Regular Article

From Short-Sighted to Far-Sighted: A Comparative Study of Recursive Machine Learning Approaches for Open Quantum Systems

Arif Ullah[1]

[1]*School of Physics and Optoelectronic Engineering, Anhui University, Hefei, 230601, Anhui, China.*

* Corresponding author: arif@ahu.edu.cn

Abstract: Accurately modeling the dynamics of open quantum systems is critical for advancing quantum technologies, yet traditional methods often struggle with balancing accuracy and efficiency. Machine learning (ML) offers a promising alternative, particularly through recursive models that predict system evolution based on the past history. While these models have shown success in predicting single observables, their effectiveness in more complex tasks, such as forecasting the full reduced density matrix (RDM), remains unclear. In this work, we extend history-based recursive ML approaches to complex quantum systems, comparing four physics-informed neural network (PINN) architectures: (i) single-RDM-predicting PINN (SR-PINN), (ii) SR-PINN with simulation parameters (PSR-PINN), (iii) multi-RDMs-predicting PINN (MR-PINN), and (iv) MR-PINN with simulation parameters (PMR-PINN). We apply these models to two representative open quantum systems: the spin-boson (SB) model and the Fenna-Matthews-Olson (FMO) complex. Our results demonstrate that single-RDM-predicting models (SR-PINN and PSR-PINN) are limited by a narrow history window, failing to capture the full complexity of quantum evolution and resulting in unstable long-term predictions, especially in nonlinear and highly correlated dynamics. In contrast, multi-RDMs-predicting models (MR-PINN and PMR-PINN) provide more accurate predictions by extending the forecast horizon, incorporating long-range temporal correlations, and mitigating error propagation. Surprisingly, including simulation parameters explicitly, such as temperature and reorganization energy, in PSR-PINN and PMR-PINN does not consistently improve accuracy and, in some cases, even reduces performance. This suggests that these parameters are already implicitly encoded in the RDM evolution, making their inclusion redundant and adding unnecessary complexity. These findings highlight the limitations of short-sighted recursive forecasting in complex quantum systems and demonstrate the superior stability and accuracy of far-sighted approaches for long-term predictions.

Open quantum systems describe quantum systems interacting with their environment, playing a fundamental role in quantum computing, quantum memory, quantum transport, proton tunneling in DNA and energy transfer in photosynthesis [1-5]. Their dynamics are captured by the reduced density matrix (RDM), which evolves under both the system internal dynamics and the influence of its environment.

Modeling the influence of environment is challenging due to its high-dimensional nature. Mixed quantumclassical methods [6-17] simplify the problem by treating the system quantum mechanically while approximating the environment classically, significantly reducing computational cost. However, these methods often struggle to capture detailed balance [18-20] or subtle quantum correlations [21]. Fully quantum approaches, including path-

integral [22-30] and quantum master equation-based methods [31-38], provide more accurate descriptions but are computationally expensive, particularly in regimes with strong system-environment coupling or where fine discretization is needed for numerical stability.

Recently, machine learning (ML) has emerged as a promising tool for learning complex spatiotemporal dynamics in high-dimensional systems [39-61]. One widely used ML strategy is the recursive approach, where the future evolution of a quantum state is predicted iteratively based on a short history of past evolution. This method has been successfully applied to the relaxation dynamics of the two-state spin-boson (SB) model [40,43,44,60], even enabling extrapolation beyond the trained time window [40]. However, previous applications have been limited to predicting a single observable—such as the population difference in the SB model—and have relied solely on singlestep prediction models.

In this work, we extend recursive ML approaches to more complex quantum systems, focusing on predicting the full RDM rather than just a single observable. We examine four physics-informed neural network (PINN)-based architectures: (i) the single-RDM-predicting PINN (SR-PINN), (ii) the SR-PINN with simulation parameters (PSR-PINN), (iii) the multi-RDMs-predicting PINN (MR-PINN), and (iv) the MR-PINN with simulation parameters (PMR-PINN). These architectures are tested on the relaxation dynamics of the SB model and the exciton energy transfer (EET) process in the Fenna-Matthews-Olson (FMO) complex.

From our results, we underscore the limitations of short-sighted, single-RDM-predicting models (SR-PINN and PSR-PINN) in capturing long-term system dynamics, especially in systems with intricate behavior. These models, constrained by a narrow history window, fail to predict long-term quantum evolution accurately, as they cannot fully capture the complexity of system evolution. In contrast, far-sighted models—such as MR-PINN and PMR-PINN—overcome these limitations by extending the forecast horizon, allowing them to incorporate long-range temporal correlations and achieve more stable predictions.

Although we initially anticipated that incorporating simulation parameters such as reorganization energy (λ), characteristic frequency (γ), and temperature (T) would improve accuracy, our findings show that these parameters do not consistently enhance performance and, in some instances, actually degrade it. This suggests that the relevant effects of these parameters are already implicitly encoded in the RDM evolution, making their explicit inclusion unnecessary in certain cases.

To build our case, let's consider an open quantum system (S), consisting of n states interacting with an external environment (E). As stated before, the dynamics of the system is governed by the RDM, which evolves non-unitarily due to environmental effects. While the full system follows unitary evolution described by the Liouville–von Neumann equation, tracing out the environmental degrees of freedom introduces a superoperator $\mathbf{R}$ that encodes dissipation and decoherence. Under the assumption that the initial state is separable between the system and environment ($\rho(0) = \rho_\text{S}(0) \otimes \rho_\text{E}(0)$), mathematically it can be described as

$$\rho_\text{S}(t) = \mathbf{Tr}_\text{E}\left(\mathbf{U}(t, 0)\,\rho(0)\,\mathbf{U}^\dagger(t, 0)\right)$$
$$= -i\,[\mathbf{H}_\text{S}, \rho_\text{S}(t)] + \mathbf{R}[\rho_\text{S}(t)], \qquad (1)$$

where $\rho_\text{S}(t)$ is the RDM of the system at time t, $\mathbf{Tr}_\text{E}$ denotes the partial trace over the environment, $\mathbf{R}$ is a superoperator that

encodes the effects of the environment and $\mathbf{U}(t, 0)$ and $\mathbf{U}^\dagger(t, 0)$ are the forward and backward time-evolution operators, respectively.

