High-performance solution of hierarchical equations of motions for studying energy-transfer in light-harvesting complexes

Christoph Kreisbeck, Tobias Kramer and Birgit Hein
Institut für Theoretische Physik, Universität Regensburg, 93040 Regensburg, Germany

Mirta Rodríguez
Instituto de Estructura de la Materia CSIC, C/ Serrano 121, 28006 Madrid, Spain

Excitonic models of light-harvesting complexes, where the vibrational degrees of freedom are treated as a bath, are commonly used to describe the motion of the electronic excitation through a molecule. Recent experiments point toward the possibility of memory effects in this process and require to consider time non-local propagation techniques. The hierarchical equations of motion (HEOM) were proposed by Ishizaki and Fleming to describe the site-dependent reorganization dynamics of protein environments (J. Chem. Phys., 130, p. 234111, 2009), which plays a significant role in photosynthetic electronic energy transfer. HEOM are often used as a reference for other approximate methods, but have been implemented only for small systems due to their adverse computational scaling with the system size. Here, we show that HEOM are also solvable for larger systems, since the underlying algorithm is ideally suited for the usage of graphics processing units (GPU). The tremendous reduction in computational time due to the GPU allows us to perform a systematic study of the energy-transfer efficiency in the Fenna-Matthews-Olson (FMO) light-harvesting complex at physiological temperature under full consideration of memory-effects. We find that approximative methods differ qualitatively and quantitatively from the HEOM results and discuss the importance of finite temperature to achieve high energy-transfer efficiencies.

I. INTRODUCTION

Light-harvesting complexes (LHC) are pigment protein-complexes that act as the functional units of photosynthetic systems, capable of absorbing the energy of a photon and transferring it towards the reaction center where it is converted into chemical energy usable for the cell. The transfer of energy in such systems is described by electronic exciton-dynamics coupled to the vibrations and other mechanical modes of the complex [1]. Laser spectroscopy shows quantum coherent effects in the energy transfer in LHC at temperatures up to 300 K [2–4].

Theoretical studies of model Hamiltonians at different levels of approximation [5–11] show that the interplay between coherent transport and dissipation leads to high efficiencies in the energy transport in these systems. LHC provide a remarkable example of systems where noise or dissipation aids the transport. Understanding these systems is relevant as it gives insight into the optimal design of artificial systems such as novel nanofabricated structures for quantum transport or optimized solar cells.

The modelling of LHC is challenging due to the lack of atomistic ab-initio methods and requires to resort to effective descriptions. This is most apparent in the treatment of the vibrational excitations, which are commonly described by a structureless mode distribution. Then the energy transfer is calculated by the time propagation of a density matrix, which couples the electronic exciton dynamics to the vibrational environment. For LHC, the rearrangement of the molecular states after the absorption of the photon has to be taken into account and is described by the reorganization energy. The hierarchical equations of motion (HEOM) [12–14] for the time evolution of the density-matrix were adapted by Ishizaki and Fleming [15] to include the reorganization process in the transport equations and is exact within the model of exciton dynamics coupled to a bath with a Drude-Lorentz spectral density.

In principle the HEOM can be extended to other spectral densities by using a superposition of Drude-Lorentz peaks [16, 17]. Previous calculations for the energy-transfer efficiency of the FMO complex did not consider memory effects and used a weak coupling perturbation theory [7, 8]. Other models try to get around these limitations by using the generalized Bloch-Redfield equations [9], but yield different results compared to the HEOM solution of the same model-system. Prolonged coherent dynamics is predicted due to the slow dissipation of reorganization energy to the vibrational environment [13]. Theoretical descriptions must go beyond the rotating-wave approximation, perturbation theory, and require a full incorporation of time non-local effects, and physiological temperature. The HEOM fulfill all these premises.

*Electronic address: tobias.kramer@mytum.de
To date, only the exciton population-dynamics for the FMO model has been studied within the full hierarchical approach [18, 19] whereas the calculation of efficiency or 2D absorption spectra have been considered out-of-range for present computational power, since they require stable algorithms to propagate enlarged system matrices over many more time-steps. The adverse computational scaling of the HEOM stems from the need to propagate a complete hierarchy of coupled auxiliary equations, which need to be simultaneously accessed in memory and propagated in time. The insufficient computational power and memory-transfer bandwidth of conventional CPU clusters [20] has limited the application of the HEOM to study energy-transfer efficiency in small dimer systems, where other methods are available for comparison around $T = 0$ K [21–24]. The advent of high-performance graphics processing units (GPU) with several hundred stream-processors working in parallel and with a high-bandwidth memory has lead us to perform the full HEOM approach for the exciton model of LHC. The efficiency calculations for the FMO system in the strong coupling regime require to propagate 240 000 auxiliary matrices up to 50 ps (corresponding to 20 000 time steps). The full HEOM approach takes only hours of computational time on a single GPU, whereas the corresponding CPU calculation would run several weeks and becomes completely unfeasible for bigger LHC due to the large communication overhead. We use the GPU algorithmic advance to characterize the exciton energy-transfer efficiency in LHC for a wide range of reorganization energies under full consideration of the memory-effects and at $T = 300$ K. Our calculations reveal several important mechanisms which are not contained within the approximative methods. The GPU-HEOM method opens the window to a wide-spread utilization of the HEOM, including the calculation of two-dimensional non-linear spectra of LHC as we will discuss elsewhere. Also the implementation of a scaled version of the HEOM [19], which reduces the number of auxiliary matrices, could be achieved on a GPU and reduces the computational effort of hierarchical methods further.

For the development of new theoretical chemistry and physics algorithms, GPU are important devices and considerably enlarge the class of solvable problems if one manages to devise a program code which takes full advantage of the GPU stream-processing architecture. For interacting many-body systems, this cannot be generally achieved by porting an existing program to the GPU, but requires to follow the vector-programming paradigm from the onset [25, 26].

