Nonadiabatic molecular dynamics with decoherence and detailed balance under a density matrix ensemble formalism

Jun Kang and Lin-Wang Wang*
Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States
(Dated: May 24, 2019)

The mixed quantum/classical nonadiabatic molecular dynamics (NAMD) is a powerful tool to study many phenomena, especially ultrafast carrier transport and cooling. Carrier decoherence and detailed balance are two major issues in NAMD. So far, there is no computationally inexpensive approach to incorporate both effects. While the decoherence effect can be easily included in the state density matrix formalism, and the detailed balance can be included in surface hopping or wave function collapsing approach, it is difficult to include both of them in a unified formalism. In this work we introduce a state density matrix formalism (referred as P-Matrix) including both the decoherence and detailed balance effects for NAMD. This method is able to explicitly treat the decoherence between different pairs of adiabatic states. Moreover, the off-diagonal density matrix elements are divided into two parts, corresponding to energy-increasing and energy-decreasing transitions respectively. The detailed balance is then enforced by a Boltzmann factor applied to the energy-increasing transition part. The P-Matrix formalism is applied to study hot-hole cooling and transfer processes in Si quantum dot (QD) systems. The calculated hot carrier relaxation time is consistent with experiments. In a QD-pair system, the hot-hole cooling time shows weak dependence with the QD spacing. However, the hot carrier transfer rate from one QD to another is found to decrease exponentially with the QD-QD distance. When the QD spacing is small ($\sim$ 1 nm), the hot-carrier transfer can be very efficient. It is also shown that the explicit treatment of decoherence time is important in order to treat this hot-carrier transfer correctly.

I. INTRODUCTION

Nonadiabatic molecular dynamics (NAMD) simulation[1,2] is a widely used approach to study carrier dynamics processes involving excited states, such as charge relaxation[3–5], recombination,[6] and transport[7–10]. NAMD simulation is often carried out in a mixed quantum-classic (MQC) fashion, in which the electron degree of freedom is described quantum mechanically following the time dependent Schrodinger's equation (TDSE), whereas nuclear movement is treated classically following the Newton's second law. There are many MQC algorithms. In principle one can also consider the dynamics of the whole open system, namely the quantum subsystem coupled with the environment. This leads to the quantum-classical Liouville equation and generalized quantum master equation approaches[11–16]. Despite their rigor, their implementation could be complicated. Mean-field Ehrenfest dynamics (MFE)[17–20] and fewest switches surface hopping (FSSH)[21–23] are two simpler and more widely used algorithms. In the MFE, the nuclear movement follows the average atomic force provided by all the electron states solved by TDSE. There is no branching either for the electron wavefunction or for the nuclear trajectory, and the electron wavefunction is always described by a single coherent electron state (a many-electron state), instead of by an ensemble of states. In the FSSH approaches, the nuclei move along one adiabatic energy surface and stochastically hop to other surfaces, and the hopping rate is determined by an auxiliary wavefunction following the TDSE.

Due to the inconsistencies between quantum and classical mechanics in MQC-NAMD, there could be several issues.[24] One difficulty is to maintain the detailed balance, i.e., to reproduce the Boltzmann quantum state population in the long-time thermal equilibrium. It is well known that the MFE lacks detailed balance[7,25,26], leading to overheating of the electronic subsystem. The reason is that MFE is based on the electronic wavefunction alone, which does not include either the quantum mechanical wavefunction of the phonon, or any electronic coupling to an outside open system. This limits the capability of MFE in the study of equilibrium properties and energy relaxation processes. Detailed balance can be forced in MFE by introducing symmetrical coupling matrix elements with quantum corrections[27], but by doing so, the transition probabilities obtained from the TDSE is changed. Another possible solution is to include both zero point electronic energy and windowing on top of MFE[28], but it suffers from numerical instabilities[29]. On the other hand, although surface hopping does not satisfy detailed balance rigorously, the energy conservation requirement during the hopping by rescaling the relevant nuclei kinetic energy provides the detailed balance in an empirical and approximate way.[30] Unfortunately, the original surface hopping algorithm does not have the proper decoherence.

