Predicting the future of excitation energy transfer in light-harvesting complex with artificial intelligence-based quantum dynamics

Arif Ullah* and Pavlo O. Dral*

State Key Laboratory of Physical Chemistry of Solid Surfaces, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Department of Chemistry, and College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China

E-mail: ua2024@xmu.edu.cn; dral@xmu.edu.cn

Abstract

Exploring excitation energy transfer (EET) in light-harvesting complexes (LHCs) is essential for understanding the natural processes and design of highly-efficient photovoltaic devices. LHCs are open systems, where quantum effects may play a crucial role for almost perfect utilization of solar energy. Simulation of energy transfer with inclusion of quantum effects can be done within the framework of dissipative quantum dynamics (QD), which are computationally expensive. Thus, artificial intelligence (AI) offers itself as a tool for reducing the computational cost. Here we suggest AI-QD approach using AI to directly predict QD as a function of time and other parameters such as temperature, reorganization energy, etc., completely circumventing the need of recursive step-wise dynamics propagation in contrast to the traditional QD and alternative, recursive AI-based QD approaches. Our trajectory-learning AI-QD approach is

able to predict the correct asymptotic behavior of QD at infinite time. We demonstrate AI-QD on seven-sites Fenna–Matthews–Olson (FMO) complex.

Introduction

From the birth of life, solar energy has been the driving force of life. Via the mechanism of photosynthesis, living organisms capture sunlight with the highly sophisticated pigments in their antenna systems and transfer sunlight energy to the reaction center (RC) in the form of electron-hole pairs (excitons), where it is stored as biochemical energy.[1] The transfer of solar energy from antenna to RC, which is also known as excitation energy transfer (EET), in the form of excitons is considered to be highly efficient with close to unit efficiency.[2] Understanding this high efficiency of the natural harvesting systems is very important because of its potential applications in designing very efficient organic solar cells and storage devices.[3] Experiments showed that the long-lasting coherence in the efficient natural light-harvesting harvesting complexes (LHCs) is preserved by the surrounding protein environments (scaffold), and this coherence may be responsible for this high efficiency.[4,5] The most well-investigated LHC is Fenna–Matthews–Olsen (FMO) complex, which is found in green sulfur bacteria.[6] The small size and simplicity of the FMO complex also makes it a testbed of simulation approaches. The FMO complex is a trimer of identical subunits, where each subunit consists of bacteriochlorophyll (BChl) molecules (system) attached to their protein environments.[7]

Enormous amount of research work has been done on light harvesting processes.[8–13] Taking FMO as an example, it is easy to see that the system (BChl molecules) is not isolated from the environment (the protein) and thus, the correct simulation of FMO should treat it as an open system rather than isolated one. In addition, many experiments suggest,[14,15] that quantum effects, particularly coherence, might play an important role in the light harvesting processes and may even be responsible for achieving the high-end efficiency. Temporal and

spatial simulation of EET with the inclusion of quantum effects can be done within many frameworks such as classical mapping-based approaches,[16–18] perturbative methods,[19–21] and dissipative quantum dynamics (QD)[22–27] adopted here.

QD simulations can be performed using the hierarchical equations of motion (HEOM)[28] and its many improvements and extensions,[8,23,29–31] the quasiadiabatic propagator path integral (QuAPI)[32] and its variant iterative QuAPI (iQuAPI),[27] the trajectory-based stochastic equation of motion (SEOM) approach,[25,33–39] the multi-layer multi-configuration time-dependent Hartree (ML-MCTDH)[26] and the local thermalising Lindblad master equation (LTLME).[22] The development of various quantum dissipative dynamics methods stirs from the fact that each of these methods has some limitations and hence there is no single universal method that works in all cases. For instance, HEOM is numerically exact but come with a very high computational cost at low temperatures, the SEOM has no explicit dependence on the temperature but has very bad convergence at long-time propagation, in the QuAPI approach all correlation effects are included over a finite time and correlation effects beyond this time are neglected. Most importantly, all these traditional QD approaches require step-wise propagation of trajectories and the next step depends on the previous steps, thus, QD simulation is an iterative, recursive process. Both calculations at each time step and recursive nature of QD makes it rather computationally expensive.

Alleviating the computational cost of QD became a target of a series of studies applying artificial intelligence (AI),[40–46] inspired by advances in application of AI employing machine learning (ML) algorithms in computational chemistry and chemical physics.[47,48] AI was also applied to investigate EET in a dimer system[44] and the FMO complex.[40] Saving of computational cost by AI in above studies is impressive, however, one of the studies[40] only focused on predicting energy transfer times and transfer efficiencies rather than temporal and spatial evolution, while other related studies[44–46] adopted basically the same recursive nature of QD trajectory propagation.

The recursive nature of the previous AI-based QD makes it prone to error accumulation.

In recursive simulations, previously predicted values are used as an input to predict the next value. Thus, the prediction error at each time-step will accumulate, which results in deterioration of accuracy. In addition, the recursive nature of predictions does not allow us to make a prediction for any arbitrary time without predicting values before that. Finally, a short-time trajectory is needed as the seed to be generated with traditional approaches such as HEOM and then provided as an input to AI model to make prediction for the next time step and ultimately propagate the long-time dynamics. Thus, even when having AI model, we still need to spend valuable computational time to generate the short-time trajectory with the traditional approaches.

Here, we suggest an AI-QD approach to directly predict QD with AI as a function of time and other parameters such as temperature, reorganization energy, etc., completely circumventing the need of recursive step-wise dynamics propagation in contrast to the traditional QD and alternative, recursive AI-based QD approaches. Our AI-QD approach is able to predict QD at infinite time with correct asymptotic behavior and can be viewed as trajectory learning, which does not need any short-time trajectory as an input, eradicates the need of traditional approaches to generate the seed, and alleviates the problem of error accumulation. We demonstrate the applicability of AI-QD on seven-sites Fenna–Matthews–Olson (FMO) complex and show how AI-QD can be used for massive, infinite-time QD simulations and provide insights into the desired range of parameters and more efficient paths followed by the transfer of excitation energy.

Results

Reference quantum dynamics of the FMO complex

We employ the Frenkel exciton Hamiltonian[49] to study EET dynamics in the FMO complex:

$$\boldsymbol{H} = \boldsymbol{H}_{\mathrm{s}} + \boldsymbol{H}_{\mathrm{env}} + \boldsymbol{H}_{\mathrm{s-env}} + \boldsymbol{H}_{\mathrm{reorg}}\,, \tag{1}$$

with all Hamiltonian terms given below

$$\boldsymbol{H}_{\mathrm{s}} = \sum_{i}^{n} |i\rangle \epsilon_i \langle i| + \sum_{i,j=1, i\neq j}^{n} |i\rangle J_{ij} \langle j| \,, \tag{2}$$

$$\boldsymbol{H}_{\mathrm{env}} = \sum_{i=1}^{n} \sum_{k=1} \left(\frac{1}{2} \boldsymbol{P}_{k,i}^2 + \frac{1}{2} \omega_{k,i}^2 \boldsymbol{Q}_{k,i}^2 \right) , \tag{3}$$

$$\boldsymbol{H}_{\mathrm{s-env}} = - \sum_{i=1}^{n} \sum_{k=1} |i\rangle c_{k,i} \boldsymbol{Q}_{k,i} \langle i| \,, \tag{4}$$

$$\boldsymbol{H}_{\mathrm{reorg}} = \sum_{i=1}^{n} |i\rangle \lambda_i \langle i| \,, \tag{5}$$

where $\boldsymbol{H}_{\mathrm{s}}$, $\boldsymbol{H}_{\mathrm{env}}$, $\boldsymbol{H}_{\mathrm{s-env}}$ and $\boldsymbol{H}_{\mathrm{reorg}}$ denote system (BChl molecules) Hamiltonian, Hamiltonian of protein-environment, system-environment interaction Hamiltonian and the reorganization term, respectively. In Eq. (1), n is the number of sites (BChl molecules), ϵ_i is the energy of the ith site and J_{ij} is the inter-site coupling between sites i and j. $\boldsymbol{P}_{k,i}$ and $\boldsymbol{Q}_{k,i}$ are respectively momentum coordinate and frequency of environment mode k associated with site i. In $\boldsymbol{H}_{\mathrm{s-env}}$, each site is connected to its own environment. The $c_{k,i}$ is the strength of coupling between site i and mode k of its environment. The reorganization term $\boldsymbol{H}_{\mathrm{reorg}}$ can be seen as a counter term that emerges from the interaction of the sites with the environment.[8,49,50] It is added to stop further renormalization of the site energy ϵ_i by the environment. In the reorganization term $\boldsymbol{H}_{\mathrm{reorg}}$, λ_i is the reorganization energy corresponding to site i,[51]

$$\lambda_i = \frac{1}{\pi} \int_0^{\infty} \frac{J_i(\omega)}{\omega} d\omega \,, \tag{6}$$

where $J_i(\omega)$ is spectral density of the environment corresponding to site i. As shown by Nalbach and Thorwart,[52] the effects of the discrete molecular modes on the population dynamics are largely irrelevant. As a result, it is acceptable to use continues environment spectral density such as Drude–Lorentz spectral density

$$J_{\mathrm{env}}(\omega) = 2\lambda \frac{\omega\gamma}{\omega^2 + \gamma^2}, \tag{7}$$

where γ and λ denote the characteristic frequency (bath relaxation rate) and the reorganization energy, respectively.

In general terms, the EET dynamics in the FMO complex can be described by Liouville–von Neumann equation

$$\frac{d}{dt}\boldsymbol{\rho}(t) = \frac{i}{\hbar}\left[\boldsymbol{H}, \boldsymbol{\rho}(t)\right] , \tag{8}$$

where $\boldsymbol{\rho}$ is the density matrix. Because of the many-body effects, direct propagation of Eq. (8) is not straightforward. Different approaches are developed to simplify and propagate Eq. (8) and interested readers are advised to look into the corresponding references.[25,30,32,53]

We use the local thermalising Lindblad master equation (LTLME)[22] to propagate the reference QD trajectories for the reduced density matrix of the system (see the Supplementary Information), where we adopt Adolphs and Renger's Hamiltonian for seven sites per subunit[54] (see Methods). The LTLME is a coherent and complete positive trace-preserving approach, but may not be that accurate as HEOM or SEOM approaches (because of approximations used in its derivation[22,55]), but here it is not the concern of our proof-of-concept paper.