The manuscript is organized as follows: in Sect. II we set up the model for energy transfer to the reaction center in the FMO complex. In Sect. III we calculate the key-quantities used to characterize the energy flow, namely the efficiency and the transfer time to the reaction center. We compute them for a wide range of reorganization energies and bath correlation-times within the hierarchical approach. This section contains a detailed discussion of the differences of the HEOM results compared to calculations based on approximative methods. We highlight the main mechanism behind the high efficiency, the delicate balance between the requirements of an energy gradient towards the reaction center and the detuning of the energies, as shown in Sect. IV. In Sect. V we discuss how the transport efficiency is optimized with respect to physiological temperature and comment on the thermalization properties of the HEOM. Finally we summarize our findings in Sect. VI. Throughout the article, we provide detailed information about the computational times and requirements and collect in the appendices additional detailed information about the algorithms used and our GPU implementation.

II. MODEL

The FMO protein is part of the light harvesting complex that appears in green sulfur bacteria. Its structure has been widely studied both with X-ray and optical spectroscopic techniques [27–29]. It has a trimer structure, with each of the monomers consisting of seven bacteriochlorophyll (BChl) pigment molecules, which are electronically excited when the energy flows from the antenna to the reaction center. An ab-initio calculation of the energy-transfer process within an atomistic model is far beyond present computational capabilities. Instead one has to develop effective model Hamiltonians such as the widely used excitonic Frenkel-Hamiltonian [1, 30, 31]. Within the Frenkel model, which assumes that excitations enter the system one at a time, the seven BChl pigments of the FMO complex are treated as individual sites which are coupled to each other and also to the protein environment. The excitonic Hamiltonian is given by

\[
\mathcal{H}_{\text{ex}} = E_0 |0\rangle \langle 0| + \sum_{m=1}^{N} (\varepsilon_m^0 + \lambda_m) |m\rangle \langle m| + \sum_{m>n} J_{mn} |m\rangle \langle n| + |n\rangle \langle m| ,
\]

where $N = 7$, $|m\rangle$ corresponds to an electronic excitation of the chromophore BChl$_m$ and $|0\rangle$ denotes the electronic ground state of the pigment protein complex where we fix the zero of energy $E_0 = 0$. The site energies $\varepsilon_m = \varepsilon_m^0 + \lambda_m$ of
the chromophores consist of the “zero-phonon energies” $\varepsilon_m^0$ and a reorganization energy $\lambda_m$, which takes into account the rearrangement of the complex during excitation due to the phonon bath:

$$H_{\text{reorg}} = \sum_{m=1}^{N} \lambda_m |m\rangle\langle m|.$$  \hspace{1cm} (2)

In the following we will consider identical couplings for all sites, $\lambda_m = \lambda$.

The inter site couplings $J_{mn}$ are obtained by fits to experimentally measured absorption spectra \cite{29}. In this contribution we use the designations and parameters of Ref. \cite{32}, table 4 (trimer column) and table 1 (column 4), summarized in \textbf{I}. A sketch of the dominant couplings is shown in \textbf{I}. The protein environment surrounding the pigments is modeled as identical featureless spectral bath densities coupled to each BChl. For simplicity, we neglect correlations between the baths. The electronic excitations at each site couple linearly with strength $d_i$ to the vibrational phonon modes $b_i^\dagger$ of frequency $\omega_i$. The coupling Hamiltonian is given by

$$H_{\text{ex-phon}} = \sum_{m=1}^{N} \left( \sum_i \hbar \omega_i d_i (b_i + b_i^\dagger) \right) |m\rangle\langle m|,$$  \hspace{1cm} (3)

where we assume identical baths at every site. Note that the reorganization energy is related to the coupling by

| BChl_1 | BChl_2 | BChl_3 | BChl_4 | BChl_5 | BChl_6 | BChl_7 |
|--------|--------|--------|--------|--------|--------|--------|
| BChl_1 | 12410  | -87.7  | 5.5    | -5.9   | 6.7    | -13.7  | -9.9   |
| BChl_2 | -87.7  | 12530  | 30.8   | 8.2    | 0.7    | 11.8   | 4.3    |
| BChl_3 | 5.5    | 30.8   | 12210  | -53.5  | -2.2   | -9.6   | 6.0    |
| BChl_4 | -5.9   | 8.2    | -53.5  | 12320  | -70.7  | -17.0  | -63.3  |
| BChl_5 | 6.7    | 0.7    | -2.2   | -70.7  | 12480  | 81.1   | -1.3   |
| BChl_6 | -13.7  | 11.8   | -9.6   | -17.0  | 81.1   | 12630  | 39.7   |
| BChl_7 | -9.9   | 4.3    | 6.0    | -63.3  | -1.3   | 39.7   | 12440  |

\textbf{TABLE I: Exciton Hamiltonian in the site basis in (cm}^{-1}). Bold font shows the dominant couplings and site energies. Values taken from Ref. \cite{32}.
\[ \lambda = \sum_i \hbar \omega_i d_i^2 / 2. \]

We model the losses due to radiative decay from the exciton to the electronic ground state \(|0\rangle\) introducing a dipole coupling to an effective radiation photon field \(a^\dagger_\nu\)

\[ H_{\text{ex-phot}} = \sum_{m=1}^{N} \sum_{\nu} (a_\nu + a^\dagger_\nu) \mu_\nu^m (|0\rangle \langle m| + |m\rangle \langle 0|), \]

which results in a finite life-time for the exciton. The reaction center (RC) is treated as a population-trapping state

\[ H_{\text{trap}} = E_{\text{RC}} |RC\rangle \langle RC| \]

and enlarges the system Hamiltonian to a \(9 \times 9\) matrix. Adolphs and Renger \cite{32} suggest that pigments 3 and 4, which have the largest overlap with the energetically lowest exciton-state, couple to the reaction center. Recent experimental evidence shows that pigment 3 is orientated towards the reaction center \cite{33}. In addition it has been proposed that an 8th pigment may play a role in the initial stages of the energy transfer \cite{34}. Here, we include the reaction center by introducing leakage rates from pigments 3 and 4 to the reaction center, which acts as a population trapping state. Thus the coupling term to the reaction center reads

\[ H_{\text{ex-RC}} = \sum_{m=3}^{4} \sum_{\nu'} (a_{\nu'} + a^\dagger_{\nu'}) \mu_{\nu'RC}^m (|RC\rangle \langle m| + |m\rangle \langle RC|) \]

where the sum runs over the photon modes at the reaction center. As shown in Sect. A, Eqs. (A7,A8), the coupling can be expressed in terms of a trapping rate \(\Gamma_{\text{phot}}\), and similarly for the radiative decay in \(\Lambda\) with the rate \(\Gamma_{\text{phot}}\). The total Hamiltonian of the system is thus given by

\[ \mathcal{H} = H_{\text{ex}} + H_{\text{trap}} + H_{\text{ex-phot}} + H_{\text{ex-RC}} + \mathcal{H}_{\text{phot}} + \rho_0 + \rho_{\text{RC}}, \]

where \(H_{\text{phot}} = \sum_{i,m} (\hbar \omega_i b^\dagger_i b_i)_m\), \(H^0_{\text{phot}} = \sum_{\nu,m} (\hbar \nu a^\dagger_{\nu} a_{\nu})_m\), and \(H_{\text{phot}}^{RC} = \sum_{\nu',m=3,4} (\hbar \nu' a^\dagger_{\nu'} a_{\nu'})_m\). The time evolution of the total density operator \(R(t)\) is described by the Liouville equation

\[ \frac{d}{dt}R(t) = -i \hbar [{\mathcal{H}}, R(t)]. \]