Decoherence is another quantum mechanical phenomenon caused by the separation of nuclear wavefunctions for different electronic states (or say potential energy surfaces). It describes the phenomenon where an original single electron state breaks down into many components which lose the ability to interfere with each other. More specifically, the system at t=0 can be described as: $\Psi(\mathbf{r},0)\Theta(\mathbf{R},0)$, where $\Psi(\mathbf{r},0)$ is the electron wavefunction, and $\Theta(\mathbf{R},0)$ is the nu-

clear wavefunction. At time t, the wavefunction will be evolved into: $\sum_i \phi_i(\mathbf{r},t)\Theta_i(\mathbf{R},t)$ (ϕ_i is a electronic adiabatic state). Due to the separation of nuclear wavefunctions, $\langle\Theta_i(\mathbf{R},t)|\Theta_j(\mathbf{R},t)\rangle$ decays with time, and the interference between ϕ_i and ϕ_j also vanishes. If $\langle\Theta_i(\mathbf{R},t)|\Theta_j(\mathbf{R},t)\rangle$ becomes zero for $i \neq j$, then $\phi_i(\mathbf{r},t)$ and $\phi_j(\mathbf{r},t)$ become decoherent (the dot product for most physical operators between these two states becomes zero). As a result, from the electron wavefunction point of view, different $\phi_i(\mathbf{r},t)$ can be described as an ensemble of wavefunctions, e.g., the wavefunction has "collapsed".

In their original forms, the separation of nuclear wavefunctions is not included in MFE and FSSH, hence the wavefunctions are fully coherent. Many empirical approaches have been proposed to introduce decoherence correction in MFE and FSSH. For example, one method is to include a coherence penalty functional that accounts for decoherence effects to the Hamiltonian in MFE.[31] Moreover, decay-of-mixing approaches have been developed for both MFE and FSSH,[32,33] in which the coefficients of the wavefunctions are modified using a decoherence time after the evolution of TDSE at each time step. Another set of popular approaches to describe decoherence is the explicit wavefunction collapsing scheme, for examples, the instantaneous decoherence[34], A-FSSH[35], DISH[36], and MF-SD[37]. In these algorithms, the wavefunction will be decomposed into several adiabatic states $\Psi(\mathbf{r},t) = \sum_i \phi_i(\mathbf{r},t)$. Then it will stochastically chose an adiabatic state $\phi_i(\mathbf{r},t)$ to be broken away from the rest of $\Psi(\mathbf{r},t)$, and one will then continue the simulation either with $\phi_i(\mathbf{r},t)$, or with $\Psi(\mathbf{r},t) - \phi_i(\mathbf{r},t)$. Much like in the surface hopping algorithm, the energy conservation requirement during the collapsing restores the detailed balance. Such wavefunction collapsing approaches have been used to study many interesting problems[38–42]. However, there could still be potential issues. The probability of collapsing for $\Psi = \phi_i + \phi_j + \phi_k$ depends on the average coherence time of each adiabatic state to the rest of the adiabatic states. But this is hardly satisfactory. For example, ϕ_k can have a short coherence time with ϕ_i, but a long coherence time with ϕ_j. Thus breaking ϕ_k away from the rest of the wavefunctions will do a disservice to the coherence between ϕ_k and ϕ_j. More deeply, this means the system cannot already be described by a single electronic wavefunction at any given time, without including the phonon wavefunction in an entangled manner. It is simple to solve this issue in a density matrix formalism, where the off-diagonal term D_{ij}, D_{ik}, and D_{jk} all decays differently following their own coherent time. However, to take into account the detailed balance, so far one has to adopt a stochastic solution like surface hopping as discussed above. In contrast, MFE is deterministic so only one trajectory is needed, hence it is computationally efficient. Unfortunately, there is yet to have a method to incorporate both decoherence and detailed balance in a unified deterministic density matrix formalism.