In the recursive ML framework, modeling the time evolution of Eq. (1) is formulated as learning a mapping function $\mathcal{M}$ that maps the input descriptors into predicted RDMs. In general, we have

$$\mathcal{M} : \{\mathbb{R}^{n \times n}\}^{k'} \rightarrow \{\mathbb{R}^{n \times n}\}^l, \qquad (2)$$

where $\{\mathbb{R}^{n \times n}\}^{k'}$ is a collection of k' input matrices (of size $n \times n$) that encode physical information such as historical RDM data, initial conditions, and simulation parameters, and $\{\mathbb{R}^{n \times n}\}^l$ is a sequence of l predicted RDMs corresponding to different time steps. In our study, we consider four distinct approaches for predicting the time evolution of RDM:

The *SR-PINN approach*: This method predicts the RDM at the next time step based solely on a fixed-length history of past RDMs. The recursive mapping function is defined as

$$\mathcal{M}_\text{rec} : \{\mathbb{R}^{n \times n}\}^{k'} \rightarrow \mathbb{R}^{n \times n}, \qquad (3)$$

with

$$\mathcal{M}_\text{rec}\left[\rho_S(t_{k-k'+1}), \rho_S(t_{k-k'+2}), \ldots, \rho_S(t_k)\right] = \rho_S(t_{k+1}). \qquad (4)$$

The procedure is applied iteratively: after predicting $\rho_S(t_{k+1})$, this new RDM is appended to the history while the oldest entry is removed, keeping the memory size constant at k'.

The *PSR-PINN approach*: To improve prediction accuracy, additional simulation parameters $\mathbf{p}$ (e.g., system–environment coupling, characteristic frequency, temperature) are incorporated into the input. The mapping function becomes

$$\mathcal{M}_\text{rec} : \mathbb{R}^p \times \{\mathbb{R}^{n \times n}\}^{k'} \rightarrow \mathbb{R}^{n \times n}, \qquad (5)$$

such that

$$\mathcal{M}_\text{rec}\left[\mathbf{p}, \left[\rho_S(t_{k-k'+1}), \ldots, \rho_S(t_k)\right]\right] = \rho_S(t_{k+1}). \qquad (6)$$

As with the standard SR-PINN, the process is applied recursively with a fixed history length.

The MR-PINN approach: Rather than predicting a single RDM at a time, the MR-PINN approach forecasts a block of future RDMs in one step. Its mapping function is defined by

$$\mathcal{M}_\text{rec} : \{\mathbb{R}^{n \times n}\}^{k'} \rightarrow \{\mathbb{R}^{n \times n}\}^{N_f}, \qquad (7)$$

with

$$\mathcal{M}_\text{rec}\left[\rho_S(t_{k-k'+1}), \ldots, \rho_S(t_k)\right]$$
$$= [\rho_S(t_{k+1}), \rho_S(t_{k+2}), \ldots, \rho_S(t_{k+N_f})]. \qquad (8)$$

In this case, the model outputs N_f future RDMs simultaneously, thus providing a multi-step prediction without requiring iterative updating.

The PMR-PINN approach: This variant extends the MRPINN method by including simulation parameters in the prediction. The mapping is defined as

$$\mathcal{M}_\text{rec} : \mathbb{R}^p \times \{\mathbb{R}^{n \times n}\}^{k'} \rightarrow \{\mathbb{R}^{n \times n}\}^{N_f}, \qquad (9)$$

so that

$$\mathcal{M}_{\text{rec}}\left[\mathbf{p}, [\rho_S(t_{k-k'+1}), \ldots, \rho_S(t_k)]\right]$$
$$= [\rho_S(t_{k+1}), \rho_S(t_{k+2}), \ldots, \rho_S(t_{k+N_f})]. \qquad (10)$$

By integrating the simulation parameters **p**, the model can adjust its predictions to account for different physical conditions while forecasting multiple future time steps concurrently.

Each of these approaches leverages the past history of RDMs (and optionally simulation parameters) to predict the future dynamics of the system, differing primarily in whether they predict a single RDM or multiple RDMs in one go.

To evaluate the proposed methods, we analyze the relaxation dynamics of the SB model and the EET process in the FMO complex (see Methods section for details). The models are implemented using a hybrid deep learning architecture that integrates convolutional neural networks (CNNs) with long short-term memory (LSTM) layers, followed by fully connected dense layers (CNN-LSTM). Following the approach outlined in Ref [39], training is optimized using a composite loss function, expressed as:

$$L = \alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3 + \alpha_4 L_4, \qquad (11)$$

where each loss term is defined as follows.

The first term, L_1, represents the mean squared error (MSE) between the predicted elements of the RDM, ρ_S, and the reference values, $\tilde{\rho}_S$:

$$L_1 = \frac{1}{N_t \cdot n^2} \sum_{t=1}^{N_t} \sum_{i,j=1}^{n} (\tilde{\rho}_{S,i,j}(t) - \rho_{S,i,j}(t))^2. \qquad (12)$$

Here, N_t denotes the number of time steps.

To ensure trace conservation of the density matrix, the second loss term, L_2, penalizes deviations of the trace from unity:

$$L_2 = \frac{1}{N_t} \sum_{t=1}^{N_t} (\mathbf{Tr}\rho_S(t) - 1)^2. \qquad (13)$$

The third term, L_3, enforces positive semi-definiteness by penalizing negative eigenvalues $\mu_i(t)$ of the density matrix:

$$L_3 = \frac{1}{N_t \cdot n} \sum_{t=1}^{N_t} \sum_{i=1}^{n} \max(0, -\mu_i(t))^2. \qquad (14)$$

Additionally, L_4 ensures that all eigenvalues remain within the valid range [0, 1], enforcing a key physical constraint of the RDM:

$$L_4 = \frac{1}{N_t \cdot n} \sum_{t=1}^{N_t} \sum_{i=1}^{n} (\text{clip}(\mu_i(t), 0, 1) - \mu_i(t))^2. \qquad (15)$$

The clipping function used here is defined as:

$$\text{clip}(\mu_i(t), 0, 1) = \begin{cases} 0, & \text{if } \mu_i(t) < 0, \\ \mu_i(t), & \text{if } 0 \le \mu_i(t) \le 1, \\ 1, & \text{if } \mu_i(t)) > 1. \end{cases} \qquad (16)$$

The weighting coefficients $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ control the relative contributions of these loss terms. In our case, we set them all to unity ($\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1.0$). Collectively, these loss components ensure that the predicted RDM satisfies key physical properties: accuracy (L_1), trace conservation (L_2), positive semi-definiteness (L_3), and eigenvalue constraints (L_4).