We assume that at initial time \(t = 0\) the total density operator factorizes in system and bath components

\[ R(t=0) = \rho(t=0) \otimes \rho_{\text{phot}}^{0} \otimes \rho_{\text{RC}}, \]

while at later times the system and the bath get entangled. Since we are only interested in the exciton dynamics, we trace out the degrees of freedom of the phonon and photon environments \(\alpha = \{\text{phon}, \text{phot}^{0}, \text{phot}^{RC}\}\) and propagate the reduced \(9 \times 9\) density matrix in the Schrödinger picture

\[ \rho(t) = \text{Tr}_\alpha (e^{-\frac{i}{\hbar} (\mathcal{L}_{\alpha} + \mathcal{L}_{\text{ex-phot}} + \mathcal{L}_{\text{ex-RC}} + \mathcal{L}_{\text{bath}}) t}) R(0) \]

for the exciton system \(\{|m\rangle\}_{m=1,...,7}\), the ground electronic state \(|0\rangle\), and the reaction center \(|RC\rangle\).

Eq. (10) is obtained by formal integration of the Liouville equation \cite{8}. The operator \(\mathcal{L}_{\alpha} = [H_{\text{ex}} + \mathcal{H}_{\text{trap}}]\) represents the coherent dynamics and \(\mathcal{L}_{\text{ex-phot}}\) accounts for dephasing and energy relaxation due to vibrations induced by the interaction with the protein environment, while the recombination and energy trapping are expressed by \(\mathcal{L}_{\text{ex-RC}}\) and \(\mathcal{L}_{\text{ex-RC}}\), respectively. The parts describing the different baths are summarized in \(\mathcal{L}_{\text{bath}} = [\mathcal{H}_{\text{phot}} + \mathcal{H}_{\text{phot}}^{0} + \mathcal{H}_{\text{phot}}^{RC}]\). The coupling to the phonon and photon baths can be studied with different degrees of approximation.

We calculate the energy flow within a hybrid formulation which treats the exciton dynamics and the vibrational environment within the HEOM and the trapping to the reaction center and the radiative decay within a Markov model. The Markovian treatment of the photon modes is justified as it occurs in a very different time scale and no backward energy flow to the system is allowed. We abbreviate our model by ME-HEOM, see Sect. A. We solve the hierarchical equations using GPUs, which are ideally suited for this task and lead to huge speed-ups of the algorithm. Details of the computational implementation are collected in Sect. B.
III. TRAPPING TIME FOR DIFFERENT REORGANIZATION ENERGIES

The strong coupling of the excitonic system to the vibrational environment, which is of the same order as the excitonic energy differences (100 cm$^{-1}$), requires a detailed treatment of the phonon bath over the time-scale of the correlations present in the system. The coupling is quantified by the parameter $\gamma$ Eq. (A17), ranging from (35-166 fs)$^{-1}$ for models of light-harvesting complexes [18]. We calculate the efficiency of the energy transfer from an initially excited site to the reaction center using the hierarchical equations [A20,A27]. The efficiency $\eta$ is defined as the population of the reaction center at long times

$$\eta = \langle RC|\rho(t \to \infty)|RC\rangle. \tag{11}$$

For the FMO complex, two sites are located near the light-absorbing antenna [32]. We consider initial excitations at either site 1 or 6, which give rise to two energy pathways to the reaction center. One pathway starts from site 1 and transfers energy via site 2 to site 3, and the second pathway starts from site 6 and the energy flows via site 7 or 5 to site 4, see [1].

We fix the upper limit of time propagation at $t_{\text{max}}$, defined such that the remaining population in the system, excluding the ground-state and reaction center, has dropped from initially 1 to $10^{-5}$. To our knowledge, no solid experimental data exists for the coupling strength in either site 1 or 6, which give rise to two energy pathways to the reaction center. One pathway starts from site 1 and transfers energy via site 2 to site 3, and the second pathway starts from site 6 and the energy flows via site 7 or 5 to site 4, see [1].

In the upper panel of 2 we compare the ME-HEOM result with the secular Redfield theory, which employs the Rotating-Wave approximation. In addition an optimal value of $\lambda$ emerges around $\gamma^{-1} = 25 - 35$ fs. For a small value $\gamma^{-1} = 5$ fs, the theory predicts a rapid loss of efficiency.

The lower panel of 2 details the changes of the trapping time for the two different pathways of the energy flow in the FMO complex as function of the reorganization energy. The optimal reorganization energy for an initial excitation of site 1 is given by $\lambda_{\text{opt}}^1 = 55$ cm$^{-1}$ ($\langle t \rangle_{\text{opt}} = 6.0$ ps), while for an initial excitation of site 6 we obtain $\lambda_{\text{opt}}^6 = 85$ cm$^{-1}$ ($\langle t \rangle_{\text{opt}}^6 = 5.4$ ps).

Optimal values of trapping times have been calculated within the generalized Bloch-Redfield (GBR) approximation [1]. Using the same parameters, couplings, and Hamiltonian as in Ref. [9], the ME-HEOM yield qualitative and quantitative differences with a 0.9 ps longer trapping time for an initial excitation of site 1. For an initial excitation located at site 6 the ME-HEOM and GBR results for the trapping time differ by 0.2 ps.