Parameters-based non-recursive training framework

In our parameters-based non-recursive AI-QD, we train ML model as a function of a parameter space $\mathcal{D}$ (used as the input to ML model) which depends on the system of interest and on the data from a limited number of QD trajectories. For the FMO complex, our parameter space $\mathcal{D}$ consists of information of sites, λ, γ and T. In addition, time also becomes a part of the input of our AI-QD model. In order to treat infinite time, instead of time, we introduce time-function $f(t) \in \mathcal{D}$, which normalizes time and for $t \to \infty$ becomes $f(t) = 1$. Such normalization, however, can effectively only discern data within rather short time-region, thus, instead of a single time-function, we introduce the set of redundant time-functions $\{f_k(t)\}$ for different regions in very long-time propagation (see Methods). The remaining input of

m_1	n_1	γ_1	λ_1	T_1	$\{f_k(t_0)\}$	$\rho_{11}(t_0), \rho_{12}(t_0), \ldots, \rho_{17}(t_0)$
m_1	n_1	γ_1	λ_1	T_1	$\{f_k(t_1)\}$	$\rho_{11}(t_1), \rho_{12}(t_1), \ldots, \rho_{17}(t_1)$
m_1	n_1	γ_1	λ_1	T_1	$\{f_k(t_2)\}$	$\rho_{11}(t_2), \rho_{12}(t_2), \ldots, \rho_{17}(t_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m_1	n_1	γ_1	λ_1	T_1	$\{f_k(t_M)\}$	$\rho_{11}(t_M), \rho_{12}(t_M), \ldots, \rho_{17}(t_M)$
m_1	n_2	γ_1	λ_1	T_1	$\{f_k(t_0)\}$	$\rho_{21}(t_0), \rho_{22}(t_0), \ldots, \rho_{27}(t_0)$
m_1	n_2	γ_1	λ_1	T_1	$\{f_k(t_1)\}$	$\rho_{21}(t_1), \rho_{22}(t_1), \ldots, \rho_{27}(t_1)$
m_1	n_2	γ_1	λ_1	T_1	$\{f_k(t_2)\}$	$\rho_{21}(t_2), \rho_{22}(t_2), \ldots, \rho_{27}(t_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m_1	n_2	γ_1	λ_1	T_1	$\{f_k(t_M)\}$	$\rho_{21}(t_M), \rho_{22}(t_M), \ldots, \rho_{27}(t_M)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Input						Target values

Figure 1: Preparation of training data using parameters in AI-QD training framework. Here $\{f_k(t)\}$ is a set of time-functions based on the logistic function $f_k(t) = 1/(1 + 15 \cdot \exp(-(t + c_k)))$ where $c_k = 5k - 1.0$ and $k \in \{0, 1, 2, \ldots, 99\}$ (see Methods). Other parameters are $t = \{t_0, t_1, t_2, \ldots t_M\}$, $\lambda = \{\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_i\}$, $\gamma = \{\gamma_1, \gamma_2, \gamma_3, \ldots \gamma_j\}$, and $T = \{T_1, T_2, T_3, \ldots T_l\}$. In addition, labels $n = \{n_1, n_2, n_3, \ldots n_7\}$ are used for corresponding rows in the density matrix and labels for sites with possible initial excitation are $m = \{m_1, m_2\}$. As the off-diagonal elements $\rho_{pq,\, p \neq q}$ are complex, we separate the real and imaginary parts.

our model is information about the initial excitation $m = \{m_1, m_2\} = \{0, 1\} \in \mathcal{D}$ (with zero corresponding to initial excitation on site-1 and 1 corresponding to site-6) and labels $n = \{n_1, n_2, n_3, \ldots n_7\} = \{0.1, 0.2, 0.3, \ldots, 0.7\} \in \mathcal{D}$ corresponding to the seven rows in the reduced density matrix. We train convolutional neural network (CNN) taking all above input elements $\{m,\, n,\, \gamma,\, \lambda,\, T,\, f(t)\} \in \mathcal{D}$ on exciton population $\rho_{nn}(t)$ (target values to learn or output of the trained model) on the site defined by input n (see Fig. 1 and Methods for details, such as CNN architecture and normalization of input elements).

Our training trajectories generated with the reference LTLME-QD approach are chosen by furthest-point sampling from the three-dimensional space of the following parameters: reorganization energy $\lambda = \{\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_i\}$, the characteristic frequency $\gamma = \{\gamma_1, \gamma_2, \gamma_3, \ldots \gamma_j\}$ and temperature $T = \{T_1, T_2, T_3, \ldots T_l\}$ (see Methods).

We should also decide up to what time-length t_M we should run reference LTLME-QD trajectories. Based on the prior knowledge that populations plateau in asymptotic limit, for each trajectory we choose a different time-length t_M using a vanishing gradient scheme, where t_M is chosen such that the gradient of population G is close to zero (see Methods). Using the vanishing gradient scheme to find different t_M for each trajectory allows us to sample more data from the training trajectories, which are hard-to-learn, while avoiding redundant sampling from trajectories, which are easy-to-learn. This also removes arbitrariness in choosing fixed t_M parameter as was done in previous studies using the recursive AI-QD scheme.[44,46]

Application to EET dynamics in FMO complex

As an application of our approach, we predict EET dynamics in the FMO complex with seven sites per subunit for parameters of the test set trajectories (none of which used in training). Site-1 (BChl molecule 1) and site-6 (BChl molecule 6) are most likely to get initially excited as they are close to the photosynthetic antenna complex called chlorosome,[6] we thus present results for both cases. For predictions, we just provide the parameters of the test trajectories (characteristic frequency, reorganization energy, temperature) as an input and predict the evolution of EET. Fig. 2 shows the evolution of excitation energy in all seven sites for both cases. In Fig. 2, we show EET for both short and long time periods, demonstrating that AI-QD is able to capture the coherent EET (aka quantum beating or modulation of amplitudes) of short-time dynamics and also can predict the asymptotic limit. Fig. 3 shows the prominent off-diagonal terms (aka coherence) of the reduced density matrix for Fig. 2. Table 1 shows mean absolute error (MAE) and root mean square error (RMSE) averaged over 600 trajectories. As AI-QD is non-recursive (non-iterative), without any trajectory propagation, we can directly predict the asymptotic behaviour. Our AI-QD performs well in all cases (from weak coherence to strong coherence, from Markovian to non-Markovian, from adiabatic to nonadiabatic situations) as can be observed for selected trajectories shown in

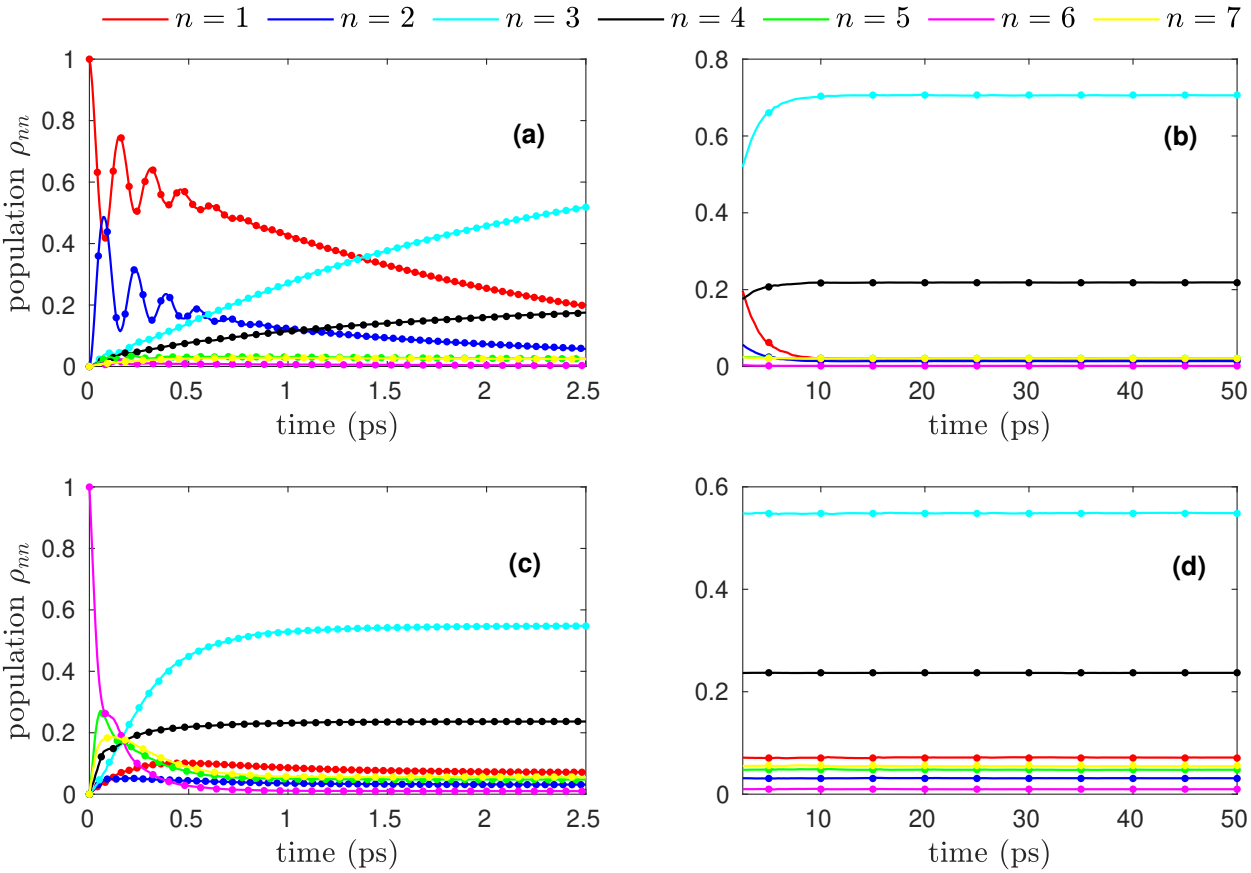

Figure 2: Population of the seven sites in the FMO complex as a function of time. In (a)-(b) the initial excitation is considered on site-1 and other parameters are $\gamma = 175$, $\lambda = 70$, $T = 70$. In (c)-(d), the initial excitation is on site-6 and other parameters are $\gamma = 75$, $\lambda = 100$, $T = 130$. (a) and (c) show a part of the population up to 2.5 ps, while the population changes beyond 2.5 ps are shown in (b) and (d), from which it is clearly seen that the population reaches plateau after a few picoseconds. The off-diagonal terms or coherences are shown in Fig. 3. The results of AI-QD are compared to the results of LTLME-QD (dots). γ and λ are in the units of cm^{-1}, while T is in the units of K.