For demonstration, we use data from the publicly available QD3SET-1 database [62] for both the SB model and the FMO complex. The models are trained on 80% of the simulations, with the remaining 20% reserved for testing. Further details on the dataset and training process can be found in the Methods section.
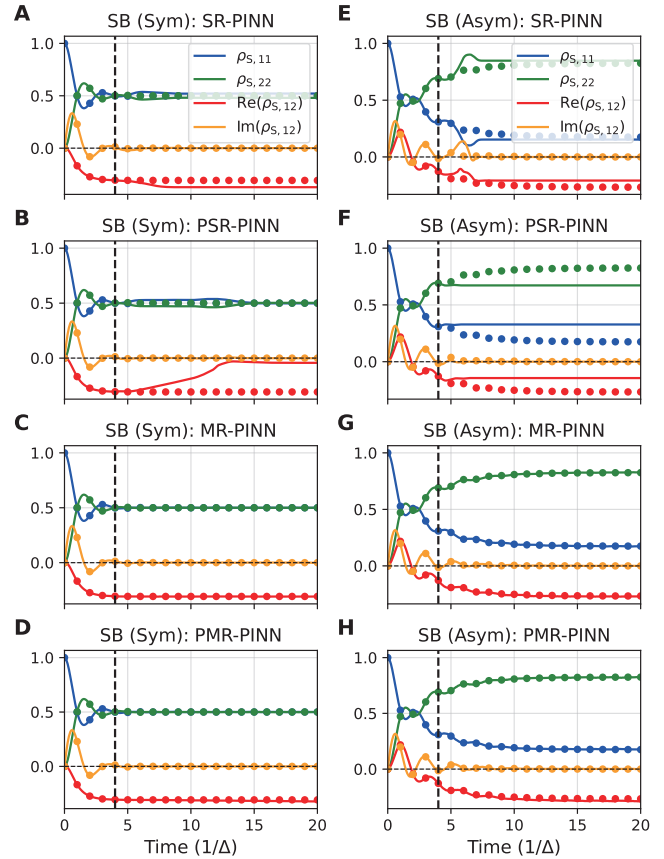

Figure 1. Time evolution of the RDM elements, including both population and coherence terms, as predicted by the SR-PINN, PSR-PINN, MR-PINN, and PMR-PINN models. The first column shows the RDM evolution for the symmetric (Sym) SB model, while the second column displays the corresponding dynamics for the asymmetric (Asym) SB model. Predictions are generated recursively, starting from an initial seed dynamics of time-length $4/\Delta$, and are compared to reference results (shown as dots). The dashed vertical line split the seed dynamics (left) and the predicted dynamics (right). For the symmetric case, the parameters used correspond to an unseen set: $\varepsilon/\Delta = 0.0$, $\gamma/\Delta = 3.0$, $\lambda/\Delta = 0.6$, and $\beta\Delta = 1.0$. In the asymmetric case, the parameters are $\varepsilon/\Delta = 1.0$, $\gamma/\Delta = 9.0$, $\lambda/\Delta = 0.6$, and $\beta\Delta = 1.0$.

Figure 1 presents the predictive performance of all the four models for the time evolution of RDM elements in both symmetric and asymmetric SB models. Each model is provided with an initial short-time seed ($4/\Delta$) and tasked with recursively forecasting the system's future evolution.

The results highlight the limitations of SR-PINN, which exhibits significant errors in both diagonal and off-diagonal
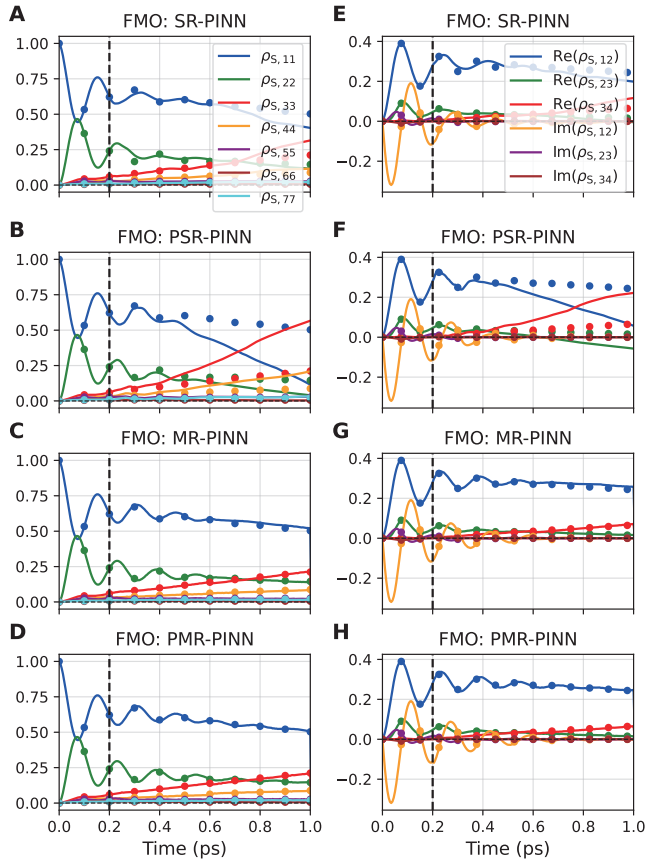
Figure 2. Time evolution of the RDM elements for the FMO complex with initial excitation on site-1, as predicted by SRPINN, PSR-PINN, MR-PINN, and PMR-PINN. The first column presents the population dynamics of exciton energy transfer (EET), while the second highlights selected coherence elements. Predictions are generated recursively using an initial seed dynamics of 0.2 ps and are compared to reference dynamics (shown as dots). The dashed vertical line split the seed dynamics (left) and the predicted dynamics (right). The test trajectory corresponds to simulation parameters $\gamma = 400$ cm^{-1}, $\lambda = 40$ cm^{-1}, and $T = 90$ K.

terms (population and coherence), leading to a rapid divergence from the expected dynamics. PSR-PINN, despite incorporating simulation parameters, further degrades accuracy, indicating that the past history window of the model remains insufficient for stable recursive predictions. In contrast, MR-PINN, which leverages a longer forecasting horizon, effectively mitigates error accumulation and successfully captures both population and coherence dynamics across the prediction window. PMR-PINN performs similarly to MR-PINN, suggesting that the inclusion of simulation parameters does not provide additional benefits in this setting.