IV. EFFICIENCY FOR REARRANGED ENERGY LEVELS

In this section we study the relevance of the spacings of the energy levels in the FMO complex to see if the experimentally obtained energy levels [1] are close to an optimal value with respect to transport efficiency at physiological
The isolated excitonic system shows coherent oscillations of energy between the initially populated site and the delocalized excitonic states. Coupling to the environment gives rise to several mechanisms leading to a non-reversible energy transfer. In the simplest Haken-Strobl model, only dephasing is incorporated [7, 35], but the temperature is fixed at $T = \infty$. Only by adjusting the dephasing rate, temperature effects can be included on a rudimentary level. The ME-HEOM approach enables us to calculate the transport at physiological temperature ($T = 300$ K) and brings into the picture another crucial mechanism to achieve highly efficient energy transfer. Namely, the temperature dependent stationary site populations. Since the system is in contact with a thermal environment at finite temperature, there is energy dissipation and the system relaxes to thermal equilibrium. This process guides the excitons to the lowest energy states (for the FMO complex within a few picoseconds) and is not contained in pure dephasing models.

For a small coupling $\lambda$ and under the assumption that the system and bath degrees of freedom factggze, the thermal

![Graph showing trapping time as a function of reorganization energy](image-url)

**FIG. 2:** Trapping time from initial site as function of reorganization energy $\lambda$ at temperature $T = 300$ K. Trapping rate to BChl 3 and 4 $\Gamma_{\text{RC}}^{-1} = 2.5$ ps and $\Gamma_{\text{phot}}^{-1} = 250$ ps. Upper panel: secular Redfield result with $\gamma^{-1} = 166$ fs and the ME-HEOM results for three different bath correlation times $\gamma^{-1} = 166$ fs, $\gamma^{-1} = 50$ fs, $\gamma^{-1} = 5$ fs. The excitation enters at site 1. Lower panel: Comparison of the trapping times for the two possible pathways in the FMO when the energy is entering the complex starting from site 1, or at site 6 for a bath correlation time of $\gamma^{-1} = 166$ fs.
FIG. 3: Upper panel: Energy transfer efficiency $\eta$ in 11 as function of temperature and site-energy shifts $\varepsilon_{3/4} \rightarrow \varepsilon_{3/4} + \Delta E$. ME-HEOM parameters: $\lambda = 35 \text{ cm}^{-1}$, $\gamma^{-1} = 166$ fs, $\Gamma_{\text{phot}}^{-1} = 250$ ps and $\Gamma_{\text{RC}}^{-1} = 2.5$ ps. The hierarchy is truncated at $N_{\text{max}} = 8$. Lower panel: energy-level shifts considered in the parameter range of the left panel.

The state of the system is given by the Gibbs measure

$$\rho_{\text{thermal}} = e^{-\beta \mathcal{H}_{\text{ex}}}/\text{Tr} \ e^{-\beta \mathcal{H}_{\text{ex}}}, \ \beta = 1/(k_B T),$$

which populates the eigenstates of $\mathcal{H}_{\text{ex}}$ according to the Boltzmann statistics. Stronger couplings lead to deviations from the Boltzmann statistics [36]. Since the coupling to the reaction center, where the system deposits its excitation, is linked to sites 3 and 4, the efficiency depends strongly on the population and actual site-energies 3 and 4. To study this relation, we shift levels $\varepsilon_{3/4} \rightarrow \varepsilon_{3/4} + \Delta E$ and compute the efficiency of the energy transfer. Figure 3 shows the efficiency evaluated with the ME-HEOM. We observe an almost symmetric behavior of the efficiency for positive and negative energy shifts, with slightly higher efficiencies towards negative energy shifts.

A shift to lower energies increases the energy gradient in the FMO as the thermal state prefers to populate the low-lying sites. This mechanism improves the transfer efficiency but shifts the two sites out of resonance and they get
decoupled from the other levels of the FMO. Thus, coherent transport becomes more difficult and the energy transfer to the reaction center is expected to slow down. Similar arguments hold when the energies $\varepsilon_3$ and $\varepsilon_4$ are shifted to higher energies. On the one hand $\Delta E > 0$ brings the sites 3 and 4 closer to resonance and increases the coupling to the remaining sites, thus enhancing coherent transport. On the other hand, the thermal state gets delocalized over all sites of the FMO complex and there is no special preference to populate site 3 and site 4. In such case the FMO loses its property to act as an energy funnel and environment-assisted transport to the reaction center is hindered.

We show how the delicate interplay between coherent delocalization and energy dissipation towards the reaction center gives rise to an optimal arrangement of site energies. We obtain maximal efficiency around $\Delta E = 0$ corresponding to the original parameters in I and the optimum value is robust against small variations in the site energies.

V. TRAPPING TIME FOR DIFFERENT TEMPERATURES

As discussed in the previous section, the environment assists the transport towards the thermal equilibrium state. In the FMO complex, the sites 3 and 4 are coupled to the reaction center and present the lowest exciton energies in the system (see I), thus the energy dissipation in the phonon environment enhances the population of those sites and hence the efficiency. With increasing temperature one might expect high transfer efficiencies because thermalisation occurs on a faster time scale. However, with increasing temperature higher energy states have a higher thermal-equilibrium population and thus the transport efficiency towards the reaction center decreases.

These two competing mechanisms result in an optimal temperature with maximal efficiency. Both mechanisms are already present in the secular Redfield limit, and the optimal energy transfer is obtained around 75 K, see Figure 4(a). Our ME-HEOM calculations predict optimal efficiency at slightly lower temperature 70 K, but this value is outside the range where our high-temperature implementation is supposed to work (see Sect. A). We obtain a steep increase of the trapping time for low temperatures shown in Figure 4(a), which is also reflected in the efficiencies 4(b). This increase in trapping time and decrease in efficiency is not present in the secular Redfield approach, which saturates for $T \to 0$. Although we take into account the lowest-order quantum correction to the Boltzmann statistics [18], at low temperatures more correction terms are required. One criteria to validate the HEOM is to check if the stationary state $\rho(t \to \infty)$ of the population dynamics of the isolated FMO, which is decoupled from the reaction center and radiative decay, approaches the thermal state. As is shown in Figure 4(c), the HEOM high-temperature implementation fails to approach the thermal state for temperatures below 100 K, where the HEOM predict an unphysical steep decent of population at low energy site 3 and hence transfer efficiency is underestimated. For temperatures above 100 K the high temperature limit agrees very well with the thermal state and the ME-HEOM results are reliable. Comparing our ME-HEOM results above 100 K to the secular Redfield ones shown in 4(a) and (b) we conclude that the Redfield approach, which is known to be valid in the weak coupling limit only, overestimates the efficiency and underestimates the trapping time.

