-D_2 *Ansatz* with $\omega = 0.1\sqrt{v/\hbar}$. The trained network shows superior capability to generalize the time evolution of the transition probability for the trained data set.

To test for the validity of the network for different conditions, we calculated another set of numerical data with a frequency $\omega = \sqrt{v/\hbar}$ to compare against. The trained network is able to mimic the numerical result to a large extent, as shown in Fig. 8. The large single point inconsistencies are due to singularities issues in the framework of numerical methods. The output data is almost identical to the input data with the error less than 10^{-4}, indicating the high accuracy of our neural network. In particular, the deviation between the output and input is negligible even for the time period around $t = 0$, pointing to the conclusion that this neural network is valid for the transition process. It is shown that with sufficient training, the neural network method is valid for both the time-independent and time-independent Hamiltonian systems.

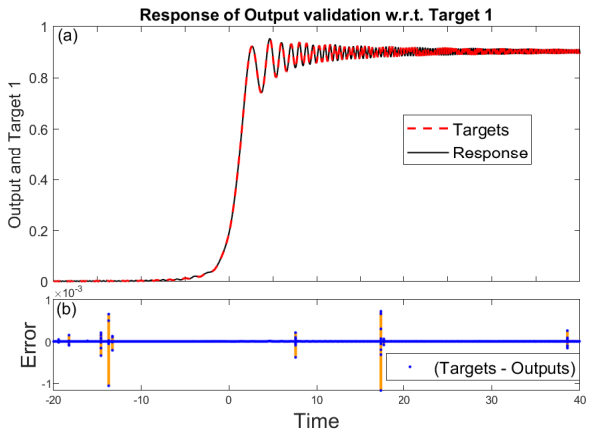

Figure 8: (a) Response of Output validation with respect to Targets. (b) The difference between the targets and the outputs. The target data is on $P_{\uparrow\to\downarrow}(t)$ for $\omega = \sqrt{v/\hbar}$, $\Delta = 0$, and $\gamma = 1.2\sqrt{\hbar v}$ calculated by the multi-D_2 *Anstaz*. Time unit is $\sqrt{\hbar\omega/v}$.

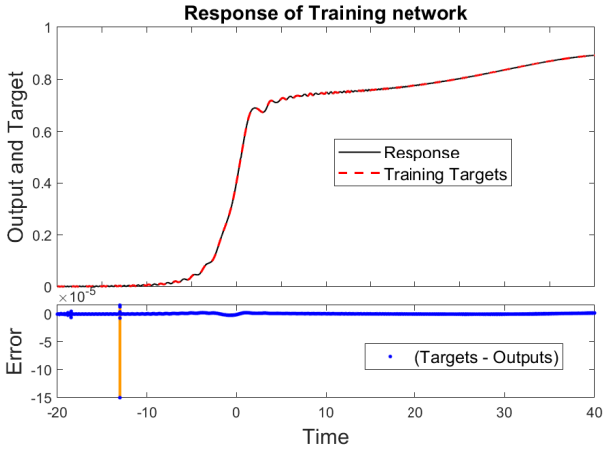

Figure 9: Response of the neural net output to the training data is shown in the upper panel. Difference between the targets and the outputs is plotted in the lower panel. The training data is on $P_{\uparrow\to\downarrow}(t)$ of the dissipative Landau-Zener model for $\omega = 0.1\sqrt{v/\hbar}$, $\Delta = 0$, and $\gamma = 1.2\sqrt{\hbar v}$ calculated by the multi-D_2 *Anstaz*. Time unit is $\sqrt{\hbar\omega/v}$.

IV. CONCLUSION

Accurate dynamics simulation of open quantum systems is often unattainable due to numerical challenges they pose. In this work, the dynamical neural network is employed to simulate open-system quantum dynamics. We first describe the training procedure and verify performance of the trained net by cross validation. This method is applied to the dynamics of the time-independent spin-boson Hamiltonian and the time-dependent Landau-Zener Hamiltonian. The trained net performed admirably under the validation testing we have carried out. We have also used this method for the prediction of time evolution of population in the spin-boson model with multiple bath modes at long times. The application of dynamical neural network on time-dependent variational parameters is of great interest and has potential to be exploited in future work. We have presented our preliminary results on using artificial intelligence to predict quantum observables, and hope that it provides inspiration to workers in the field.