Fig. S1 with corresponding errors reported in Table S1 of the Supplementary Information. From Table S1, we observe that our AI-QD approach is comparatively more accurate in strongly coherent cases (large value of γ and small values of λ and T) which can be seen as a consequence of the vanishing gradient scheme which may favor these challenging cases due to a larger number of training points sampled from such trajectories. AI-QD approach can even extrapolate to a good degree as its error for the test trajectories propagated with parameters outside the training parameter space is of a similar order of magnitude to the test

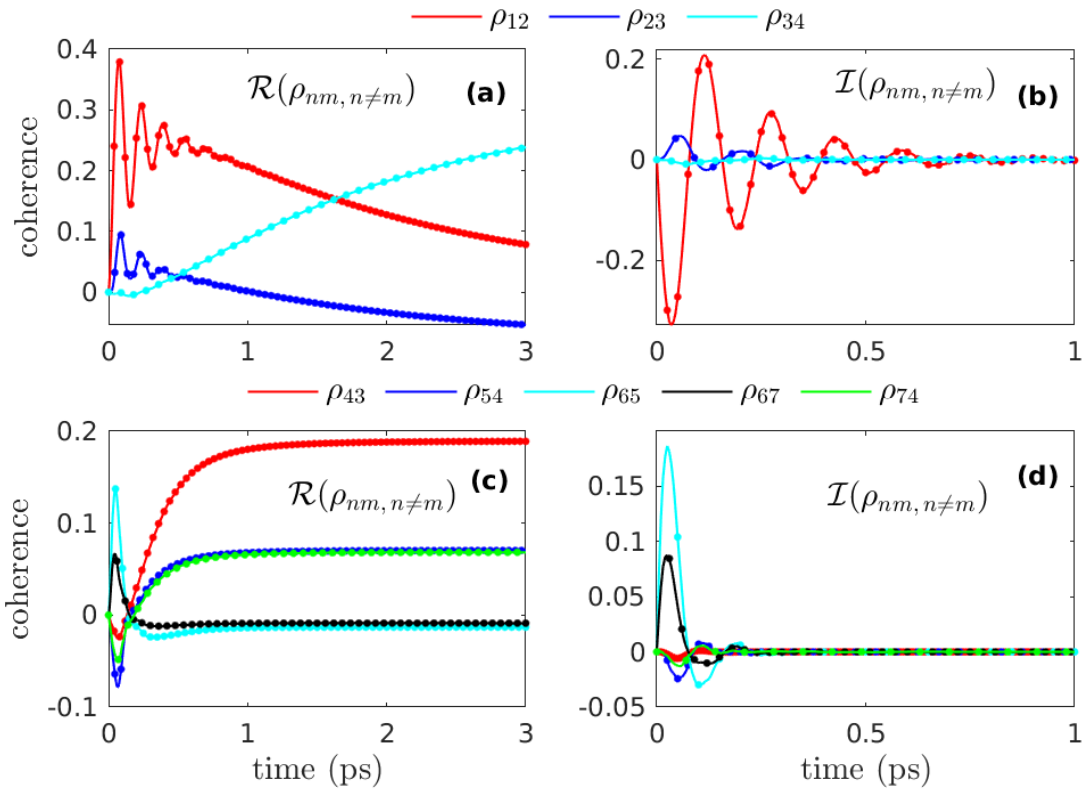

Figure 3: Electronic coherence as a function of time. (a) and (b) respectively show the real and imaginary parts of the prominent off-diagonal terms for Fig. 2(a)-(b) where $\gamma = 75$, $\lambda = 100$, $T = 130$ with the initial excitation on site-1. (c) and (d) respectively show the real and imaginary part of the prominent off-diagonal terms for Fig. 2(c)-(d) where $\gamma = 75$, $\lambda = 100$, $T = 130$ with the initial excitation on site-6. The results of AI-QD are compared to the results of LTLME-QD (dots). γ and λ are in unit of cm^{-1}, while T is in the units of K.

trajectories propagated with parameters inside the training parameter space (interpolation) as shown in the Supplementary Information (see additional set of test trajectories in Fig. S2 and Table S2).

It was shown,[8,56,57] that the transfer of excitation energy in the seven-sites FMO complex follows mainly two paths, i.e., site-1 $\rightarrow$ site-2 $\rightarrow$ site-3 $\leftrightarrow$ site-4 and site-6 $\rightarrow$ site-5, site-7, site-4 $\rightarrow$ site-3, here the $\leftrightarrow$ shows that the excitation energy equilibrates between site-3 and site-4 after site-3 is populated (see Fig. 3) Among the seven sites, the sites 1 and 6 are close to the baseplate protein, while the sites 3 and 4 are near to the target RC complex.[54,58] It has been proposed that the quantum coherence allows the FMO complex to quickly sample

Table 1: Mean absolute error (MAE) and root mean square error (RMSE) averaged over 600 randomly generated test trajectories propagated up to 1 ns. $\mathcal{R}\{\rho_{mn,n\neq m}\}$ and $\mathcal{I}\{\rho_{mn,n\neq m}\}$ represent the real and imaginary part of the off-diagonal terms, respectively.

	diagonal terms	off-diagonal terms	
Error	ρ_{nn}	$\mathcal{R}\{\rho_{mn,n\neq m}\}$	$\mathcal{I}\{\rho_{mn,n\neq m}\}$
MAE	$1.3{\cdot}10^{-3}$	$5.1{\cdot}10^{-4}$	$2.4{\cdot}10^{-4}$
RMSE	$2.1{\cdot}10^{-3}$	$8.1{\cdot}10^{-4}$	$3.6{\cdot}10^{-4}$

several routes (paths) in search for site-3.[5] In Fig. 4, we show the population of site-3 at t = 0.5 ps (500 fs) as a function of γ, λ and T. From Fig. 4(a), we observe that at room temperature $T = 300$, the ETT to site-3 or, in other words, to RC complex gets slow as the characteristic frequency γ increases. In contrast, the ETT to site-3 increases with the increase in reorganization energy λ as shown in Fig. 4(b). Similar trend can be observed with the increase in temperature T as can be seen in Fig. 4(c).

In order to find the optimum parameters for the fastest transfer of excitation energy, we have calculated population of site-3 at 0.5 ps (500 fs) for a massive set of ca. 0.57 million possible combinations (site-1 + site-6) of the γ, λ, T with the search space $\gamma = 25, 30, 35, \ldots, 245$, $\lambda = 10, 15, 20, \ldots, 345$ and $T = 25, 30, 35, \ldots, 345$. We report the fastest EET of 0.761 to site-3 for path-2 with $\gamma = 30$, $\lambda = 310$, $T = 25$, while for path-1 for the same parameters EET is 0.626. From Figs. 2, 4 and from the optimum parameters, we notice that following path-1, i.e., site-1 $\rightarrow$ site-2 $\rightarrow$ site-3 $\leftrightarrow$ site-4, the EET shows more coherence and is slow compared to excitation transfer following path-2, i.e., site-6 $\rightarrow$ site-5, site-7, site-4 $\rightarrow$ site-3. From Eq. (9) (Methods), energy of the site-1 (12410 cm^{-1}) is lower than the baseplate, which has been reported to be 12500 cm^{-1}.[59,60] This allows a quick transfer of the excitation energy to site-1 from the baseplate. However, the energy of site-2 (12530 cm^{-1}) is higher than site-1 and also than site-3 (12210 cm^{-1}), which on the one hand stops backward transfer from site-3, but on the other hand creates a local minimum on site-1. Despite the local minimum on site-1, the excitation energy is not trapped because of the quantum coherent wave-like motion between site-1 and site-2. Following path-2, the energy of site-6 (12630

cm^{-1}) is higher than the energy of baseplate. To stop backward transfer of excitation energy from site-6 to baseplate, site-6 should quickly transfer excitation energy to other sites such as site-5, site-7 and site-4. This quick transfer from site-6 to site-5, site-7 and site-4 is only possible by the strong coupling of site-6 to site-5 and site-7, which in return are strongly coupled to site-4.

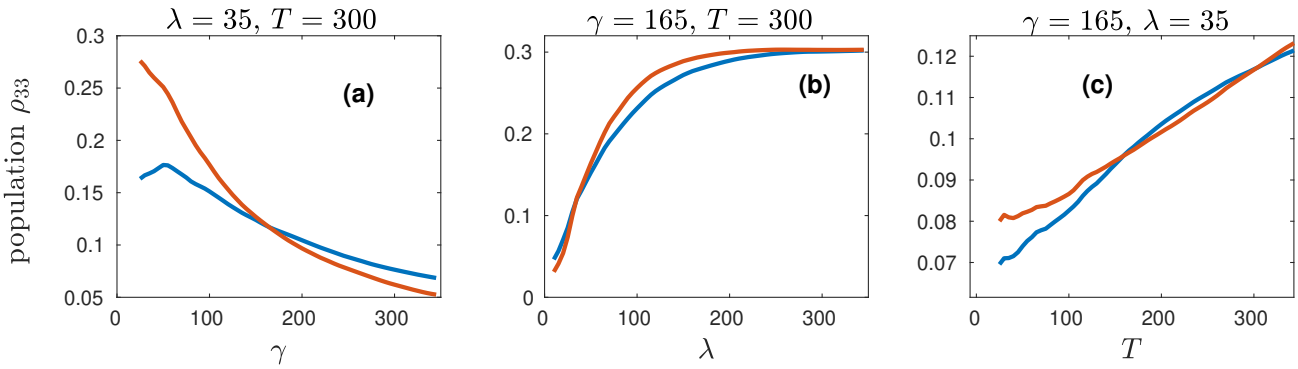

Figure 4: The evolution of site-3 population at t = 0.5 ps (500 fs) as a function of (a) characteristic frequency of the environment γ (b) reorganization energy λ and (c) temperature T. The blue line corresponds to the case with initial excition on site-1 while the red line is for the case with initial excition on site-6. γ and λ are in the units of cm^{-1} while T is in the units of K.