VI. CONCLUSIONS

We have shown that the HEOM are computationally feasible for calculating the energy transfer for large systems following our GPU implementation. This algorithmic advance allowed us to calculate the efficiency and trapping time of the energy transfer in the FMO complex for a wide range of parameters. The results point to qualitative and quantitative deficiencies of approximative methods and show that an accurate treatment of memory effects and reorganization processes in the system-bath coupling of LHC is needed to evaluate the precise role of temperature, exciton energy-differences, the coupling strength, and the time correlations in the bath. The ME-HEOM yield longer trapping times and indicate the importance of memory effects and correlations in order to maximize the efficiency in the FMO complex at physiological temperature. Interestingly, the zero-shift energies of the FMO complex provide an almost optimal arrangement for funneling the energy flow to the reaction center at $T = 300$ K. Beyond the results for the FMO complex, our fast computational GPU-algorithm for the HEOM provides a robust and scalable way to treat bigger systems and allows us to calculate two-dimensional spectra of LHC, which requires to enlarge the dimension of the density matrix by taking into account double-excitonic states.

Acknowledgement

This work has been supported by the DAAD project 50240755 and the Spanish MINECO AI DE2009-0088 (Acciones Integradas Hispano-Alemanas), the Emmy-Noether program of the DFG, KR 2889/2, the Spanish MICINN project FIS2010-18799 and the Ramón y Cajal program.
FIG. 4: Energy transfer as a function of temperature for the secular Redfield approximation and the exact ME-HEOM calculation with $\gamma^{-1} = 166$ fs, $\Gamma^{-1}_\text{phot} = 250$ ps, $\Gamma^{-1}_\text{RC} = 2.5$ ps and truncation $N_{\text{max}} = 8$. Both approaches use a reorganization energy of $\lambda = 35$ cm$^{-1}$ and start with initial population at site 1. (a) Trapping time as a function of temperature. (b) Efficiency as a function of temperature. (c) Population of site 3 for different temperatures in the Boltzmann thermal equilibrium state $\rho_{\text{thermal}}$ and for the isolated FMO (decoupled from the reaction center and the radiative decay) using the ME-HEOM, $\rho(t \to \infty)$. Note that the ME-HEOM needs further corrections at temperatures below 100 K in order to reach the thermal state.
Appendix A: Hybrid Markov-HEOM approach

1. Secular Redfield approximation

We follow Ref. [37] and expand the phonon part $L_{\text{phon}}$ up to second order in the exciton-phonon coupling. We use the Born-Markov and secular approximation to obtain

$$-\frac{i}{\hbar} \mathcal{L}_{\text{exc-phon}} = \sum_{m, \omega} i L_m(\omega) \langle V_m^\dagger(\omega) V_m(\omega), \cdot \rangle + \sum_{m, \omega} i \gamma_m(\omega) D(V_m(\omega)), \quad (A1)$$

where $V_m = |m\rangle \langle m|$ stands for the exciton operators and $D(V)\rho = V\rho V^\dagger - \frac{1}{2} V^\dagger V \rho - \frac{1}{2} \rho V^\dagger V$. The Lamb shift reads

$$L_m(\omega) = \text{Im} \int_0^\infty dt e^{-i \omega t} \langle u_m(t) u_m(0) \rangle_{\text{phon}}, \quad (A2)$$

and the decoherence rates are given by

$$\gamma_m(\omega) = 2 \text{Re} \int_0^\infty dt e^{-i \omega t} \langle u_m(t) u_m(0) \rangle_{\text{phon}}, \quad (A3)$$

with the phonon operators $u_{m,\text{phon}} = \left( \sum_i \hbar \omega_i d_i(b_i + b_i^\dagger) \right)_m$. The exciton operators

$$V_m(\omega) = \sum_{\omega, M, N} c_{m}^\dagger(M) c_N(N) |M \rangle \langle N| \delta(\omega - E_M + E_N) \quad (A4)$$

are evaluated in the excitonic eigenbasis $|M\rangle = \sum_m c_m(M) |m\rangle$ with $\mathcal{H}_{\text{exc}}|M\rangle = E_M |M\rangle$. For simplicity we assume that the phonon environments of the individual chromophores are uncorrelated. We additionally neglect the Lamb-type renormalization term. For the explicit evaluation of the decoherence rates, we quantify the strength of the exciton-phonon coupling and introduce a Drude-Lorentz spectral density

$$J_m(\omega) = 2 \lambda_m \frac{\omega \gamma_m}{\omega^2 + \gamma_m^2}. \quad (A5)$$

The decoherence rates are then given by

$$\gamma(\omega) = \begin{cases} 2 \pi J(-\omega)(n(-\omega) + 1), & \text{if } \omega < 0 \\ 2 \pi \frac{\hbar \omega T}{k_B} \frac{dJ(\omega)}{d\omega}, & \text{if } \omega = 0 \\ 2 \pi J(\omega)n(\omega), & \text{if } \omega > 0 \end{cases} \quad (A6)$$

where $n(\omega) = (\exp(h\omega/k_B T) - 1)^{-1}$ corresponds to the phonon statistics. Note that we neglect the index $m$ and use the same parameters for all sites.