Acknowledgments

This paper is dedicated to Prof. Wolfgang Domcke on the occasion of his 70th birthday. Support from the Singapo e National Research Foundation through the Competitive Research Programme (Project No. NRF-CRP5-2009-04) and the Singapore Ministry of Education Academic Research Fund (Grant No. RG106/15) is gratefully acknowledged. K. Sun also thanks the National Natural Science Foundation of China (Grant No. 11574052) and the Natural Science Foundation of Zhejiang Province (Grant No. LY18A040005) for partial support.

Appendix A: DYNAMICAL NEURAL NET TRAINING AND TESTING FOR LANDAU-ZENER PROBLEMS

To study the applicability of neural network to various complex systems, the dynamics for the Landau-Zener problem was chosen as a suitable candidate. The data was obtained from the time-dependent variational calculation using the multi D_2 *Ansatz*. The training data is the transition probability for $\omega = 0.1\sqrt{v/\hbar}$, $\Delta = 0$, and $\gamma = 1.2\sqrt{\hbar v}$.

Following a procedure similar to the spin boson system, we created another neural network and trained it. We have used five neurons with 10 delays in the hidden layer to form this network. The transfer function used for the hidden layer is the hyperbolic tangent sigmoid function. This was the least network size for stable analysis of the data.

Fig. 9 shows the response of the network to this training data and the error between the response and training data. As presented in the upper panel, the output of the network and the training data match almost perfectly in the tested regime. In the lower panel, we can see that the network adjusts extremely fast to the transition to as the error is very low. The network training can smear the sharp jumps in its output around $t = -13/\sqrt{\hbar\omega/v}$, which is related to the numerical singularity issue in the preliminary calculation. It can be found that even for this more challenging data set, the network validation shows significantly high accuracy.