To test a larger system, in Figure 2 we showcase the predicted evolution of RDM elements for the FMO complex under initial excitation on site-1. The models are trained with an initial short-time seed (0.2 ps) and recursively predict the system's future dynamics.

SR-PINN exhibits considerable inaccuracies, particularly in long-term dynamics, leading to deviations from the expected population transfer trends. PSR-PINN, despite integrating simulation parameters, fails to improve performance and even amplifies errors, especially in diagonal elements. As in the SB model, MR-PINN achieves significantly enhanced accuracy, demonstrating robust predictions of both energy transfer and coherence decay. PMR-PINN yields results comparable to MR-PINN, reinforcing the observation that the longer forecasting window is the primary factor driving predictive stability.

A quantitative analysis of model performance, summarized in the accompanying table (Table I), further substantiates these findings. As in the SB model, SRPINN struggles to maintain accuracy, particularly for coherence elements, and PSR-PINN further aggravates errors. MR-PINN consistently outperforms both singletime- step approaches, achieving the lowest mean absolute errors (MAE) across all RDM elements. The inclusion of simulation parameters in PMR-PINN does not lead to meaningful improvements over MR-PINN.

Notably, the errors are higher for the asymmetric SB model compared to the symmetric case, indicating that prediction accuracy degrades as system complexity increases. This trend suggests that more intricate dynamical behaviors impose additional challenges for PINN-based models, particularly when using single-RDM-predicting training strategies.

For the FMO complex, the disparity between methods remains

Table 1. Time-averaged mean absolute error (MAE) for the diagonal (Diag) and off-diagonal (Off-diag) elements of the RDMs predicted by the SR-PINN, PSR-PINN, MR-PINN, and PMR-PINN models for the test trajectory of the SB model and FMO complex. Off-diagonal errors represent the average MAE for both real and imaginary components. Values are expressed in the form 10^x.

Model	SB Model (Sym)		SB Model (Asym)	
	Diag	Off-diag (Real, Imag)	Diag	Off-diag (Real, Imag)
SR-PINN	1.6e-2	(4.4e-2, 1.7e-3)	2.9e-2	(3.8e-2, 7.8e-3)
PSR-PINN	1.3e-2	(1.3e-1, 1.1e-3)	1.0e-1	(7.9e-2, 3.6e-3)
MR-PINN	4.9e-4	(7.3e-4, 5.1e-4)	1.4e-3	(3.9e-3, 9.1e-4)
PMR-PINN	6.7e-4	(5.1e-3, 8.8e-4)	1.3e-3	(1.2e-2, 1.2e-3)
Model	FMO	Complex (site-1)	FMO	Complex (site-6)
	Diag	Off-diag (Real, Imag)	Diag	Off-diag (Real, Imag)
SR-PINN	1.4e-2	(4.1e-3, 8.6e-4)	1.3e-2	(2.3e-3, 1.6e-4)
PSR-PINN	6.3e-2	(1.6e-2, 4.8e-4)	1.1e-1	(2.9e-2, 2.9e-3)
MR-PINN	2.6e-2	(6.1e-3, 2.8e-4)	2.5e-2	(2.3e-3, 1.5e-4)
PMR-PINN	2.5e-2	(5.9e-3, 4.2e-4)	2.6e-2	(2.3e-3, 1.4e-4)

evident. SR-PINN and PSR-PINN perform poorly, with PSR-PINN producing the highest errors for both population and coherence terms, particularly when the initial excitation occurs at site-6 (as shown in the table). MR-PINN and PMR-PINN provide a marked improvement, although predictive errors remain higher compared to the SB model, reflecting the increased complexity of the system. The trend observed in the SB model, where errors increase with system complexity, is also evident in the FMO complex. The lack of significant gains from PMR-PINN over MR-PINN suggests that the longer prediction window is the dominant factor in improving accuracy, while the inclusion of simulation parameters has a limited effect.

In summary, this work investigated four PINN-based architectures for predicting the time evolution of the RDM in selected open quantum systems: single-RDM-predicting models (SR-PINN, PSR-PINN) and multi-RDMs-predicting models (MR-PINN, PMR PINN). These models use historical RDM data to predict future dynamics, with some incorporating environmentspecific parameters such as temperature and system-bath coupling.

Our findings reveal the limitations of single-RDM-predicting models (SR-PINN and PSR-PINN) in capturing long-term quantum dynamics. These models, constrained by a narrow history window, struggle to predict long-term quantum evolution accurately, as they fail to capture the full complexity of system evolution. However, as demonstrated in previous works [40,43,44,60], when trained on simpler observables—such as the population difference in the spin-boson model—they yield reasonable predictions, suggesting that single-variable evolution is easier to propagate recursively.

In contrast, multi-RDMs-predicting models (MRPINN and PMR-PINN) consistently provide stable and accurate long-term predictions across various scenarios. By predicting multiple RDMs in one step, these models mitigate cumulative errors and better capture long-range temporal correlations, improving their ability to generalize to unseen conditions. This emphasizes that extending the forecast horizon is more effective than merely increasing the historical input length, as explicitly forecasting future states stabilizes predictions more effectively than relying solely on past dynamics. Surprisingly, explicitly incorporating simulation parameters—such as reorganization energy, characteristic frequency, and temperature (in PSR-PINN and PMR-PINN)—did not consistently improve predictive accuracy and, in some cases, slightly reduced performance. This suggests that the effects of these parameters are already implicitly captured in the RDM evolution, rendering their explicit inclusion redundant and potentially introducing unnecessary complexity.

Overall, this work underscores the limitations of shortsighted, single-step recursive models in complex quantum systems and reinforces the advantages of far-sighted, multi-step approaches for robust, long-term predictions. Our findings highlight that incorporating a longer predictive horizon is key to improving prediction stability, capturing complex dynamics, and reducing the impact of short-term fluctuations in open quantum systems.