For the original MFE and FSSH, the effect of the wavefunction evolution to the nuclear movement is explicitly included. Such "back reaction" is necessary for small systems like molecules, or cases where the trajectory of the nuclei is the main concern (e.g., in a chemical reaction with branching). But MFE and FSSH can also be combined with the neglect of back-reaction approximation (NBRA)[43], in which the "back reaction" is explicitly ignored (but may be implicitly included by correction terms), and one just takes an average nuclear trajectory, e.g., from the conventional ground state Born-Oppenheimer molecular dynamics (BO-MD). This is a good approximation for many large systems where the nuclear movement will not be dramatically altered by a single hot electron wavefunction, and the focus of the study is not on the nuclear movement, but on the electron dynamics, like the case for hot carrier cooling or transfer in quantum dot, bulk, surface or large molecule. The NBRA brings a considerable computational saving since the trajectory can be pre-calculated using normal MD before doing the electronic dynamics. It has made it possible to calculate electron dynamics for systems with several hundred atoms at the first-principles DFT level. Thus, to achieve efficient and accurate NAMD simulation, it is highly desirable to develop density matrix MFE formalism under NBRA, and including both decoherence and detailed balance.

In this work, we will use the NBRA since the focus of our study is the carrier dynamics. We will modify the conventional density matrix approach, so it can incorporate the detailed balance element in the formalism. Including this effect allows us to study carrier cooling and charge transfer, which are among the most interesting topics under NBRA and for large systems. The resulting approach takes the output of a conventional BO-MD (e.g, under density functional theory, DFT), and calculate the NAMD as a post-process. We will then apply this new formalism (P-Matrix) to study the hot-carrier relaxation and transfer in Si QDs. Although our P-Matrix new formalism describes an ensemble of the carrier dynamics, it does not carry out the calculation using explicit stochastic process like in surface hopping or wavefunction collapsing approach. As a result, it is computationally efficient.

II. METHOD

In the NAMD approach, the single-particle state $\psi_l(t)$, which satisfies the TDSE, is usually expanded by the adiabatic eigenstates $\phi_i(t)$, namely:

$$i\frac{\partial\psi_l(t)}{\partial t} = H(t)\psi_l(t), \tag{1}$$

$$\psi_l(t) = \sum_i C_i^l(t)\phi_i(t), \tag{2}$$

$$H(t)\phi_i(t) = \varepsilon_i(t)\phi_i(t). \tag{3}$$

Here $H(t)$ is the single-electron Hamiltonian, which under NBRA only depends on the nuclear position $\mathbf{R}(t)$. In the density matrix formalism, the density matrix $D_{ij}(t)$ can represent an ensemble of single particle states $\{\psi_l, w_l\}$. Here w_l is the statistical weight of ψ_l. Using $\phi_i(t)$ as the basis, we have:

$$D_{ij}(t) = \sum_l w_l(t) C_i^l(t) C_j^l(t)^*, \tag{4}$$

with $D_{ij}(t) = D_{ji}(t)^*$, and $D_{ii}(t) \geq 0$. It is easy to show that, the time evolution of $D_{ij}(t)$ satisfies the following equation (See SI for details):

$$\begin{aligned} \frac{\partial}{\partial t} D_{ij}(t) =& -i \sum_k [V_{ik}(t) D_{kj}(t) - D_{ik}(t) V_{kj}(t)] \\ =& -i[V, D]_{ij}, \end{aligned} \tag{5}$$

with

$$V_{ij}(t) = \delta_{ij} \varepsilon_i(t) - i \langle \phi_i(t) | \frac{\partial \phi_j(t)}{\partial t} \rangle. \tag{6}$$

The off-diagonal term in $D_{ij}(t)$ represents coherent coupling between adiabatic states ϕ_i and ϕ_j. A straight forward approach commonly used to include the decoherence phenomenon is to introduce a decay term into Eq. 5:[21,44]

$$\frac{\partial}{\partial t} D_{ij}(t) = -i[V, D]_{ij} - (1 - \delta_{ij}) \frac{D_{ij}(t)}{\tau_{ij}(t)}, \tag{7}$$

where τ_{ij} is the pairwise decoherence time. τ_{ij} can be calculated either in advance or on the fly. Here it enters only as a parameter. In principle, any method for calculating τ_{ij} can be used. By approximating the nuclear wavefunction as a product of frozen Gaussian wave packets, Wong and Rossky proposed an instantaneous decoherence time as[44]:

$$\tau_{ij}(t) = [\sum_n \frac{1}{2 a_n \hbar^2} [\mathbf{F}_i^n(t) - \mathbf{F}_j^n(t)]^2]^{-1/2}. \tag{8}$$