Discussion

In this work, we have presented a non-recursive (non-iterative) AI-QD approach for blazingly fast prediction of quantum dynamics, as predictions can be made for any time step up to asymptotic limit completely circumventing the need of recursive trajectory propagation. This can be used, as we demonstrated here, for massive quantum dynamics simulations, for example, in search for the best conditions required for efficient energy transfer in designed photovoltaic devices. Just to put things into perspective, our AI-QD approach can predict the entire 2.5 ps trajectory within ca. 2 min on a single core of Intel(R) Core(TM) i7-10700 CPUs @ 2.90 GHz, independent of the reference method used for generating training trajectories, while the same propagation with the traditional recursive approaches such as HEOM would

take hours, and the cost would exponentially increase for low temperatures. The high cost of accurate approaches such as HEOM was also a reason why we used a much faster LTLME for this proof-of-concept study to extensively test our approach (propagation of an entire trajectory takes only 3 min with LTLME on a single CPU of the above computer architecture. It is worth emphasizing that AI-QD is embarrassingly parallel and the calculations can be further significantly sped up by using multiple CPUs or GPUs, because predictions with AI-QD for different time steps are independent of each other and different segments of trajectories can be distributed for independent calculations on many threads.

We demonstrated the feasibility of AI-QD approach on an example of the FMO complex, but this approach is general enough to be used for any other complex after retraining. It remains to be seen how well the AI-QD approach can be extended to describe several LHCs at the same time — a topic of our ongoing research. One could use the LHC' Hamiltonian elements as a representation of LHC complexes and an early encouraging study [42] has shown that by using Hamiltonian elements as input of an ML model, one can successfully describe scalar properties (energy transfer times and transfer efficiencies) for different Hamiltonians. However, open question remains how successful would be such an approach to learn dynamics and in addition, how to circumvent different dimensionalities of Hamiltonians of different complexes.

Methods

Training data

In the seven-sites FMO complex (apo-FMO), where seven BChl molecules (seven sites) exist per subunit, the inter-subunit interaction is very small and each subunit can be considered relatively isolated.[61] Here we adopt Adolphs and Renger's Hamiltonian for seven sites per

subunit[54]

$$\mathbf{H}_{\mathrm{s}} = \begin{bmatrix} 12410 & -87.7 & 5.5 & -5.9 & 6.7 & -13.7 & -9.9 \\ -87.7 & 12530 & 30.8 & 8.2 & 0.7 & 11.8 & 4.3 \\ 5.5 & 30.8 & 12210 & -53.5 & -2.2 & -9.6 & 6.0 \\ -5.9 & 8.2 & -53.5 & 12320 & -70.7 & -17.0 & -63.6 \\ 6.7 & 0.7 & -2.2 & -70.7 & 12480 & 81.1 & -1.3 \\ -13.7 & 11.8 & -9.6 & -17.0 & 81.1 & 12630 & 39.7 \\ -9.9 & 4.3 & 6.0 & -63.3 & -1.3 & 39.7 & 12440 \end{bmatrix}, \tag{9}$$

where energies are given in cm^{-1}. Each site is coupled to its own environment characterized by the Drude–Lorentz spectral density given by Eq. (7). Not long-ago, an eighth BChl molecule (site-8) has been discovered,[11] however as has been mentioned by Jia et al.,[62] the role of the eighth BChl molecule (site-8) in the transfer of excitation energy in the FMO complex is negligible.

Trajectories for the reduced density matrix have been generated with the local thermalising Lindblad master equation (LTLME)[22] (see the Supplementary Information) implemented in quantum_HEOM package[63] with QuTip[64] in the back-end with all the possible combinations of the following parameters: $\lambda = \{10, 40, 70, 100, 130, 160, 190, 220, 250, 280, 310\}$ cm^{-1}, $\gamma = \{25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300\}$ cm^{-1} and $T = \{30, 50, 70, 90, 110, 130, 150, 170, 190, 210, 230, 250, 270, 290, 310\}$ K. We consider that all these combinations of parameters make a parameter space $\mathcal{D}$. The time-step used for propagation is 5 fs and the trajectory is propagated up to $t_M = 1$ ns (10^6 fs). With the possibility of initial excitation on site-1 and site-6, we generate 1980 trajectories for each excitation case.

Data preparation

With all the possible combinations of the parameters ($\mathcal{D}$), we have 3960 total number of trajectories $\mathrm{N}_{\mathrm{traj}}$ (1980 (site-1) + 1980 (site-6), all these trajectories correspond to their respective combination of parameters in parameter space $\mathcal{D}$). Using farthest-point sampling[65]

in the three-dimensional space of λ, γ and T, we choose 1000 trajectories as our training space $\mathcal{TS}$ (500 (site-1) + 500 (site-6) , ca. 25% of space $\mathcal{D}$)), 200 trajectories as the validation set $\mathcal{VS}$ (ca. 5% of space $\mathcal{D}$)) and the rest of trajectories, we keep as the test set $\mathcal{STP}$ (ca. 70% of space $\mathcal{D}$). For each trajectory, we choose a different time-length t_M using a vanishing gradient scheme. In this scheme, we take the gradient G of the population of each site (ρ_{nn}, $n = 1, 2, 3, \ldots, 7$) for 10 consecutive time-steps and if all of them remain less than the threshold value of $G_{\text{th}} = 1 \times 10^{-10}$, we choose our t_M. We find t_M for all seven sites and then choose the maximum value among them, thus we keep a single value of asymptotic limit (t_M) for all seven-sites. By analyzing the gradients, we find the region of the trajectory, where the change in population of the site is very small. By knowing that, we keep the time-length of our trajectory t_M up to that region, because beyond t_M the change in population is very small, and ML is able to predict it. As the asymptotic limit for each trajectory is different, we have different value of t_M for each trajectory. In our training, we have included $t \to \infty$, corresponding to the asymptotic behaviour at long-time. Using the strategy of different t_M for each trajectory allows us to include more sampling in our training set from hard-to-learn trajectories, while avoiding redundant sampling from easy-to-learn trajectories. For training, sampling is done with different training time-steps Δt_{train} in different regions of the trajectory. We sample our training points from ||0ps–1ps||, ||1ps–1.5ps||, ||1.5ps–2.5ps||, ||2.5ps–5ps||, ||5ps–25ps||, ||25ps–50ps||, ||50ps–250ps|| ||250ps–t_M|| regions with $\Delta t_{\text{train}} = 5, 10, 25, 50, 100, 200, 500, 1000$ fs, respectively. The number of training points depends on the number of trajectories N_{traj} chosen for training, training time-step Δt_{train} and time-length of trajectories t_M, which in turn depends on G_{th}.

Training architecture

We use convolutional neural network (CNN) architecture, because the importance of convolutional layers is much explored for image analysis, where these layers extract important features such as edges, textures, objects, and scenes. When it comes to time-series data, we

Table 2: Summary of the optimized neural network architecture with layers, output shape (OS), number of parameters (NP), activation function (AF), number of filters (NF), kernel size (KS) and number of neurons (NN).

Layers (type)	OS	NP	AF	NF	KS	NN
First hidden convolutional layer (1D)	(None, 103, 90)	360	relu	90	3	×
Second hidden convolutional layer (1D)	(None, 103, 70)	18970	relu	70	3	×
Maximum pooling layer	(None, 51, 70)	0	×	×	×	×
Flatten layer	(None, 3570)	0	×	×	×	×
First hidden dense layer	(None, 512)	1828352	relu	×	×	512
Second hidden dense layer	(None, 512)	262656	relu	×	×	512
Third hidden dense layer	(None, 512)	262656	relu	×	×	512
Dense output layer	(None, 13)	6669	linear	×	×	13

Total parameters: 2,379,663
Trainable parameters: 2,379,663
Non-trainable parameters: 0

are using convolutional layers in the hope to extract some important features from the data (such as the time influence). After learning those features, when we provide a test trajectory, the trained ML model will look for those features in that test trajectory.[66] Though we have used the CNN model, other ANN architectures such as long short-term memory (LSTM) is also an option. LSTM is considered to be more suitable for extracting long-time temporal dependencies in contrast to convolutional neural networks (CNNs) which are more local. However, CNNs are easy to train and in many studies, they have outperformed LSTM for future forecasting.[67,68]

We use 1000 trajectories as our training set $\mathcal{TS}$ and 200 trajectories as the validation set $\mathcal{VS}$. After preparation of the input following Fig. 1, we build a CNN architecture and optimize it with hyperopt library.[69] The optimization was carried out only on 300 training

trajectories from the training set $\mathcal{TS}$. After optimization, our training architecture consists of two one-dimensional (1D) hidden convolutional layers, one maximum pooling layer, one flatten layer, three fully connected hidden dense layers and one output dense layer. The convolutional layers extract time-dependent correlations from a moving window, while maximum pooling layer pulls out the important information and decreases the size of the feature map which leads to reducing the computational cost. The flatten layer converts the output from the maximum pooling layer into 1D format as the fully connected dense layers, which are the traditional networks, can only work with 1D data. We train our CNN architecture using Keras software package[70] with the TensorFlow in the backend.[71] Activation function, number of filters, kernel size and number of neurons for the respective convolutional and dense layers are given in Table 2. In our study, we train a single CNN model and with ca. 3.2 million training points and 900 epochs, training takes ca. 42 hrs on 32 Intel(R) Xeon(R) Gold 6226R CPUs @ 2.90GHz. The optimized learning rate is 1×10^{-3} with adoptive mean optimizer and the batch size is 512. Using mean squared error function as a loss, we report 1.86×10^{-7} as the validation loss. The mean absolute error (MAE) and root mean square error (rmse) averaged over 600 randomly chosen test trajectories (which were not part of the training process) are given in Table 1.