Methods

Hamiltonians of the SB model and FMO complex: The SB model describes a two-level system interacting with an environment composed of independent harmonic oscillators. The Hamiltonian of the system is given by:

$$\mathbf{H} = \epsilon\sigma_z + \Delta\sigma_x + \sum_k \omega_k \mathbf{b}_\mathbf{k}^\dagger \mathbf{b}_\mathbf{k} + \sigma_z \sum_k c_k(\mathbf{b}_\mathbf{k}^\dagger + \mathbf{b}_\mathbf{k}), \qquad (17)$$

where σ_z and σ_x are Pauli matrices, ϵ denotes the energy difference between the two states, and Δ represents their coupling strength. The surrounding environment consists of harmonic oscillators characterized by creation and annihilation operators $\mathbf{b}_\mathbf{k}^\dagger$ and $\mathbf{b}_\mathbf{k}$, corresponding to mode k with frequency ω_k. The system-bath interaction is governed by the coupling coefficient c_k for each mode.

Our next system of interest, the FMO complex is a trimeric protein found in green sulfur bacteria, where it plays a crucial role in photosynthetic energy transfer. Each monomer of the FMO complex contains multiple chlorophyll molecules—typically seven or eight—that facilitate exciton transport [63]. The excitonic dynamics within a monomer can be described by the Frenkel exciton model Hamiltonian [64]:

$$\begin{aligned}\mathbf{H} &= \sum_{i=1}^n |i\rangle\, \epsilon_i\, \langle i| + \sum_{i\neq j}^n |i\rangle\, J_{ij}\, \langle j| \\ &+ \sum_{i=1}^j \sum_{k=1}^n \left(\frac{1}{2}\mathbf{P}_{k,i}^2 + \frac{1}{2}\omega_{k,i}^2 \mathbf{Q}_{k,i}^2\right)\mathbf{I} \\ &- \sum_{i=1}^n \sum_{k=1}^n |i\rangle\, c_{k,i}\, \mathbf{Q}_{k,i}\, \langle i| + \sum_{i=1}^n |i\rangle\, \lambda_i\, \langle i|, \qquad (18)\end{aligned}$$

where n denotes the number of chlorophyll sites, ϵ_i is the site energy, and J_{ij} represents the electronic coupling between sites i and j. The operators $\mathbf{P}_{k,i}$ and $\mathbf{Q}_{k,i}$ correspond to the momentum and position of the k-th vibrational mode associated with site i, while $\omega_{k,i}$ is its frequency. The identity matrix $\mathbf{I}$ ensures proper dimensional consistency in the model. The coupling strength between site i and the k-th vibrational mode is given by $c_{k,i}$, and λ_i represents the reorganization energy of site i.

In both the SB model and the FMO complex, the environmental influence is characterized by the Debye spectral density:

$$J(\omega) = 2\lambda\, \frac{\gamma\omega}{\omega^2 + \gamma^2}, \qquad (19)$$

where λ is the reorganization energy, and γ is the characteristic frequency, defined as the inverse of the relaxation time ($\gamma = 1/\tau$). For the FMO complex, we assume that all chlorophyll sites experience the same environmental conditions.

Data extraction: For training our models, we utilized precomputed RDMs provided by the QD3SET-1 database [62], which contains simulations based on the hierarchical equations of motion (HEOM) method [22,27,65,66]. In the case of the SB model, our dataset, labeled D_{sb}, comprises 1000 simulations covering a four-dimensional parameter space defined by ε/Δ, λ/Δ, γ/Δ, and $\beta\Delta$, corresponding to the system bias, bath reorganization energy, bath relaxation rate, and inverse temperature, respectively. For the seven-site FMO complex, we also used 1000 simulations from QD3SET-1 that detail the exciton dynamics starting from excitations at site-1 and site-6, spanning the parameter set (λ, γ, T). In this dataset, the dynamics was generated using the trace-conserving local thermalizing Lindblad master equation (LTLME) [67], with Hamiltonian parameters taken from the work of Adolphs and Renger [68]. Specifically, the FMO Hamiltonian, $\mathbf{H}_\mathrm{S}$, is expressed as

$$\mathbf{H}_S = \begin{pmatrix} 200 & -87.7 & 5.5 & -5.9 & 6.7 & -13.7 & -9.9 \\ -87.7 & 320 & 30.8 & 8.2 & 0.7 & 11.8 & 4.3 \\ 5.5 & 30.8 & 0 & -53.5 & -2.2 & -9.6 & 6.0 \\ -5.9 & 8.2 & -53.5 & 110 & -70.7 & -17.0 & -63.6 \\ 6.7 & 0.7 & -2.2 & -70.7 & 270 & 81.1 & -1.3 \\ -13.7 & 11.8 & -9.6 & -17.0 & 81.1 & 420 & 39.7 \\ -9.9 & 4.3 & 6.0 & -63.6 & -1.3 & 39.7 & 230 \end{pmatrix}, \quad (20)$$

with an added diagonal offset of 12210 cm^{-1}.

Data Preparation: To construct the training dataset, each RDM, $\rho_S(t)$, along with its associated coefficients, is flattened into a one-dimensional vector. Given the Hermitian property of the RDM ($\rho_{S,ij}(t) = \rho_{S,ji}(t)^*$), we retain only the real components of the diagonal elements while including both the real and imaginary parts of the upper triangular off-diagonal elements.

Each simulation trajectory is then divided into multiple training samples. In the recursive training framework, an initial segment of the system's dynamics, $\{\rho_S(t_0), \rho_S(t_1), \ldots, \rho_S(t_k)\}$, serves as the input sequence. For the single-RDM-predicting models (SR-PINN and PSR-PINN), the dataset is structured to predict the immediate next RDM, $\rho_S(t_{k+1})$. The input sequence is updated at each step by appending the newly predicted RDM while discarding the earliest one ($\rho_S(t_0)$), maintaining a fixed sequence length. This iterative process continues until the final time step t_K is reached.

For the multi-RDMs-predicting models (MR-PINN and PMR-PINN), the output at each prediction step is a block of N_f future RDMs, predicted simultaneously. The number of predicted steps, N_f, is determined by the length of the prediction window t_w and the time-step dt used in propagating the dynamics, via the relation $N_f = t_w/dt$.