Here a_n is the width of the Gaussian wave packets for nuclei, $\mathbf{F}_i^n$ is the Hellmann-Feynman force from state i, and n runs over all phonon modes. Based on Eq. 8 and under thermal equilibrium approximation, we have derived a simplified expression of τ_{ij} (See SI for details):

$$\tau_{ij}(t) = \sqrt{\frac{24 (k_B T)^2}{\langle | \frac{\partial}{\partial t} [\widetilde{\varepsilon}_i(t, t') - \widetilde{\varepsilon}_j(t, t')] |^2 \rangle_{t'}}}, \tag{9}$$

where the brackets $\langle \rangle_{t'}$ indicate the average over $t' < t$, $\widetilde{\varepsilon}_i(t, t') = \langle \phi_i(t) | H(t') | \phi_i(t) \rangle$, and T is the temperature. Our tests show that Eq. 9 and Eq. 8 give the same magnitude of τ_{ij} (Fig. S1). Eq. 9 simply uses the adiabatic state eigenenergy from the BO-MD simulation to calculate the decoherence time between ϕ_i and ϕ_j, thus is suitable for NBRA. However, unlike Eq. 8 which is an instantaneous formula at time t, Eq. 9 requires a time average with t'. To avoid the situation where index of i changes with time for a same characteristic adiabatic state ϕ_i (e.g., due to state crossing), we have used $\phi_i(t)$ to calculate the expectation value of $H(t')$ to define $\widetilde{\varepsilon}_i(t, t')$. As is shown in the SI, this $\widetilde{\varepsilon}_i(t, t')$ can also be calculated from the BO-MD output.

Besides the quantum decoherence, the other major issue of Eq. 5 under NBRA is the lack of detailed balance. Here we propose a correction to this problem. To restore the detailed balance between the i to j transition, and j to i transition, one can multiply the probability for the energy-increasing transition by the Boltzmann factor $\exp(-\frac{\Delta E}{k_B T})$, similar to the FSSH and DISH under the NBRA limit. Therefore, to correct Eq. 5, the key point is to distinguish the energy-increasing and energy-decreasing transitions in the density matrix formalism. The off-diagonal term D_{ij}, and its corresponding term $V_{ji} D_{ij}$ in Eq. 5 describes the occupied state transition between state i and j. Unfortunately, this single $D_{ij} = D_{ji}^*$ term includes both the transitions from i to j, and from j to i. Our key observation is that, we can break it into two terms: $D_{ij} = P_{ij} + P_{ji}^*$ with $P_{ij} \neq P_{ji}^*$, and P_{ij} describes the pumping from state j to i. More specifically, we have their time evolution equation:

$$\frac{\partial}{\partial t} P_{ij} = -i[V, P]_{ij} - i V_{ij} (P_{ii} + P_{jj}^*), \quad (i \neq j) \tag{10}$$

$$\frac{\partial}{\partial t} P_{ii} = -i[V, P]_{ii}. \tag{11}$$

Keeping in mind that $V_{ij} = V_{ji}^*$, one can show that Eqs. 10 and 11 can reproduce the original Eq. 5. Note in the above formula the imaginary part of the diagonal term P_{ii} does not play any role. So we can always enforce P_{ii} to be a real number, then $D_{ii} = 2 P_{ii}$. To understand Eq. 10, we can focus on the terms involving only i and j states at the right side, then it becomes: $\partial P_{ij} / \partial t = -i(\varepsilon_i - \varepsilon_j) P_{ij} - i V_{ij} D_{jj}$, while the first term is a simple phase evolution term, the second term is obviously the term which pumps the charge from state j to state i. The charge change in state i is described by Eq. 11: $\partial D_{ii} / \partial t = -2\mathrm{Re}(i V_{ij} P_{ji}) + 2\mathrm{Re}(i P_{ij} V_{ji})$(the "Re" comes from the fact we only keep the real part of P_{ii}), where the first term represents the lose of charge due to the pumping from i to j, and the second term represents the increase of charge due to the pumping from j to i. Having distinguished the pumping from i to j and from j to i, we can now introduce an empirical constraint to force the detailed balance. If transition from i to j increases the energy, e.g. $\varepsilon_i < \varepsilon_j$ for electrons (or $\varepsilon_i > \varepsilon_j$ for holes), we then multiple $\mathrm{Re}(i P_{ji} V_{ij})$ by a Boltzmann factor $\exp(-\frac{\Delta E}{k_B T})$ so that the i to j transition is suppressed. Meanwhile the $\mathrm{Re}(i P_{ij} V_{ji})$ term will be kept unchanged. Similarly, if j to i transition increases the energy, the Boltzmann factor should be applied to $\mathrm{Re}(i P_{ij} V_{ji})$. This is much like the FSSH or wavefunction collapsing scheme, where if the hoping (or collapsing) causes an electron energy decrease, the event is al-