Input normalization and redundant time-functions

As we have multiple input elements, we need to normalize them all. In normalized input, we have $\lambda = \{\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_j\}/\lambda_{\max}$, $\gamma = \{\gamma_1, \gamma_2, \gamma_3, \ldots \gamma_k\}/\gamma_{\max}$ and $T = \{T_1, T_2, T_3, \ldots T_l\}/T_{\max}$, where $\lambda_{\max}$, $\gamma_{\max}$ and $T_{\max}$ represent the maximum values of λ, γ and T, respectively. We divide $n = \{n_1, n_2, n_3, \ldots n_7\} = \{1, 2, 3, \ldots, 7\}$ (labels corresponding to the seven rows in the reduced density matrix) by 10 to normalize their values, i.e., the input elements corresponding to the rows in the reduced density matrix are $\{0.1, 0.2, 0.3, \ldots, 0.7\}$. Labels for sites with possible initial excitation are $m = \{0, 1\}$, which respectively represent initial excitation on site-1 and site-6. The input time is represented by a set of redundant time-functions $\{f_i(t)\}$,

each of which is logistic function $f(t)$ normalizing time. We use a set of 100 logistic functions $f_k(t) = 1/(1 + 15 \cdot \exp(-(t + c_k)))$, where $c_k = 5k - 1.0$ and $k \in \{0, 1, 2, \ldots, 99\}$, i.e., each logistic function has the same shape and designed to cover the corresponding ≈ 5 ps region and is shifted with respect to the next logistic function by 5 ps, as shown in Fig. S3 of the Supplementary Information. The infinity limit is given by all redundant time-functions set to one.

Acknowledgements

P.O.D. acknowledges funding by the National Natural Science Foundation of China (No. 22003051), the Fundamental Research Funds for the Central Universities (No. 20720210092), and via the Lab project of the State Key Laboratory of Physical Chemistry of Solid Surfaces.

Authors contribution

A.U. conceived the idea of investigating excitation energy transfer in the Fenna–Matthews–Olsen complex. P.O.D. conceived the idea of learning trajectories as a function of time and parameters. Both authors developed the method. A.U. did all the implementations, calculations, analysis of data, and wrote the original version of the manuscript. Both authors revised the manuscript.

Competing interests

The authors declare no competing interests.

References

(1) Fassioli, F.; Dinshaw, R.; Arpin, P. C.; Scholes, G. D. Photosynthetic light harvesting: excitons and coherence. *Journal of The Royal Society Interface* **2014**, *11*, 20130901.

(2) Blankenship, R. E. *Molecular mechanisms of photosynthesis*; John Wiley & Sons, 2021.

(3) Bhatia, S. *Advanced renewable energy systems,(Part 1 and 2)*; CRC Press, 2014.

(4) Olson, J. M. The FMO protein. *Discoveries in Photosynthesis* **2005**, 421–427.

(5) Engel, G. S.; Calhoun, T. R.; Read, E. L.; Ahn, T.-K.; Mančal, T.; Cheng, Y.-C.; Blankenship, R. E.; Fleming, G. R. Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature* **2007**, *446*, 782–786.

(6) Karafyllidis, I. G. Quantum transport in the FMO photosynthetic light-harvesting complex. *Journal of biological physics* **2017**, *43*, 239–245.

(7) Tronrud, D. E.; Wen, J.; Gay, L.; Blankenship, R. E. The structural basis for the difference in absorbance spectra for the FMO antenna protein from various green sulfur bacteria. *Photosynthesis research* **2009**, *100*, 79–87.

(8) Ishizaki, A.; Fleming, G. R. Theoretical examination of quantum coherence in a photosynthetic system at physiological temperature. *Proceedings of the National Academy of Sciences* **2009**, *106*, 17255–17260.

(9) Collini, E.; Wong, C. Y.; Wilk, K. E.; Curmi, P. M.; Brumer, P.; Scholes, G. D. Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature* **2010**, *463*, 644–647.

(10) Milder, M. T.; Brüggemann, B.; van Grondelle, R.; Herek, J. L. Revisiting the optical properties of the FMO protein. *Photosynthesis research* **2010**, *104*, 257–274.

(11) Schmidt am Busch, M.; Müh, F.; El-Amine Madjet, M.; Renger, T. The eighth bacteriochlorophyll completes the excitation energy funnel in the FMO protein. *The journal of physical chemistry letters* **2011**, *2*, 93–98.

(12) Olbrich, C.; Strümpfer, J.; Schulten, K.; Kleinekathöfer, U. Theory and simulation of the environmental effects on FMO electronic transitions. *The journal of physical chemistry letters* **2011**, *2*, 1771–1776.

(13) Chenu, A.; Scholes, G. D. Coherence in energy transfer and photosynthesis. *Annual review of physical chemistry* **2015**, *66*, 69–96.

(14) Cheng, Y.; Silbey, R. J. Coherence in the B800 ring of purple bacteria LH2. *Physical review letters* **2006**, *96*, 028103.

(15) Lee, H.; Cheng, Y.-C.; Fleming, G. R. Coherence dynamics in photosynthesis: protein protection of excitonic coherence. *Science* **2007**, *316*, 1462–1465.

(16) Cotton, S. J.; Miller, W. H. The symmetrical quasi-classical model for electronically non-adiabatic processes applied to energy transfer dynamics in site-exciton models of light-harvesting complexes. *Journal of chemical theory and computation* **2016**, *12*, 983–991.

(17) Mannouch, J. R.; Richardson, J. O. A partially linearized spin-mapping approach for nonadiabatic dynamics. I. Derivation of the theory. *The Journal of Chemical Physics* **2020**, *153*, 194109.

(18) Liu, J.; He, X.; Wu, B. Unified Formulation of Phase Space Mapping Approaches for Nonadiabatic Quantum Dynamics. *Accounts of chemical research* **2021**, *54*, 4215–4228.

(19) Hwang-Fu, Y.-H.; Chen, W.; Cheng, Y.-C. A coherent modified Redfield theory for excitation energy transfer in molecular aggregates. *Chemical Physics* **2015**, *447*, 46–53.

(20) Jang, S.; Cheng, Y.-C.; Reichman, D. R.; Eaves, J. D. Theory of coherent resonance energy transfer. *The Journal of chemical physics* **2008**, *129*, 101104.

(21) Wu, J.; Liu, F.; Shen, Y.; Cao, J.; Silbey, R. J. Efficient energy transfer in light-harvesting systems, I: optimal temperature, reorganization energy and spatial–temporal correlations. *New Journal of Physics* **2010**, *12*, 105012.

(22) Mohseni, M.; Rebentrost, P.; Lloyd, S.; Aspuru-Guzik, A. Environment-assisted quantum walks in photosynthetic energy transfer. *The Journal of chemical physics* **2008**, *129 17*, 174106.

(23) Wilkins, D. M.; Dattani, N. S. Why quantum coherence is not important in the Fenna–Matthews–Olsen complex. *Journal of chemical theory and computation* **2015**, *11*, 3411–3419.

(24) Strümpfer, J.; Schulten, K. Open quantum dynamics calculations with the hierarchy equations of motion on parallel computers. *Journal of chemical theory and computation* **2012**, *8*, 2808–2816.

(25) Imai, H.; Ohtsuki, Y.; Kono, H. Application of stochastic Liouville–von Neumann equation to electronic energy transfer in FMO complex. *Chemical Physics* **2015**, *446*, 134–141.

(26) Schulze, J.; Shibl, M. F.; Al-Marri, M. J.; Kühn, O. Multi-layer multi-configuration time-dependent Hartree (ML-MCTDH) approach to the correlated exciton-vibrational dynamics in the FMO complex. *The Journal of chemical physics* **2016**, *144*, 185101.

(27) Richter, M.; Fingerhut, B. P. Coarse-grained representation of the quasi adiabatic propagator path integral for the treatment of non-Markovian long-time bath memory. *The Journal of chemical physics* **2017**, *146*, 214101.

(28) Tanimura, Y.; Kubo, R. Time evolution of a quantum system in contact with a nearly Gaussian-Markoffian noise bath. *Journal of the Physical Society of Japan* **1989**, *58*, 101–114.

(29) Xu, R.-X.; Cui, P.; Li, X.-Q.; Mo, Y.; Yan, Y. Exact quantum master equation via the calculus on path integrals. *The Journal of chemical physics* **2005**, *122*, 041103.

(30) Tanimura, Y. Stochastic Liouville, Langevin, Fokker–Planck, and master equation approaches to quantum dissipative systems. *Journal of the Physical Society of Japan* **2006**, *75*, 082001.

(31) Zhang, H.-D.; Cui, L.; Gong, H.; Xu, R.-X.; Zheng, X.; Yan, Y. Hierarchical equations of motion method based on Fano spectrum decomposition for low temperature environments. *The Journal of chemical physics* **2020**, *152*, 064107.

(32) Makarov, D. E.; Makri, N. Path integrals for dissipative systems by tensor multiplication. Condensed phase quantum dynamics for arbitrarily long time. *Chemical physics letters* **1994**, *221*, 482–491.

(33) Stockburger, J. T.; Grabert, H. Non-Markovian quantum state diffusion. *Chemical Physics* **2001**, *268*, 249–256.

(34) Shao, J. Decoupling quantum dissipation interaction via stochastic fields. *The Journal of chemical physics* **2004**, *120*, 5053–5056.

(35) Ke, Y.; Zhao, Y. An extension of stochastic hierarchy equations of motion for the equilibrium correlation functions. *The Journal of chemical physics* **2017**, *146*, 214105.

(36) McCaul, G.; Lorenz, C.; Kantorovich, L. Partition-free approach to open quantum systems in harmonic environments: An exact stochastic Liouville equation. *Physical Review B* **2017**, *95*, 125124.