In the SB model, the prediction window is set to $t_w = 2/\Delta$, leading to $N_f = 40$ time steps given the chosen $dt = 0.05/\Delta$. Similarly, for the FMO complex, the prediction window is $t_w = 0.4$ ps, corresponding to $N_f = 80$ time steps with $dt = 0.005$ ps.

More generally, the number of training samples that can be generated from a single simulation depends on the propagation interval $[t_k, t_K]$ and the prediction window t_w, and is approximately given by $(t_K - t_k)/t_w$. These parameters vary depending on the system and the prediction strategy.

For the single-RDM-predicting approaches, a single time-step prediction window is used. In the SB model, we use $t_k = 4/\Delta$, $t_K = 20/\Delta$, and $t_w = dt = 0.05/\Delta$; for the FMO complex, the settings are $t_k = 0.2$ ps, $t_K = 1$ ps, and $t_w = dt = 0.005$ ps. In contrast, for the multi-RDMs models, the effective t_w corresponds to the length of the output sequence—$2/\Delta$ for SB and 0.4 ps for FMO—resulting in fewer but richer training samples, each covering a longer prediction horizon.

It is important to emphasize that the selection of t_k (input sequence length) and N_f (number of predicted steps) is informed by system complexity, physical intuition, and empirical testing. The value of t_k should be sufficient to encode the temporal structure and variability of the input RDMs, allowing the model to distinguish between different states. The choice of N_f depends on the desired prediction horizon and the system's complexity—for instance, $N_f = 40$ is adequate for the SB model, while the more intricate FMO complex benefits from a longer horizon ($N_f = 80$). As a practical guideline, we recommend setting N_f equal to the time span of the input trajectory to maintain consistency and provide a balanced context for prediction.

It is important to note that in the PSR-PINN and PMR-PINN models, simulation parameters were normalized by their respective maximum values. The normalized simulation parameters are expressed as $\lambda/\lambda_{\max}$, $\gamma/\gamma_{\max}$, $\beta/\beta_{\max}$ and $T/T_{\max}$, where $\lambda_{\max}$, $\gamma_{\max}$, $\beta_{\max}$ and $T_{\max}$ correspond to the maximum values of λ, γ, β and T, respectively.

Training and Prediction Strategies: To improve training efficiency, we utilize farthest point sampling [41,69] to select a representative subset of simulation trajectories. For each case in SB model ($\varepsilon/\Delta = 0$ and 1) and the FMO complex (initial excitations on site-1 and site-6), 400 simulation trajectories are allocated for training, with the remaining data reserved for testing. The training is conducted using a CNN-LSTM architecture, where convolutional layers are followed by LSTM layers and fully connected dense layers. For the SB model, a single CNNLSTM model is trained for both cases ($\varepsilon/\Delta = 0$ and 1), whereas for the FMO complex, separate models are trained for initial excitations on site-1 and site-6. To ensure a fair comparison, all models share an identical architecture, and during inference, models are selected based on comparable training and validation loss values.

Inference follows the same approach as the training data preparation. In single RDM prediction, a short sequence of past RDMs, $\{\rho_S(t_0), \rho_S(t_1), \ldots, \rho_S(t_k)\}$, serves as the input seed. The model predicts the next RDM, $\rho_S(t_{k+1})$, which is then appended to the input sequence while the oldest RDM is removed. This iterative process continues until the entire trajectory is predicted.

For multi-RDMs prediction, a similar strategy is employed, but instead of predicting a single RDM, the model generates a window of N_f future RDMs in each step. The input sequence is then updated with the last N_k RDMs from the newly predicted block, allowing for the efficient generation of extended dynamics.

Acknowledgments

Supporting information

The code and data supporting this work are available at https://github.com/Arif-PhyChem/rc-pinn-comparison.

References

[1] Breuer H. P., Laine E. M., Piilo J., Vacchini B. Colloquium: nonmarkovian dynamics in open quantum systems. *Rev. Mod. Phys.*, **88** (2) (2016), 021002.

[2] Khodjasteh K., Sastrawan J., Hayes D., Green T. J., Biercuk M. J., Viola L. Designing a practical high-fidelity long-time quantum memory. *Nat. Commun.*, **4** (1) (2013), 2045.

[3] Cui P., Li X. Q., Shao J., Yan Y. Quantum transport from the perspective of quantum open systems. *Phys. Lett. A*, **357** (6) (2006), 449-453.

[4] Slocombe L., Sacchi M., Al-Khalili J. An open quantum systems approach to proton tunnelling in DNA. *Commun. Phys.*, **5** (1) (2022), 109.

[5] Zerah Harush E., Dubi Y. Do photosynthetic complexes use quantum coherence to increase their efficiency? probably

not. *Sci. Adv.*, **7** (8) (2021), eabc4631.

[6]   Miller W. H. The semiclassical initial value representation: a potentially practical way for adding quantum effects to classical molecular dynamics simulations. *J. Phys. Chem. A*, **105** (13) (2001), 2942-2955.

[7]   Cotton S. J., Miller W. H. Symmetrical windowing for quantum states in quasi-classical trajectory simulations: application to electronically non-adiabatic processes. *J. Chem. Phys.*, **139** (23) (2013).

[8]   Liu J., He X., Wu B. Unified formulation of phase-space mapping approaches for nonadiabatic quantum dynamics. *Acc. Chem. Res.*, **54** (23) (2021), 4215-4228.

[9]   Runeson J. E., Richardson J. O. Spin-mapping approach for nonadiabatic molecular dynamics. *J. Chem. Phys.*, **151** (4) (2019), 044119.

[10] Runeson J. E., Richardson J. O. Generalized spin mapping for quantum-classical dynamics. *J. Chem. Phys.*, **152** (8) (2020).

[11] Mannouch J. R., Richardson J. O. A partially linearized spin-mapping approach for nonadiabatic dynamics. I. Derivation of the theory. *J. Chem. Phys.*, **153** (19) (2020).

[12] Mannouch J. R., Richardson J. O. A partially linearized spin-mapping approach for nonadiabatic dynamics. II. Analysis and comparison with related approaches. *J. Chem. Phys.*, **153** (19) (2020).