ways allowed. But if it causes an electron energy increase ΔE, then it depends on the corresponding phonon transition degree of freedom. If this phonon degree of freedom has a kinetic energy larger than ΔE, then this process is allowed, and the kinetic energy will be rescaled. If the kinetic energy is less than ΔE, then this stochastic event will be abandoned. Since the probability for this phonon degree of freedom to have energy larger than ΔE is $\exp(-\frac{\Delta E}{k_B T})$ (assuming a thermal equilibrium), the allowing probability (for this event to happen) should also be proportional to $\exp(-\frac{\Delta E}{k_B T})$. This is exactly the requirement for detailed balance. To put everything together, we finally have:

$$\frac{\partial}{\partial t}P_{ij} = -i[V,P]_{ij} - iV_{ij}(P_{ii}+P^*_{jj}) - \frac{P_{ij}}{\tau_{ij}}, \quad (i\neq j) \quad (12)$$

$$\begin{aligned}\frac{\partial}{\partial t}P_{ii} =& -\mathrm{Re}(i[V,P]_{ii}) \\ &+\sum_j \mathrm{Re}(iP_{ij}V_{ji})f(\Delta\varepsilon_{ij})(\exp(\frac{-|\Delta\varepsilon_{ij}|}{k_B T})-1) \\ &-\sum_j \mathrm{Re}(iP_{ji}V_{ij})(1-f(\Delta\varepsilon_{ij}))(\exp(\frac{-|\Delta\varepsilon_{ij}|}{k_B T})-1).\end{aligned} \quad (13)$$

in which $\Delta\varepsilon_{ij} = \varepsilon_i - \varepsilon_j$ and $f(x) = 1(0)$ for $x > 0$, and $f(x) = 0(1)$ for $x < 0$ for electron (hole).

To carry out Eqs. 12 and 13, the only thing needed is the $V_{ij}(t)$ of Eq. 6. This can be obtained from the conventional BO-MD simulation. In the BO-MD, as implemented in the DFT plane wave pseudopotential code PWmat[45,46], the overlap matrix $\langle\phi_i(t-\Delta t)|\phi_j(t)\rangle$ between two consecutive steps $t-\Delta t$ and t is output for every time step t along with the eigenenergy $\varepsilon_i(t)$. Here, typically Δt is about 1-2 fs. A fixed number of adiabatic states are used in Eq. 2, e.g., 40 for the example studied below. Here is the only place where the calculation might be slightly more expensive than the usual BO-MD simulation since one might need more unoccupied states. These information can then be used to construct a linear interpolating Hamiltonian from $t-\Delta t$ to t.[7,20]. This Hamiltonian will be used to integrate Eqs. 12 and 13 from $t-\Delta t$ to t with a much smaller time step (e.g., as small as 10^{-5} fs). Since this is a small dimension Hamiltonian, the integration is fast. The technical details of the integration is given in the SI. Overall, this postprocess NAMD simulation does not take much time compared to the original BO-MD simulation. Note for pure FSSH with NBRA but without decoherence, the computational cost is similar to our method, because in this case the solution of the TD-SE (eq. 5) is independent of the hopping events, thus many surface hopping trajectories can be calculated with the TD-SE solved only once. However, unlike the pure FSSH-NBRA, when decoherence is included via wave function collapsing (like the DISH method), the integration of the TD-SE is required for each stochastic realization of the surface hopping trajectory, because the wavefunction collapse will affect the solution of TD-SE. The cost of solving the TD-SE increases linearly with the number of realizations of the stochastic process, which can become an issue as discussed in Ref.[47]. Therefore, the P-Matrix method has the advantage in terms of computational cost.