[1] A. Thiel, J. Phys. G Nucl. Part. Phys. **16**, 867 (1990).
[2] R. J. Lipert, G. Bermudez, and S. D. Colson, J. Phys. Chem. **92**, 3801 (1988).
[3] W. Xie and W. Domcke, J. Chem. Phys. **147**, 184114 (2017).
[4] D. Bouwmeester, N. H. Dekker, F. E. v. Dorsselaer, C. A. Schrama, P. M. Visser, and J. P. Woerdman, Phys. Rev. A **51**, 646 (1995).
[5] W. Wernsdorfer, R. Sessoli, A. Caneschi, D. Gatteschi, and A. Cornia, Europhysics Lett. **50**, 552 (2000).
[6] L. Zhu, A. Widom, and P. M. Champion, J. Chem. Phys. **107**, 2859 (1997).
[7] G. D. Fuchs, G. Burkard, P. V Klimov, and D. D. Awschalom, Nat. Phys. **7**, 789 (2011).
[8] P. Benioff, J. Stat. Phys., **22**, 5 563C591, 1980.
[9] L. Chen, P. Shenai, F. Zheng, A. Somoza, and Y. Zhao, Molecules **20**, 15224 (2015)
[10] Z. Huang and Y. Zhao, Phys. Rev. A **97**, 13803 (2018).
[11] H. Wang, J. Phys. Chem. A **119**, 7951 (2015).
[12] E. Jeckelmann and S. R. White, Phys. Rev. B **57**, 6376 (1998).
[13] N. Zhou, Z. Huang, J. Zhu, V. Chernyak, and Y. Zhao, J. Chem. Phys. **143**, 014113 (2015).
[14] Z. Huang, L. Chen, N. Zhou, and Y. Zhao, Ann. Phys. **529**, 1600367 (2017).
[15] P. Nalbach, N. Klinkenberg, T. Palm, and N. Mller, Phys. Rev. E **96**, 42134 (2017).
[16] L. Chen, Y. Zhao, and Y. Tanimura, J. Phys. Chem. Lett. **6**, 3110 (2015).
[17] H. M. Wiseman, Quantum Semiclassical Opt. J. Eur. Opt. Soc. Part B **8**, 205 (1996).
[18] S. Popescu, A. J. Short, and A. Winter, Nat. Phys. **2**, 754 (2006).
[19] X. Wang and T. Xiang, Phys. Rev. B **56**, 5061 (1997).
[20] V. May and O. Kuhn, Charge and Energy Transfer Dynamics in Molecular Systems, Wiley-VCH, Weinheim, 2004.
[21] H. Carmichael, An Open Systems Approach to Quantum Optics, Springer, Berlin, 1993
[22] K. Saito, M. Wubs, S. Kohler, Y. Kayanuma, and P. Hänggi, Phys. Rev. B **75**, 214308 (2007).
[23] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 21021 (2017).
[24] X. Gao and L.-M. Duan, Nat. Commun. **8**, 662 (2017).
[25] C. Bény, arXiv Prepr. arXiv1301.3124 (2013).
[26] C. Edward Rasmussen, Bayesian Integrals. **7**, 651 (2003).
[27] J. Cerrillo and J. Cao, Phys. Rev. Lett. **112**, 110401 (2014).
[28] M. Buser, J. Cerrillo, G. Schaller, and J. Cao, Phys. Rev. A **96**, 62122 (2017).
[29] R. Rosenbach, J. Cerrillo, S. F. Huelga, J. Cao, and M. B. Plenio, New J. Phys. **18**, 23035 (2016).
[30] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[31] U. Hübner, N. B. Abraham, and C. O. Weiss, Phys. Rev. A **40**, 6354 (1989).
[32] U. Hübner, W. Klische, N. B. Abraham, and C. O. Weiss, in Coherence Quantum Opt. VI, edited by J. H. Eberly, L. Mandel, and E. Wolf (Springer US, Boston, MA, 1989), pp. 517C520.
[33] D. J. C. MacKay, Neural Comput., **4**, 3 415-447, May 1992.
[34] K. Levenberg, Quarterly of Applied Mathematics, **2**. Brown University, 164C168.
[35] D. W. Marquardt, J. Soc. Ind. Appl. Math., **11**, 2 431C441, Jun. 1963.
[36] M. T. Hagan and M. B. Menhaj, IEEE Trans. Neural Networks, **5**, 6 989C993, 1994.
[37] L. Wang, Y. Fujihashi, L. Chen, and Y. Zhao, J. Chem. Phys. **146**, 124127 (2017).
[38] Y. Zhao, B. Luo, Y. Zhang, and J. Ye, J. Chem. Phys. **137**, 084113 (2012).
[39] Y. Zhao, D. W. Brown, and K. Lindenberg, J. Chem. Phys. **107**, 3159 (1997); **107**, 3179 (1997).
[40] Z. Huang, L. Wang, C. Wu, L. Chen, F. Grossmann, and Y. Zhao, Phys. Chem. Chem. Phys. **19**, 1655 (2017).
[41] Z. Huang, Y. Fujihashi, and Y. Zhao, J. Phys. Chem. Lett. **8**, 3306 (2017).
[42] L. Wang, L. Chen, N. Zhou, Y. Zhao, J. Chem. Phys., **144**, 024101 (2016).
[43] P. Bocchieri and A. Loinger,Phys. Rev. **107**, 337 (1957); L. Barreira, "Poincaré recurrence: old and new," in Proceedings of the Fourth International Congress on Mathematical Physics (World Scientific, 2006), pp. 415C422
[44] N. Wu, L. Duan, X. Li, and Y. Zhao, 84111, (2014).
[45] C. Zener, Proc. R. Soc. London A **137**, 696 (1932).
[46] L. D. Landau, Phys. Z. **2**, 46 (1932)
[47] I. Chiorescu, P. Bertet, K. Semba, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij, Nature **431**, 159 (2004).
[48] J. Zhou, P. Huang, Q. Zhang, Z. Wang, T. Tan, X. Xu, F. Shi, X. Rong, S. Ashhab, and J. Du, Phys. Rev. Lett. **112**, 10503 (2014).
[49] P. Ao and J. Rammer, Phys. Rev. Lett. **62**, 3004 (1989).