(37) Han, L.; Chernyak, V.; Yan, Y.-A.; Zheng, X.; Yan, Y. Stochastic Representation of Non-Markovian Fermionic Quantum Dissipation. *Physical review letters* **2019**, *123*, 050601.

(38) Han, L.; Ullah, A.; Yan, Y.-A.; Zheng, X.; Yan, Y.; Chernyak, V. Stochastic equation of motion approach to fermionic dissipative dynamics. I. Formalism. *The Journal of Chemical Physics* **2020**, *152*, 204105.

(39) 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. *The Journal of Chemical Physics* **2020**, *152*, 204106.

(40) Häse, F.; Kreisbeck, C.; Aspuru-Guzik, A. Machine learning for quantum dynamics: deep learning of excitation energy transfer properties. *Chem. Sci.* **2017**, *8*, 8419–8426.

(41) Hartmann, M. J.; Carleo, G. Neural-Network Approach to Dissipative Quantum Many-Body Dynamics. *Physical Review Letters* **2019**, *122*, 250502.

(42) Häse, F.; Roch, L. M.; Friederich, P.; Aspuru-Guzik, A. Designing and understanding light-harvesting devices with machine learning. *Nature Communications* **2020**, *11*, 4587.

(43) Secor, M.; Soudackov, A. V.; Hammes-Schiffer, S. Artificial Neural Networks as Propagators in Quantum Dynamics. *J. Phys. Chem. Lett.* **2021**, *12*, 10654−10662.

(44) Herrera Rodríguez, L. E.; Kananenka, A. A. Convolutional Neural Networks for Long Time Dissipative Quantum Dynamics. *The Journal of Physical Chemistry Letters* **2021**, *12*, 2476–2483.

(45) Lin, K.; Peng, J.; Gu, F. L.; Lan, Z. Simulation of Open Quantum Dynamics with Bootstrap-Based Long Short-Term Memory Recurrent Neural Network. *The Journal of Physical Chemistry Letters* **2021**, *12*, 10225.

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

(47) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K.-R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems. *Chemical Reviews* **2021**, *121*, 9816 – 9872.

(48) Shi, Y.; Prieto, P. L.; Zepel, T.; Grunert, S.; Hein, J. E. Automated experimentation powers data science in chemistry. *Accounts of Chemical Research* **2021**, *54*, 546–555.

(49) Ishizaki, A.; Fleming, G. R. Unified treatment of quantum coherent and incoherent hopping dynamics in electronic energy transfer: Reduced hierarchy equation approach. *The Journal of chemical physics* **2009**, *130*, 234111.

(50) Mühlbacher, L.; Kleinekathöfer, U. Preparational effects on the excitation energy transfer in the FMO complex. *The Journal of Physical Chemistry B* **2012**, *116*, 3900–3906.

(51) Zhong, X.; Zhao, Y. Charge carrier dynamics in phonon-induced fluctuation systems from time-dependent wavepacket diffusion approach. *The Journal of chemical physics* **2011**, *135*, 134110.

(52) Nalbach, P.; Thorwart, M. The role of discrete molecular modes in the coherent exciton dynamics in FMO. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2012**, *45*, 154009.

(53) Stock, G.; Thoss, M. Semiclassical description of nonadiabatic quantum dynamics. *Physical review letters* **1997**, *78*, 578.

(54) Adolphs, J.; Renger, T. How proteins trigger excitation energy transfer in the FMO complex of green sulfur bacteria. *Biophysical journal* **2006**, *91*, 2778–2797.

(55) Worster, S. B.; Stross, C.; Vaughan, F.; Linden, N.; Manby, F. Structure and Efficiency

in Bacterial Photosynthetic Light-Harvesting. *The journal of physical chemistry letters* **2019**, *10*, 7383−7390.

(56) Brixner, T.; Stenger, J.; Vaswani, H. M.; Cho, M.; Blankenship, R. E.; Fleming, G. R. Two-dimensional spectroscopy of electronic couplings in photosynthesis. *Nature* **2005**, *434*, 625–628.

(57) Cho, M.; Vaswani, H. M.; Brixner, T.; Stenger, J.; Fleming, G. R. Exciton Analysis in 2D Electronic Spectroscopy. *The journal of physical chemistry. B* **2005**, *109*, 10542–10556.

(58) Wen, J.; Zhang, H.; Gross, M. L.; Blankenship, R. E. Membrane orientation of the FMO antenna protein from Chlorobaculum tepidum as determined by mass spectrometry-based footprinting. *Proceedings of the National Academy of Sciences* **2009**, *106*, 6134–6139.

(59) Francke, C.; Amesz, J. Isolation and pigment composition of the antenna system of four species of green sulfur bacteria. *Photosynthesis research* **1997**, *52*, 137–146.

(60) Frigaard, N.-U.; Li, H.; Martinsson, P.; Das, S. K.; Frank, H. A.; Aartsma, T. J.; Bryant, D. A. Isolation and characterization of carotenosomes from a bacteriochlorophyll c-less mutant of Chlorobium tepidum. *Photosynthesis research* **2005**, *86*, 101–111.

(61) Ke, Y.; Zhao, Y. Hierarchy of forward-backward stochastic Schrödinger equation. *The Journal of chemical physics* **2016**, *145*, 024101.

(62) Jia, X.; Mei, Y.; Zhang, J. Z.; Mo, Y. Hybrid QM/MM study of FMO complex with polarized protein-specific charge. *Scientific reports* **2015**, *5*, 1–10.

(63) Abbott, J. W. quantum_HEOM. `https://github.com/jwa7/quantum_HEOM`, (accessed on Nov 22, 2021), 2019.

(64) Johansson, J. R.; Nation, P. D.; Nori, F. QuTiP: An open-source Python framework for the dynamics of open quantum systems. *Computer Physics Communications* **2012**, *183*, 1760–1772.

(65) Dral, P. O. MLatom: A program package for quantum chemical research assisted by machine learning. *Journal of computational chemistry* **2019**, *40*, 2339–2347.

(66) Ismail Fawaz, H.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.-A. Deep learning for time series classification: a review. *Data Min. Know. Disc.* **2019**, *33*, 917–963.

(67) Tran, D.; Bourdev, L.; Fergus, R.; Torresani, L.; Paluri, M. Learning Spatiotemporal Features with 3D Convolutional Networks. IEEE International Conference on Computer Vision (ICCV). 2015; pp 4489–4497.

(68) Lea, C.; Flynn, M.; Vidal, R.; Reiter, A.; Hager, G. Temporal convolutional networks for action segmentation and detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017; p 156–165.

(69) Bergstra, J.; Komer, B.; Eliasmith, C.; Yamins, D. L. K.; Cox, D. D. Hyperopt: a Python library for model selection and hyperparameter optimization. *Computational Science & Discovery* **2015**, *8*, 014008.

(70) Keras, a deep learning API. `https://keras.io`, 2014.

(71) Abadi, M. et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. `http://tensorflow.org/`, 2016.

Supplementary Information

Predicting the future of excitation energy transfer in light-harvesting complex with artificial intelligence-based quantum dynamics

A. Ullah* and P. O. Dral*

E-mail: ua2024@xmu.edu.cn; dral@xmu.edu.cn

Lindblad equation of motion for EET in FMO complex

The theory of local thermalising Lindblad equation of motion is well-documented in Ref. 1, however for the sake of completeness, we briefly outline it here. We begin from system Hamiltonian given in Eq. (2) of the main text, and rewrite it as

$$\boldsymbol{H}_{\mathrm{s}} = \sum_{i=1}^{n} \epsilon_i a_i^{\dagger} a_i + \sum_{j<i}^{n} J_{ij} \left(a_i^{\dagger} a_j + a_j^{\dagger} a_i \right) . \tag{1}$$

The $a_i^{\dagger}$ and a_i are the exciton creation and annihilation operators at site-i. Because of the fast EET, only single-excitation is considered, hence we have two possible states $|0\rangle$ and $|i\rangle = a_i^{\dagger}|0\rangle$. The exciton basis $|e\rangle = \sum_i c_i(e)|i\rangle$ is considered as eigenbasis of the Hamiltonian $\boldsymbol{H}_{\mathrm{s}}$, i.e., $\boldsymbol{H}_{\mathrm{s}}|e\rangle = \epsilon_i|e\rangle$. With the Born-Markov and secular approximations, the Lindblad master equation (to second order in the system-bath coupling) for the dynamics of the

reduced density matrix is written as

$$\frac{\partial \rho_{\rm s}(t)}{\partial t} = -\frac{i}{\hbar}\left[\boldsymbol{H}_{\rm s} + \boldsymbol{H}_{\rm reorg}, \rho_{\rm s}(t)\right] + L_{\rm env}(\rho_{\rm s}(t)) + L_{\rm s-env}(\rho_{\rm s}(t)), \tag{2}$$

where the Lindblad superoperators $L_{\rm env}$ and $L_{\rm s-env}$ are given by ($k = {\rm env},\, {\rm s-env}$)

$$L_k(\rho_{\rm s}) = \sum_{\omega}\sum_{i,j}\kappa_{ij}^{k}(\omega)\left[A_i^k(\omega)\rho_{\rm s}A_j^{k\dagger}(\omega) - \frac{1}{2}A_i^k(\omega)A_j^{k\dagger}(\omega)\rho - \frac{1}{2}\rho_{\rm s}A_i^k(\omega)A_j^{k\dagger}(\omega)\right].$$

In local thermalising approach, the corresponding Lindblad generators for $\boldsymbol{H}_{\rm env}$ are

$$A_i^{\rm env}(\omega) = \sum_{\omega-\omega'} c_i^*\left(e_\omega\right) c_i\left(e_{\omega'}\right) \left|e_\omega\right\rangle \left\langle e_{\omega'}\right|, \tag{3}$$

where the summation is over all transitions between eigenstates $|e_\omega\rangle$ and $|e_{\omega'}\rangle$ with frequency ω. The $c_i^*\left(e_\omega\right) c_i\left(e_{\omega'}\right)$ is a factor weighting transfer between the two eigen states $|e_\omega\rangle$ and $|e_{\omega'}\rangle$. The $c_i\left(e_\omega\right)$ and $c_i\left(e_{\omega'}\right)$ are the ith site coefficient in eigenstate $|e_\omega\rangle$ and $|e_{\omega'}\rangle$, respectively. The rate κ in Eq.(2) is considered site-independent with the following expression $\gamma^{\rm env} = 2\pi\left[J(\omega)(1+n(\omega)) + J(-\omega)n(-\omega)\right]$ where $J(\omega)$ is spectral density and $n(\omega) = 1/\left[\exp\left(\frac{\hbar\omega}{k_{\rm B}T}\right) - 1\right]$ is bosonic distribution.