III. SIMULATIONS AND RESULTS

In the following we apply the above method (referred as P-Matrix method) to investigate hot-hole relaxation process in silicon quantum dot systems. Density functional calculations were preformed using the PWmat[45,46], a GPU-based code with plane-wave basis. The planewave cutoff energy is 30 Ry. Test calculations show that this cutoff is sufficient to obtained converged eigenenergies and non-adiabatic coupling coefficients, as presented in the SI. Molecular dynamics (MD) simulations were preformed at ∼300 K under the *NVE* ensemble with the Verlet algorithm. The time step for MD is 1 fs. ∼40 states are used in the valence band to expand the hot carrier wavefunction in Eq. 2. The norm-conserving pseudopotentials[48] and the LDA functional[49] are adopted. Although our formalism can be equally applied to different functional, e.g. LDA or HSE, in actual simulations, the choice of the functional should be considered carefully. For example, a previous study shows LDA and HSE give very different nonadiabatic couplings for Si_7 or Si_{26} clusters.[50] One possible reason is that these systems are not passivated, so there are many highly localized states induced by the dangling bonds. In such cases the LDA will predict much delocalization states compared to HSE results. On the other hand, ref.[50] also showed that for well passivated systems like SiH_4 or Si_2H_6, the difference between LDA and HSE results are quite small, because the wavefunction localization in these systems are provided by the spatial confinement of the QD. Our current fully passivated Si QD belongs to the latter case, thus LDA should be good enough.

We first look at the single silicon QD as shown in Fig. 1a. The constructed QD contains 87 Si atoms, and the surface is passivated by H atoms. At t=0, a hot hole is excited to 0.66 eV below VBM. The population-weighted average energy E_{ave} of the hot hole is calculated for $t > 0$ to study the hot-hole cooling process using P-Matrix. For comparison, we also used Eq. 5 (TDSE-NBRA) and the modified Ehrenfest (ME) algorithm in Ref.[7]. The TDSE-NBRA does not include either the decoherence effect or the detailed balance. The ME method is similar to the FSSH under NBRA approximation. It has the detailed balance correction (DBC) but the decoherence is missing. In ME, the Boltzmann factor is applied in the TDSE to the charge transfer rate between states i and j at every instant of t. This is unlike the P-matrix formalism where the Boltzmann is applied to P_{ij} in Eq. 13, which is a time accumulated quantity (see Eq. 24 in SI). In

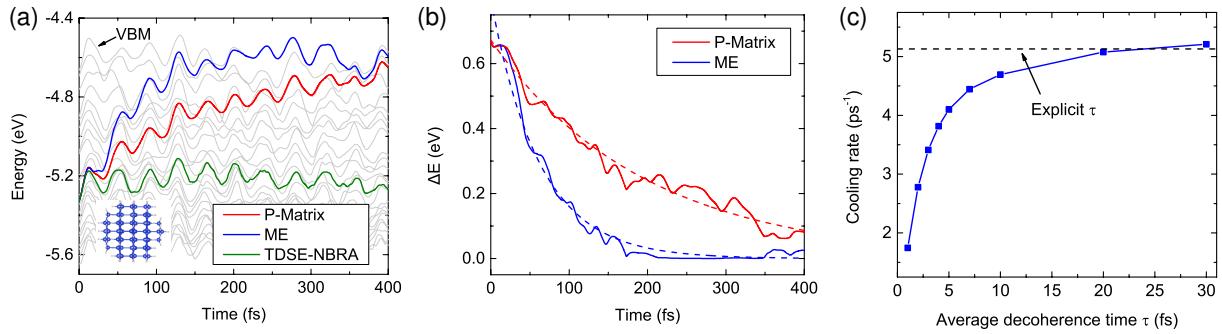

FIG. 1. (a) The population-weighted average energy (colored lines) of the hot hole calculated by the P-Matrix, ME, and ED, as well as the adiabatic eigenenergies (gray lines). The inset shows the structure of the Si quantum dot. (b) The excess energy calculated by P-Matrix and ME. Dashed lines are exponential decay fitting. (c) The hot-hole cooling rates calculated by fixing the decoherence time τ at different values using the P-Matrix method. The dashed line indicates the result obtained using explicit τ.