We describe the radiative decay and trapping to the reaction center by a Lindblad ansatz

$$-\frac{i}{\hbar} \mathcal{L}_{\text{exc-phot}} = \sum_{m=1}^N \Gamma_{\text{phot}} D(|0\rangle \langle m|), \quad (A7)$$

and

$$-\frac{i}{\hbar} \mathcal{L}_{\text{exc-RC}} = \sum_{m=3}^4 \Gamma_{\text{RC}} D(|RC\rangle \langle m|), \quad (A8)$$

with identical trapping rates for all sites $\Gamma_{\text{phot}} = 2 \pi |\mu|^2$ and $\Gamma_{\text{RC}} = 2 \pi |\mu_{RC}|^2$, respectively. $A7$ and $A8$ are derived by a master equation approach employing the rotating wave approximation for Hamiltonians $\mathcal{H}_{\text{phot}}^0$ and $\mathcal{H}_{\text{phot}}^{\text{RC}}$, respectively. The strength of the exciton-photon coupling is defined by a constant spectral density $J(\omega) = 2 \pi$. We further assume that the photon field cannot create excitations. The same holds for the trapping to the reaction center, as no backward energy flow to the system is allowed. This is equivalent to previous works, where trapping and exciton recombination are included in the Hamiltonian in the form of anti-Hermitian parts $\mathcal{L}_{\text{RC}}$. The equation of motion for the density operator in secular Redfield form reads

$$\frac{d}{dt} \rho(t) = -\frac{i}{\hbar} (\mathcal{L}_0 + \mathcal{L}_{\text{exc-phon}} + \mathcal{L}_{\text{exc-phot}} + \mathcal{L}_{\text{exc-RC}}) \rho(t). \quad (A9)$$
2. Combined master-equation-HEOM approach (ME-HEOM)

For the FMO complex, the coupling of the excitons to the phonon environment is large and the reorganization process occurs on a time scale which is comparable to the system dynamics. The secular Redfield approximation is not valid for the coupling to the phonon bath and a non-perturbative treatment is required. We will follow the derivation in Ref. [15] and introduce a set of hierarchically coupled equations. For trapping-time and efficiency calculations we introduce slight modifications and in particular we include the coupling to the reaction center and the radiative decay. We derive a combined ME-HEOM approach which treats the exciton-phonon coupling exactly, whereas the leakage to the reaction center and exciton ground state is described in the Born-Markov limit.

We start with the Liouville equation for the total density operator $\mathcal{R}$ and assume that the total density operator factorizes $\mathcal{R} = \mathcal{R}_0 \otimes \rho_{\text{photon}} \otimes \rho_{\text{reaction center}}$. In the interaction picture with

$$\mathcal{H}_0 = \mathcal{H}_\text{ex} + \mathcal{H}_\text{trap} + \mathcal{H}_\text{phon} + \mathcal{H}_\text{phot}^0 + \mathcal{H}_\text{RC}^\text{phot},$$

where we denote operators with $\tilde{O}(t) = e^{i\mathcal{H}_0 t/\hbar} O e^{-i\mathcal{H}_0 t/\hbar}$,

the Liouville equation reads

$$\frac{d}{dt} \tilde{R}(t) = -\frac{i}{\hbar} [\tilde{\mathcal{H}}_{\text{ex-phon}} + \tilde{\mathcal{H}}_{\text{ex-phot}} + \tilde{\mathcal{H}}_{\text{ex-RC}}, \tilde{R}(t)]$$

$$= -\frac{i}{\hbar} \left( \tilde{\mathcal{L}}_{\text{ex-phon}} + \tilde{\mathcal{L}}_{\text{ex-phot}} + \tilde{\mathcal{L}}_{\text{RC}} \right) \tilde{R}(t).$$

(A12)

After formal integration and tracing out the bath degrees of freedom $\alpha = \{\text{phon, phot}^0, \text{phot}^\text{RC}\}$ we get a formal solution for the reduced density operator describing the exciton degrees of freedom

$$\tilde{\rho}(t) = \tilde{U}(t) \tilde{\rho}(0)$$

(A13)

with time evolution operator

$$\tilde{U}(t) = \text{Tr}_\alpha \left( T_+ e^{-\frac{i}{\hbar} \int_0^t ds \left( \tilde{\mathcal{L}}_{\text{ex-phon}}(s) + \tilde{\mathcal{L}}_{\text{ex-phot}}(s) + \tilde{\mathcal{L}}_{\text{ex-RC}}(s) \right) \rho_{\text{photon}} \otimes \rho_{\text{phot}^0} \otimes \rho_{\text{phot}^\text{RC}} \right).$$

(A14)

We make use of the Gaussian nature of the harmonic baths to reduce the bath expectation values to two-time correlation functions. Hence the influence of the environment is characterized by the symmetrized correlation

$$S_{m,\alpha}(t) = \frac{1}{2} \langle [\tilde{u}_{m,\alpha}(t), \tilde{u}_{m,\alpha}(0)]_+ \rangle,$$

(A15)

and the response function

$$\chi_{m,\alpha}(t) = \frac{1}{2} \langle [\tilde{u}_{m,\alpha}(t), \tilde{u}_{m,\alpha}(0)] \rangle,$$

(A16)

where $u_{m,\text{phot}} = \sum_\nu \mu_{m,\nu}^\alpha (a_\nu + a_\nu^\dagger)$. We assume that each site is coupled to an independent phonon bath and that there are no correlations between the radiative decay and trapping at different sites. For the exciton-phonon coupling we employ a Drude-Lorentz spectral density

$$J_m(\omega) = 2\lambda_m \frac{\omega \gamma_m}{\omega^2 + \gamma_m^2},$$

(A17)

and obtain, in the high temperature limit

$$S_m(t) \simeq \frac{2\lambda_m}{k_B T} e^{-\gamma_m t},$$

(A18)

$$\chi_m(t) = 2\lambda_m \gamma_m e^{-\gamma_m t}.$$
from the time correlation scale of the bath $\gamma$.

The time evolution operator becomes

$$
\hat{U}(t) = T_+ \prod_{m=1}^N e^{\int_0^t ds \hat{W}_{m,\text{phon}}(s)} \prod_{m=1}^N e^{\int_0^t ds \hat{W}_{m,\text{phot}}(s)} \prod_{m=3}^4 e^{\int_0^t ds \hat{W}_{m,\text{photRC}}(s)}
$$

(A20)

with

$$
\hat{W}_{m,\alpha} = -\frac{1}{\hbar} \int_0^t ds \hat{V}_{m,\alpha}(t) [S_{m,\alpha}(t-s)\hat{V}_{m,\alpha}(s)^\times - i\hbar \frac{1}{2} \chi_{m,\alpha}^{\times} \hat{V}_{m,\alpha}(s)^\circ].
$$

(A21)

We denote the commutation relations by $O^x f = [O, f]$ and $O^\circ f = [O, f]_+$. The time evolution of the reduced density matrix is given by

$$
\frac{d}{dt} \hat{\rho}(t) = T_+ \left( \sum_{m=1}^7 \hat{W}_{m,\text{phon}}(t) + \sum_{m=1}^7 \hat{W}_{m,\text{phot}}(t) + \sum_{m=3}^4 \hat{W}_{m,\text{photRC}}(t) \right) \hat{\rho}(t).
$$