For $\boldsymbol{H}_{\rm s-env}$, $A_i^{\rm s-env}(\omega_e) = c_i(e)|0\rangle\langle e|$ where $\hbar\omega_e$ is the molecular transition frequency. The respective rate $\kappa_{ij}^{\rm s-env}(\omega)$ is diagonal $\kappa_{ij}^{\rm s-env}(\omega) = \delta_{ij}\kappa_{ij}^{\rm s-env}(\omega)$ and site-independent $\kappa_{ii}^{\rm s-env}(\omega) = \kappa^{\rm s-env}(\omega)$. The $\boldsymbol{H}_{\rm reorg}$ is divided into two parts $\boldsymbol{H}_{\rm reorg}^{\rm env} + \boldsymbol{H}_{\rm reorg}^{\rm s-env}$ where $\boldsymbol{H}_{\rm reorg}^{k}$ $= \Sigma_{\omega,i,j}S_{ij}^k(\omega)A_i^{k\dagger}(\omega)A_j^k(\omega)$ (for $k = {\rm env}, {\rm s-env}$) with $S_{ij}(\omega)$ as the imaginary part of the half-sided Fourier transform of the bath correlation function.

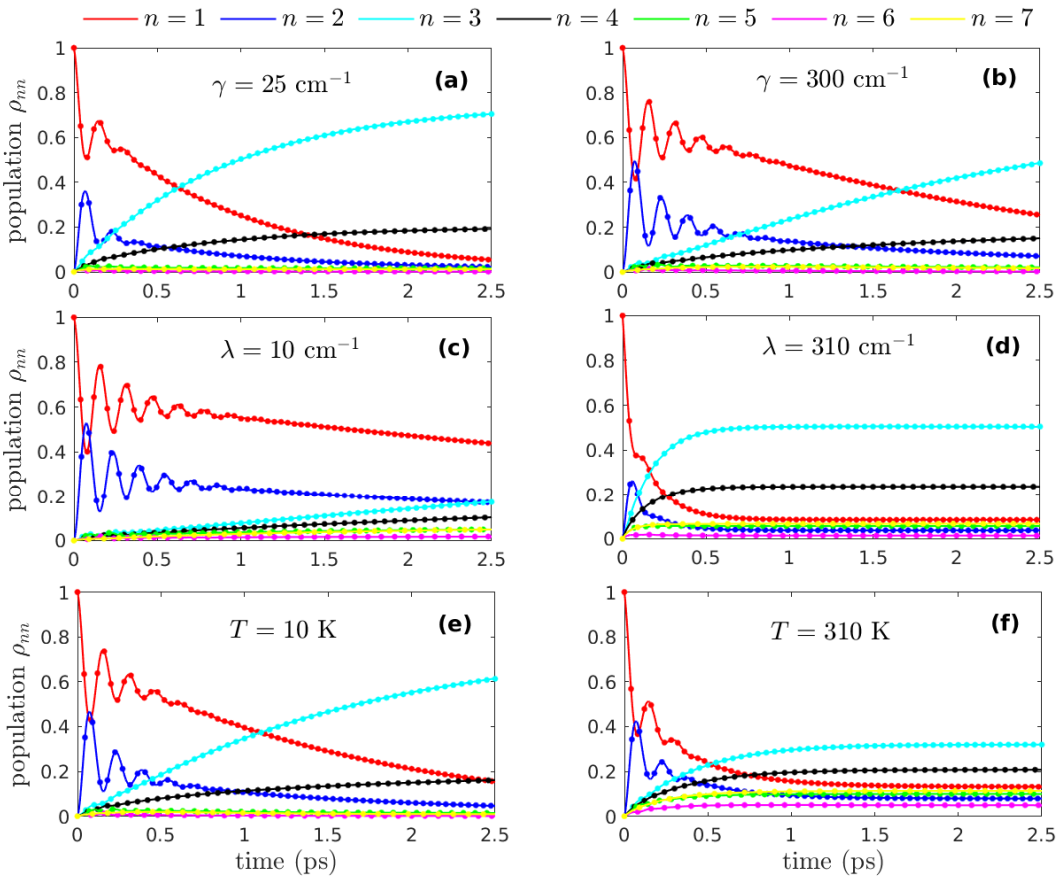

Figure S1: Population of the seven sites in the FMO complex as a function of time. The results are shown for extreme values of γ (as a scale of non-Markovian character of the bath), λ (as a scale of decoherence strength) and T in our parameter space with the aim to see their effect on the accuracy (see Table S1). Parameters are (a) $\gamma = 25$, $\lambda = 100$, $T = 50$, (b) $\gamma = 300$, $\lambda = 100$, $T = 50$, (c) $\gamma = 100$, $\lambda = 10$, $T = 150$, (d) $\gamma = 100$, $\lambda = 310$, $T = 150$, (e) $\gamma = 175$, $\lambda = 100$, $T = 30$ and (f) $\gamma = 175$, $\lambda = 100$, $T = 310$. The initial excitation is considered on site-1. The results of AI-QD are compared to the results of LTLME-QD (dots). γ and λ are in the units of cm^{-1}, while T is in the units of K.

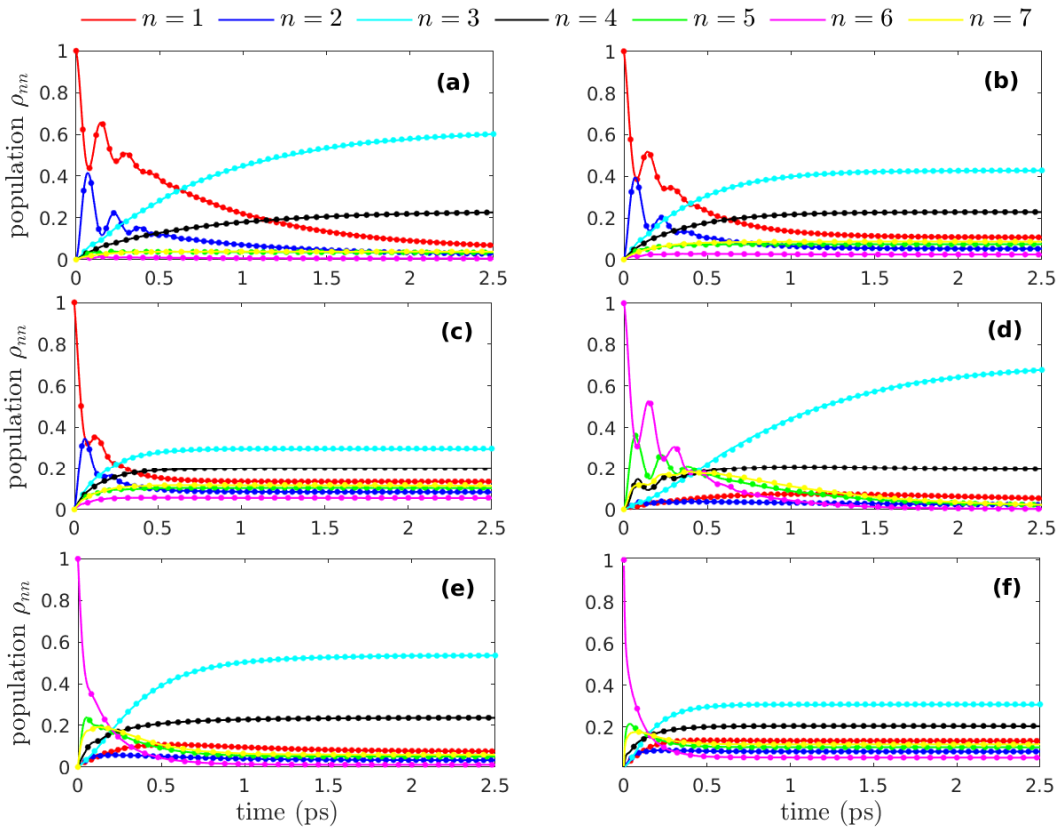

Figure S2: Population of the seven sites in the FMO complex as a function of time. The results are shown for parameters which do not appear in the training set at all. In (a), (b) and (c), the initial excitation is considered on site-1. Other parameters are (a) $\gamma = 80$, $\lambda = 85$, $T = 100$, (b) $\gamma = 205$, $\lambda = 185$, $T = 195$ and (c) $\gamma = 350$, $\lambda = 350$, $T = 350$. In (d), (e) and (f), the initial excitation is on site-6 and other parameters are (d) $\gamma = 65$, $\lambda = 35$, $T = 55$, (e) $\gamma = 155$, $\lambda = 145$, $T = 135$ and (f) $\gamma = 330$, $\lambda = 330$, $T = 330$. The results of AI-QD are compared to the results of LTLME-QD (dots). γ and λ are in the units of cm^{-1}, while T is i the units of K. The respective errors are given in Table S2. Note that (c) and (f) are outside the range of parameters used for training.