[13] Mannouch J. R., Richardson J. O. A partially linearized spin-mapping approach for simulating nonlinear optical spectra. *J. Chem. Phys.*, **156** (2) (2022).

[14] Tao G. A multi-state trajectory method for non-adiabatic dynamics simulations. *J. Chem. Phys.*, **144** (9) (2016).

[15] Mannouch J. R., Richardson J. O. A mapping approach to surface hopping. *J. Chem. Phys.*, **158** (10) (2023).

[16] Crespo-Otero R., Barbatti M. Recent advances and perspectives on nonadiabatic mixed quantum–classical dynamics. *Chem. Rev.*, **118** (15) (2018), 7026-7068.

[17] Qiu J., Lu Y., Wang L. Multilayer subsystem surface hopping method for large-scale nonadiabatic dynamics simulation with hundreds of thousands of states. *J. Chem. Theory Comput.*, **18** (5) (2022), 2803-2815.

[18] Schmidt J. R., Parandekar P. V., Tully J. C. Mixed quantum-classical equilibrium: surface hopping. *J. Chem. Phys.*, **129** (4) (2008), 044104.

[19] Amati G., Runeson J. E., Richardson J. O. On detailed balance in nonadiabatic dynamics: from spin spheres to equilibrium ellipsoids. *J. Chem. Phys.*, **158** (6) (2023), 064113.

[20] Amati G., Mannouch J. R., Richardson J. O. Detailed balance in mixed quantum–classical mapping approaches. *J. Chem. Phys.*, **159** (2023), 214114.

[21] Mannouch J. R., Kelly A. Toward a correct description of initial electronic coherence in nonadiabatic dynamics simulations. *J. Phys. Chem. Lett.*, **15** (2024), 11687-11695.

[22] Tanimura Y., Kubo R. Time evolution of a quantum system in contact with a nearly Gaussian-Markoffian noise bath. *J. Phys. Soc. Jpn.*, **58** (1) (1989), 101-114.

[23] Makarov D. E., Makri N. Path integrals for dissipative systems by tensor multiplication: condensed-phase quantum dynamics for arbitrarily long time. *Chem. Phys. Lett.*, **221** (5-6) (1994), 482-491.

[24] Su Y., Chen Z. H., Wang Y., Zheng X., Xu R. X., Yan Y.

Extended dissipaton equation of motion for electronic open quantum systems: application to the Kondo impurity model. *J. Chem. Phys.*, **159** (2) (2023).

[25] Yan Y., Xu M., Li T., Shi Q. Efficient propagation of the hierarchical equations of motion using the Tucker and hierarchical Tucker tensors. *J. Chem. Phys.*, **154** (19) (2021).

[26] Gong H., Ullah A., Ye L., Zheng X., Yan Y. Quantum entanglement of parallel-coupled double quantum dots: a theoretical study using the hierarchical equations of motion approach. *Chin. J. Chem. Phys.*, **31** (4) (2018), 510.

[27] Xu M., Yan Y., Shi Q., Ankerhold J., Stockburger J. Taming quantum noise for efficient low-temperature simulations of open quantum systems. *Phys. Rev. Lett.*, **129** (23) (2022), 230601.

[28] Bai S., Zhang S., Huang C., Shi Q. Hierarchical equations of motion for quantum chemical dynamics: recent methodology developments and applications. *Acc. Chem. Res.*, **57** (21) (2024), 3151-3160.

[29] Wang Y., Mulvihill E., Hu Z., Lyu N., Shivpuje S., Liu Y., et al. Simulating open quantum system dynamics on NISQ computers with generalized quantum master equations. *J. Chem. Theory Comput.*, **19** (15) (2023), 4851-4862.

[30] Makri N. Quantum dynamics methods based on the real-time path integral. In: *Compr. Comput. Chem.*, Elsevier (2023), V4-293.

[31] Han L., Chernyak V., Yan Y. A., Zheng X., Yan Y. Stochastic representation of non-Markovian fermionic quantum dissipation. *Phys. Rev. Lett.*, **123** (5) (2019), 050601.

[32] Han L., Ullah A., Yan Y. A., Zheng X., Yan Y., Chernyak V. Stochastic equation of motion approach to fermionic dissipative dynamics. I. Formalism. *J. Chem. Phys.*, **152** (20) (2020), 204105.

[33] Ullah A., Han L., Yan Y. A., Zheng X., Yan Y., Chernyak V. Stochastic equation of motion approach to fermionic dissipative dynamics. II. Numerical implementation. *J. Chem. Phys.*, **152** (20) (2020), 204106.

[34] Chen L., Bennett D. I., Eisfeld A. Simulation of absorption spectra of molecular aggregates: a hierarchy of stochastic pure-state approach. *J. Chem. Phys.*, **156** (12) (2022).

[35] Dan X., Xu M., Yan Y., Shi Q. Generalized master equation for charge transport in a molecular junction: exact memory kernels and their high-order expansion. *J. Chem. Phys.*, **156** (13) (2022).

[36] Stockburger J. T. Exact propagation of open quantum systems in a system-reservoir context. *EPL*, **115** (4) (2016), 40010.

[37] Lyu N., Mulvihill E., Soley M. B., Geva E., Batista V. S. Tensor-train thermo-field memory kernels for generalized quantum master equations. *J. Chem. Theory Comput.*, **19** (4) (2023), 1111-1129.

[38] Liu Y. Y., Yan Y. M., Xu M., Song K., Shi Q. Exact generator and its high-order expansions in time-convolutionless generalized master equation: applications to spin-boson model and excitation energy transfer. *Chin. J. Chem. Phys.*, **31** (4) (2018), 575-583.

[39] Ullah A., Richardson J. O. Machine learning meets su(n) Lie algebra: enhancing quantum dynamics learning with exact trace conservation. *J. Chem. Phys.*, 162, 244106 (2025).

[40] Ullah A., Dral P. O. Speeding up quantum dissipative dynamics of open systems with kernel methods. *New J.*

Phys., **23** (2021).

[41] Ullah A., Dral P. O. Predicting the future of excitation energy transfer in light-harvesting complex with artificial-intelligence-based quantum dynamics. *Nat. Commun.*, **13** (2022), 1930.