(A22)

Note that due to the time ordering operator affecting the integration in Eqs. (A21, A22) is time non-local. In the following we treat the exciton-photon and exciton-reaction center couplings in the Born-Markov limit. That is, the time non-local operators $T_+ \sum_{m=1}^7 \hat{W}_{m,\text{phot}}(t)$ and $T_+ \sum_{m=3}^4 \hat{W}_{m,\text{photRC}}(t)$ are replaced by their time-local Born-Markov limit $\mathcal{L}_{\text{ex-phot}}$ and $\mathcal{L}_{\text{ex-RC}}$ defined in Eqs. (A7, A8), respectively. Eq. (A22) finally reduces to

$$
\frac{d}{dt} \hat{\rho}(t) = -\frac{i}{\hbar} \mathcal{L}_{\text{phot}} \hat{\rho}(t) - \frac{i}{\hbar} \mathcal{L}_{\text{RC}} \hat{\rho}(t) + T_+ \sum_{m=1}^7 \hat{W}_{m,\text{phon}}(t) \hat{\rho}(t).
$$

(A23)

We define auxiliary operators

$$
\hat{\sigma}^{(n_1,\ldots,n_\tau)}(t) = T_+ \prod_{m,k,l} \left( \int_0^t ds e^{-\gamma_m(t-s)} \hat{\theta}_m(s) \right) \hat{W}_{m,\text{phot}}(s) \hat{W}_{k,\text{phot}}(s) \hat{W}_{l,\text{photRC}}(s)
$$

(A24)

with

$$
\hat{\theta}_m(s) = i \left( \frac{2\lambda_m}{k_B T \hbar^2} \hat{V}_{m,\text{phon}}(s) - \frac{\lambda_m}{\hbar} \gamma_m \hat{V}_{m,\text{phon}}(s) \right),
$$

(A25)

and rewrite the time non-local effects into hierarchically coupled equations of motion

$$
\frac{d}{dt} \hat{\rho}(t) = -\frac{i}{\hbar} \left( \mathcal{L}_{\text{ex}} + \mathcal{L}_{\text{phot}} + \mathcal{L}_{\text{RC}} \right) \hat{\rho}(t) + \sum_m i \hat{V}_{m,\text{phot}} \sigma^{(n_1,\ldots,n_{m+1},\ldots,n_\tau)}(t)
$$

(A26)

with

$$
\frac{d}{dt} \sigma^{(n_1,\ldots,n_\tau)}(t) = \left[ -\frac{i}{\hbar} \mathcal{L}_{\text{ex}} + \mathcal{L}_{\text{phot}} + \mathcal{L}_{\text{RC}} \right] \sigma^{(n_1,\ldots,n_\tau)}(t) + \sum_m i \hat{V}_{m,\text{phot}} \sigma^{(n_1,\ldots,n_{m+1},\ldots,n_\tau)}(t)
$$

(A27)

where we again have used the Born-Markov limit for the trapping and radiative decay. The auxiliary operators keep track of the memory effects of the bath and account for the removal of the reorganization energy. The $\sigma$-matrices are initially set to zero. For a sufficiently large $N_{\text{max}} = \sum_m n_m$, the diagonal coupling in Eq. (A27) becomes the dominant term and we can truncate the hierarchy.
| $N_{\text{max}}$ | #σ-matr. | CPU (s) | GPU (s) | speed up | GPU utilization |
|-----------------|----------|---------|---------|----------|-----------------|
| 4               | 330      | 120     | 1       | $\times120$ | 22%             |
| 6               | 1,716    | 676     | 3       | $\times225$ | 56%             |
| 8               | 6,435    | 2,636   | 7       | $\times376$ | 82%             |
| 10              | 19,448   | 8,275   | 19      | $\times435$ | 93%             |
| 12              | 50,388   | 21,972  | 48      | $\times458$ | 97%             |

TABLE II: Comparison of CPU and GPU computation time of the population dynamics of the isolated FMO complex. We propagate 1000 time steps, the GPU (NVIDIA C2050) calculation are performed in single precision. Double precision (not required here for converged results) increases the GPU computation time by a factor of two.

Appendix B: Algorithm for implementing the hierarchical method on graphics processing units

For the large reorganization energies typically found in LHC one needs to go beyond the Born-Markov approach and to consider non-local temporal effects. We do this by solving the system dynamics within the hierarchical approach shown in the previous section. The method requires considerable memory and computational efforts and a large number of auxiliary matrices is needed to store the time non-local bath properties. Since all auxiliary matrices have to be accessed to perform the next propagation step, the huge communication overhead renders conventional parallelization schemes, where distributed computing nodes are connected by Ethernet [20], ineffective. GPUs have the twofold advantage of a fast memory bandwidth and the availability of several hundred stream processors. By assigning one stream-processor to each auxiliary matrix we obtain a speedup of the hierarchical method by the number of processors. The numerical calculations in this manuscript are performed on a NVIDIA Fermi C2050 GPU with 448 processors (1.15 GHz) and 3 gigabytes of ECC-protected on-board memory. The first step of the algorithm initializes the system of the $\sigma$-matrices of the hierarchy. With increasing truncation $N_{\text{max}}$, the total number of $\sigma$-matrices grows factorially $N_{\text{tot}} = (N + N_{\text{max}})!/(N!N_{\text{max}})!$, where $N$ corresponds to the number of sites [15]. As shown in [11] the calculation of a population dynamics of the FMO complex with $N = 7$, $N_{\text{max}} = 12$ requires already 50,388 matrices whereas 330 matrices are sufficient for a truncation at $N_{\text{max}} = 4$. The memory of the $\sigma$-matrices is allocated on the graphics-board and initialized to zero. It is not necessary to transfer the $\sigma$-matrices to the main-processor memory at any time during the calculation. The only memory transfer between CPU and GPU involves the $N \times N$ entries of the reduced density operator $\rho$. To advance the propagation one time-step in eq. (A27) requires a large number of matrix multiplications. Each single $\sigma$-matrix is connected to $2N$ neighbors, these connections are stored in GPU memory in a linked-list. The GPU uses 448 cores in parallel with fast GPU memory transfer and thus provides an immense reduction of the computation time up to a factor of 458 for the matrix multiplications. For benchmarking the algorithm, we propagate 1000 time steps using a 4th order Runge-Kutta integrator. For the final output into files requires a short memory transfer from the GPU to the CPU.