Table S1: Mean absolute error (MAE) and root mean square error (RMSE) for the test trajectories presented in Fig. S1. Here we show the dependence of accuracy on parameters γ, λ and T. The test trajectories were propagated up to 1 ns. $\mathcal{R}\{\rho_{mn,\,n\neq m}\}$ and $\mathcal{I}\{\rho_{mn,\,n\neq m}\}$ respectively represent the real and imaginary part of the off-diagonal terms.

parameters				mean absolute error			root mean square error		
IES	γ	λ	T	ρ_{nn}	$\mathcal{R}\{\rho_{mn,\,n\neq m}\}$	$\mathcal{I}\{\rho_{mn,\,n\neq m}\}$	ρ_{nn}	$\mathcal{R}\{\rho_{mn,\,n\neq m}\}$	$\mathcal{I}\{\rho_{mn,\,n\neq m}\}$
1	25	100	50	$8.7{\cdot}10^{-4}$	$9.0{\cdot}10^{-4}$	$3.6{\cdot}10^{-4}$	$1.2{\cdot}10^{-3}$	$1.5{\cdot}10^{-3}$	$6.4{\cdot}10^{-4}$
1	300	100	50	$4.6{\cdot}10^{-4}$	$2.5{\cdot}10^{-4}$	$1.3{\cdot}10^{-4}$	$6.8{\cdot}10^{-4}$	$3.7{\cdot}10^{-4}$	$1.9{\cdot}10^{-4}$
6	25	100	50	$8.3{\cdot}10^{-4}$	$7.1{\cdot}10^{-4}$	$2.3{\cdot}10^{-4}$	$1.1{\cdot}10^{-3}$	$1.1{\cdot}10^{-3}$	$3.5{\cdot}10^{-4}$
6	300	100	50	$6.7{\cdot}10^{-4}$	$3.0{\cdot}10^{-4}$	$1.4{\cdot}10^{-4}$	$1.0{\cdot}10^{-4}$	$4.4{\cdot}10^{-4}$	$2.0{\cdot}10^{-4}$
1	100	10	150	$6.5{\cdot}10^{-4}$	$2.8{\cdot}10^{-4}$	$1.3{\cdot}10^{-4}$	$1.3{\cdot}10^{-3}$	$5.0{\cdot}10^{-4}$	$1.9{\cdot}10^{-4}$
1	100	310	150	$2.3{\cdot}10^{-3}$	$1.0{\cdot}10^{-3}$	$3.1{\cdot}10^{-4}$	$3.6{\cdot}10^{-3}$	$1.7{\cdot}10^{-3}$	$5.1{\cdot}10^{-4}$
6	100	10	150	$7.4{\cdot}10^{-4}$	$2.8{\cdot}10^{-4}$	$1.6{\cdot}10^{-4}$	$1.4{\cdot}10^{-3}$	$5.1{\cdot}10^{-4}$	$2.5{\cdot}10^{-4}$
6	100	310	150	$1.6{\cdot}10^{-3}$	$6.8{\cdot}10^{-4}$	$2.3{\cdot}10^{-4}$	$2.6{\cdot}10^{-3}$	$1.0{\cdot}10^{-3}$	$3.5{\cdot}10^{-4}$
1	175	100	30	$4.3{\cdot}10^{-4}$	$3.5{\cdot}10^{-4}$	$1.4{\cdot}10^{-4}$	$6.3{\cdot}10^{-4}$	$5.8{\cdot}10^{-4}$	$2.0{\cdot}10^{-4}$
1	175	100	310	$1.6{\cdot}10^{-3}$	$4.7{\cdot}10^{-4}$	$2.0{\cdot}10^{-4}$	$2.5{\cdot}10^{-3}$	$7.5{\cdot}10^{-4}$	$2.9{\cdot}10^{-4}$
6	175	100	30	$5.1{\cdot}10^{-4}$	$3.0{\cdot}10^{-4}$	$1.4{\cdot}10^{-4}$	$8.5{\cdot}10^{-4}$	$4.9{\cdot}10^{-4}$	$2.2{\cdot}10^{-4}$
6	175	100	310	$1.3{\cdot}10^{-3}$	$3.8{\cdot}10^{-4}$	$2.0{\cdot}10^{-4}$	$2.1{\cdot}10^{-3}$	$6.3{\cdot}10^{-4}$	$3.0{\cdot}10^{-4}$

Table S2: Mean absolute error (MAE) and root mean square error (RMSE) for the test trajectories presented in Fig. S2. The test trajectories were propagated up to 1 ns. $\mathcal{R}\{\rho_{mn,\,n\neq m}\}$ and $\mathcal{I}\{\rho_{mn,\,n\neq m}\}$ respectively represent the real and imaginary part of the off-diagonal terms.

parameters				mean absolute error			root mean square error		
IES	γ	λ	T	ρ_{nn}	$\mathcal{R}\{\rho_{mn,\,n\neq m}\}$	$\mathcal{I}\{\rho_{mn,\,n\neq m}\}$	ρ_{nn}	$\mathcal{R}\{\rho_{mn,\,n\neq m}\}$	$\mathcal{I}\{\rho_{mn,\,n\neq m}\}$
1	80	85	100	$9.5{\cdot}10^{-4}$	$4.5{\cdot}10^{-4}$	$1.9{\cdot}10^{-4}$	$1.5{\cdot}10^{-3}$	$7.0{\cdot}10^{-4}$	$2.6{\cdot}10^{-4}$
1	205	185	195	$1.2{\cdot}10^{-3}$	$4.5{\cdot}10^{-4}$	$2.0{\cdot}10^{-4}$	$2.1{\cdot}10^{-3}$	$7.4{\cdot}10^{-4}$	$2.9{\cdot}10^{-4}$
1	350	350	350	$4.9{\cdot}10^{-3}$	$1.2{\cdot}10^{-3}$	$5.0{\cdot}10^{-4}$	$6.9{\cdot}10^{-3}$	$1.6{\cdot}10^{-3}$	$5.8{\cdot}10^{-4}$
6	65	35	55	$7.8{\cdot}10^{-4}$	$4.4{\cdot}10^{-4}$	$2.1{\cdot}10^{-4}$	$1.2{\cdot}10^{-3}$	$6.2{\cdot}10^{-4}$	$3.4{\cdot}10^{-4}$
6	155	145	135	$1.1{\cdot}10^{-3}$	$5.0{\cdot}10^{-4}$	$1.9{\cdot}10^{-4}$	$1.9{\cdot}10^{-3}$	$7.5{\cdot}10^{-4}$	$2.8{\cdot}10^{-4}$
6	330	330	330	$3.8{\cdot}10^{-3}$	$9.0{\cdot}10^{-4}$	$5.3{\cdot}10^{-4}$	$5.5{\cdot}10^{-3}$	$1.2{\cdot}10^{-3}$	$7.6{\cdot}10^{-4}$

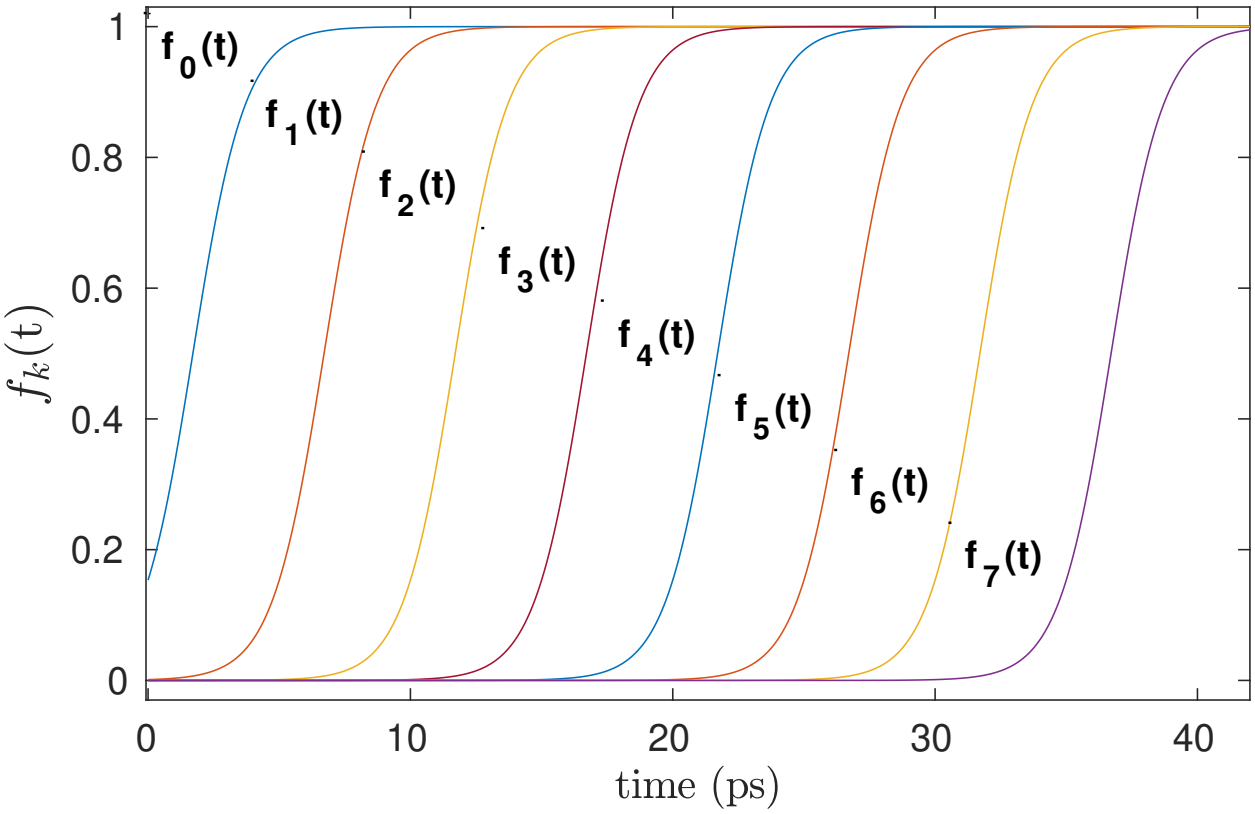

Figure S3: Logistic functions $f_k(t) = 1/(1 + 15 \cdot \exp(-(t + c_k)))$, where $c_k = 5k - 1.0$ and $k \in \{0, 1, 2, \ldots, 99\}$. Each logistic function is designed to cover the corresponding ≈ 5 ps region. Here only the first eight logistic functions are shown covering a time-region of 40 ps.

References

(1) Mohseni, M.; Rebentrost, P.; Lloyd, S.; Aspuru-Guzik, A. Environment-assisted quantum walks in photosynthetic energy transfer. *The Journal of chemical physics* **2008**, *129 17*, 174106.