[42] Ullah A., Dral P. O. One-shot trajectory learning of open quantum systems dynamics. *J. Phys. Chem. Lett.*, **13** (26) (2022), 6037-6041.

[43] Herrera Rodríguez L. E., Ullah A., Espinosa K. J. R., Dral P. O., Kananenka A. A. A comparative study of different machine-learning methods for dissipative quantum dynamics. *Mach. Learn. Sci. Technol.*, **3** (4) (2022), 045016.

[44] Herrera Rodríguez L. E., Kananenka A. A. Convolutional neural networks for long-time dissipative quantum dynamics. *J. Phys. Chem. Lett.*, **12** (9) (2021), 2476-2483.

[45] Ge F., Zhang L., Hou Y. F., Chen Y., Ullah A., Dral P. O. Four-dimensional-spacetime atomistic artificial intelligence models. *J. Phys. Chem. Lett.*, **14** (34) (2023), 7732-7743.

[46] Zhang L., Ullah A., Pinheiro M., Dral P. O., Barbatti M. Excited-state dynamics with machine learning. In: *Quantum Chemistry in the Age of Machine Learning*, Elsevier (2023), 329-353.

[47] Wu D., Hu Z., Li J., Sun X. Forecasting nonadiabatic dynamics using hybrid convolutional neural network/long short-term memory network. *J. Chem. Phys.*, **155** (22) (2021), 224104.

[48] Lin K., Peng J., Xu C., Gu F. L., Lan Z. Automatic evolution of machine-learning-based quantum dynamics with uncertainty analysis. *J. Chem. Theory Comput.*, **18** (10) (2022), 5837-5855.

[49] Bandyopadhyay S., Huang Z., Sun K., Zhao Y. Applications of neural networks to the simulation of dynamics of open quantum systems. *Chem. Phys.*, **515** (2018), 272-278.

[50] Yang B., He B., Wan J., Kubal S., Zhao Y. Applications of neural networks to dynamics simulation of Landau-Zener transitions. *Chem. Phys.*, **528** (2020), 110509.

[51] Lin K., Peng J., Xu C., Gu F. L., Lan Z. Trajectory propagation of symmetrical quasi-classical dynamics with Meyer–Miller mapping Hamiltonian using machine learning. *J. Phys. Chem. Lett.*, **13** (2022), 11678-11688.

[52] Tang D., Jia L., Shen L., Fang W. H. Fewest-switches surface hopping with long short-term memory networks. *J. Phys. Chem. Lett.*, **13** (44) (2022), 10377-10387.

[53] Shakiba M., Philips A. B., Autschbach J., Akimov A. V. Machine-learning mapping approach for computing spin relaxation dynamics. *J. Phys. Chem. Lett.*, in press (2024).

[54] Lin K., Gao X. Enhancing open quantum dynamics simulations using neural-network-based non-Markovian stochastic Schrödinger equation method. *J. Chem. Theory Comput.*, **21** (12) (2025), 5803-5814.

[55] Zeng H., Kou Y., Sun X. How sophisticated are neural networks needed to predict long-term nonadiabatic dynamics? *J. Chem. Theory Comput.*, **20** (22) (2024), 9832-9848.

[56] Long C., Cao L., Ge L., Li Q. X., Yan Y., Xu R. X., et al. Quantum neural network approach to Markovian dissipative dynamics of many-body open quantum systems. *J. Chem. Phys.*, **161** (8) (2024), 084105.

[57] Cao L., Ge L., Zhang D., Li X., Wang Y., Xu R. X., et al. Neural network approach for non-Markovian dissipative

dynamics of many-body open quantum systems. arXiv preprint, arXiv:2404.11093 (2024).

[58] Zhang J., Chen L. A non-Markovian neural quantum propagator and its application in the simulation of ultrafast nonlinear spectra. *Phys. Chem. Chem. Phys.*, **27** (2025), 182-189.

[59] Zhang J., Benavides-Riveros C. L., Chen L. Artificial-intelligence-based surrogate solution of dissipative quantum dynamics: physics-informed reconstruction of the universal propagator. *J. Phys. Chem. Lett.*, **15** (13) (2024), 3603-3610.

[60] Herrera Rodríguez L. E., Kananenka A. A. A short trajectory is all you need: a transformer-based model for long-time dissipative quantum dynamics. *J. Chem. Phys.*, **161** (17) (2024).

[61] Zhang J., Benavides-Riveros C. L., Chen L. Neural quantum propagators for driven-dissipative quantum dynamics. *Phys. Rev. Res.*, **7** (2025), L012013.

[62] Ullah A., Rodríguez L. E. H., Dral P. O., Kananenka A. A. QD3SET1: a database with quantum dissipative dynamics data sets. *Front. Phys.*, **11** (2023), 1223973.

[63] Am Busch M. S., Müh F., Madjet M. E. A., Renger T. The eighth bacteriochlorophyll completes the excitation-energy funnel in the FMO protein. *J. Phys. Chem. Lett.*, **2** (2) (2011), 93-98.

[64] Ishizaki A., Fleming G. R. Unified treatment of quantum coherent and incoherent hopping dynamics in electronic energy transfer: reduced hierarchy equation approach. *J. Chem. Phys.*, **130** (23) (2009), 234111.

[65] Shi Q., Chen L., Nan G., Xu R. X., Yan Y. Efficient hierarchical Liouville-space propagator to quantum dissipative dynamics. *J. Chem. Phys.*, **130** (8) (2009), 084105.

[66] Chen Z. H., Wang Y., Zheng X., Xu R. X., Yan Y. Universal time-domain Prony-fitting decomposition for optimized hierarchical quantum master equations. *J. Chem. Phys.*, **156** (2022), 221102.

[67] Mohseni M., Rebentrost P., Lloyd S., Aspuru-Guzik A. Environment-assisted quantum walks in photosynthetic energy transfer. *J. Chem. Phys.*, **129** (17) (2008), 11B603.

[68] Adolphs J., Renger T. How proteins trigger excitation-energy transfer in the FMO complex of green-sulfur bacteria. *Biophys.* J., **91** (8) (2006), 2778-2798.

[69] Dral P. O. MLatom: a program package for quantum-chemical research assisted by machine learning. *J. Comput. Chem.*, **40** (26) (2019), 2339-2347.