In [11] we summarize the computational speed-up of the C2050-GPU compared to a standard CPU (Intel 2.40GHz). The GPU computation is performed using single precision, which yields sufficient accuracy for the problem at hand. For the population dynamics of the FMO complex using $\lambda = 35 \, \text{cm}^{-1}$, $\gamma^{-1} = 166 \, \text{fs}$, temperature of 300 K, propagation time of 10 ps with step size $\Delta t = 10 \, \text{fs}$ and truncation $N_{\text{max}} = 12$ the populations are accurate within single precision to six digits $|\rho_{\text{single}}^{\text{t+1}} - \rho_{\text{double}}^{\text{t+1}}| < 5 \times 10^{-7}$. This switch from single for double precision increases the computation time approximately by a factor of two on the C2050-GPU.

[1] May, V.; Kühn, O. In Charge and Energy Transfer Dynamics in Molecular Systems; Weinheim: Wiley-VCH, 2004, p. 62.
[2] Engel, G. S.; Calhoun, T. R.; Read, E. L.; Ahn, T.-K.; Mancal, T.; Cheng, Y.-C.; Blankenship, R. E.; Fleming, G. R. Nature 2007, 446, 782.
[3] Collini, E.; Wong, C. Y.; Wilk, K. E.; Curmi, P. M. G.; Brumer, P.; Scholes, G. D. Nature 2010, 463, 644.
[4] Panitchayangkoon, G.; Hayes, D.; Fransted, K.; Caram, J.; Harel, E.; Wen, J.; Blankenship, R.; Engel, G. In PNAS 2010, 107, 12766.
[5] Gaab, K. M.; Bardeen, C. J. J. Chem. Phys. 2004, 121, 7813.
[6] Plenio, M. B.; Huelga, S. F. New J. Phys. 2008, 10, 113019.
[7] Rebentrost, P.; Mohseni, M.; Kassal, I.; Lloyd, S.; Aspuru-Guzik, A. New J. Phys. 2009, 11, 033003.
[8] Fussioli, F.; Olaya-Castro, A. New J. Phys. 2010, 12, 085006.
[9] Wu, J.; Liu, F.; Shen, Y.; Cao, J.; Silbey, R. J. New J. Phys. 2010, 12, 105012.
[10] Hsin, J.; Strümpfer, J.; Şener, M.; Qian, P.; Hunter, C. N.; Schulten, K. New J. Phys. 2010, 12, 085005.
[11] Hoyer, S.; Sarovar, M.; Whaley, K. B. New J. Phys. 2010, 12, 065041.
[12] Yan, Y.; Yang, F.; Liu, Y.; Shao, J. Chem. Phys. Lett. 2004, 395, 216.
[13] Xu, R.-X.; Cui, P.; Li, X.-Q.; Mo, Y.; Yan, Y. J. Chem. Phys. 2005, 122, 041103.
[14] Ishizaki, A.; Tanimura, Y. J. Phys. Soc. Jpn. 2005, 74, 3131.
[15] Ishizaki, A.; Fleming, G. R. J. Chem. Phys. 2009, 130, 234111.
[16] Meier, C.; Tannor, D. J. J. Chem. Phys. 1999, 111, 3365.
[17] Kleinekathöfer, U. J. Chem. Phys. 2004, 121, 2505–14.
[18] Ishizaki, A.; Fleming, G. R. PNAS 2009, 106, 17255.
[19] Zhu, J.; Kais, S.; Rebentrost, P.; Aspuru-Guzik, A. J. Chem. Phys. B 2011, 115, 1531–1537.
[20] Strümpfer, J.; Schulten, K. J. Chem. Phys. 2009, 131, 225101.
[21] Anders, F. B.; Bulla, R.; Vojta, M. Phys. Rev. Lett. 2007, 98, 210402.
[22] Thorwart, M.; Eckel, J.; Reina, J.; Nalbach, P.; Weiss, S. Chem. Phys. Lett. 2009, 478, 234.
[23] Roden, J.; Schulz, G.; Eisfeld, A.; Briggs, J. J. Chem. Phys. 2009, 131, 044909.
[24] Prior, J.; Chin, A. W.; Huelga, S. F.; Plenio, M. B. Phys. Rev. Lett. 2010, 105, 050404.
[25] Olivares-Amaya, R.; Watson, M. A.; Edgar, R. G.; Vogt, L.; Shao, Y.; Aspuru-Guzik, A. JCTC 2010, 6, 135–144.
[26] Kramer, T.; Kruelkl, V.; Heller, E.; Parrott, R. Phys. Rev. B: Condens. Matter 2010, 81, 205306.
[27] Olson, J. Photosynth. Research 2004, 80, 181–187.
[28] Brixner, T.; Stenger, J.; Vaswani, H. M.; Cho, M.; Blankenship, R. E.; Fleming, G. R. Nature 2005, 434, 625–8.
[29] Milder, M. T. W.; Brüggemann, B.; van Grondelle, R.; Herek, J. L. Photosynth. Res. 2010, 104, 257.
[30] Leegwater, J. A. J. Phys. Chem. 1996, 100, 14403.
[31] Ritz, T.; Park, S.; Schulten, K. J. Phys. Chem. B 2001, 105, 8259.
[32] Adolphs, J.; Renger, T. Biophys. J. 2006, 91, 2778.
[33] Wen, J.; Zhang, H.; Gross, M. L.; Blankenship, R. E. In PNAS 2009, 106, 6134.
[34] Schmidt am Busch, M.; Mühl, F.; El-Amine Madjet, M.; Renger, T. J. Phys. Chem. Lett. 2011, 2, 93.
[35] Chin, A. W.; Datta, A.; Caruso, F.; Huelga, S. F.; Plenio, M. B. New J. Phys. 2010, 12, 065002.
[36] Zürcher, U.; Talkner, P. Phys. Rev. A 1990, 42, 3267–3277.
[37] Rebentrost, P.; Chakraborty, R.; Aspuru-Guzik, A. J. Chem. Phys. 2009, 131, 184102.