Fig. 1a it can be seen that, due to the lack of DBC in the TDSE-NBRA, the hot hole stays around 0.6 eV below the VBM and doesn't cool down. In contrast, the calculated $E_{\rm ave}$ from both ME and P-Matrix show clear hot-hole cooling behavior. However, the cooling rates are quite different. The hot-hole relaxation time T_r is calculated by fitting the excess energy $\Delta E(t)$, defined as $\Delta E(t) = E_{\rm VBM}(t) - E_{\rm ave}(t)$, into an exponential decay function $A\cdot\exp(-t/T_r)$. The results are presented in Fig. 1b. ME gives a T_r of 62 fs, whereas P-Matrix gives 195 fs. Test calculations for a larger QD with 175 Si atoms were also performed, and the fitted T_r is 248 fs from the P-Matrix method (See Fig. S6 in the supporting information). Thus further increasing the size will not induce order-of-magnitude change in T_r. In experiment, the hot-carrier relaxation time of Si is determined to be 240-260 fs[51,52]. Other theoretical calculations also suggests that in Si nanostructures the carrier cools down in a few hundred femtoseconds.[53,54] These results are consistent with our P-Matrix result. The reason for the overestimation of the decay rate by ME is the following. In the perturbation treatment of the quantum mechanical transition between two states i and j, the energy conservation is required through the Fermi's Golden rule. When this requirement is satisfied (through another phonon mode energy), the charge transfer from state i to j will accumulate linearly. On the other hand, when this conservation is not satisfied, the charge transfer oscillates as a sinusoidal function (thus averaged transition rate is zero). In the ME (or FSSH treatment for this matter), this charge transfer is treated instantaneously at every time t. For the case of carrier cooling, if i to j charge transfer decreases the energy, every time it is positive, it is accepted 100%. While for negative transfer, it is suppressed by the Boltzmann factor. This leads to a large net charge transfer even when the energy conservation is not satisfied. This problem is avoided in our P-Matrix algorithm. As we show in SI, in P-Matrix, if one ignores the indirect $i \to k \to j$ transition, the transition between i and j does satisfy the Fermi's Golden rule but with a broadening of the δ function by the dephasing time $1/\tau_{ij}$. Besides, our test calculations using a three-level model system show that the P-Matrix method is able to reproduce the Boltzmann distribution at the equilibrium (See SI for details).

For further validation of its accuracy, a direct comparison between the P-Matrix method and other existing schemes with both decoherence and detail balance could be informative. In a recent study, the carrier relaxation times in fluorinated silicon QDs were calculated using FSSH under NBRA, and with the decay-of-mixing scheme for decoherence.[55] The calculated cooling time for a hot electron with an initial energy of ∼1 eV in a $Si_{66}F_{40}$ QD is 493 fs. We took this study as a benchmark, and did the same calculation for the same system using the P-Matrix method. The resulting cooling time is 590 fs, in reasonable agreement with the reported value.

The calculated τ_{ij} for the 87-Si QD using Eq. 9 varies from 7 fs to over 80 fs, with a peak distribution around 20 fs (see Fig. S2 in the SI for the histogram). We noted that in Ref.[56] a simple method is proposed to calculate τ_{ij}, using the standard deviation of the energy gap. This method has similarity with Eq. 9, and both give the same magnitudes of τ_{ij} (see SI for more discussions). Furthermore It would be interesting to see how the decoherence time affects the hot-hole cooling rate. To explore this, we did test calculations using a constant τ for all τ_{ij}. Fig. 1c shows the cooling rate as a function of different τ values. It is seen that as τ increases, the cooling rate $1/T_r$ also increases. This is more significant when τ is small. In a previous study, using a model two-level system, Pradhan et al. showed that the upper bound for the electronic transition rates is proportional to the decoherence time.[57] The reason is that the coherent evolution (accumulation) of electronic states is limited within the decoherence time scale. The observed positive correlation between the cooling rates and the decoherence time here is in agreement with their conclusions. From Fig. 1c, one also found that the cooling rate calculated using the explicit τ_{ij} is similar to that using a constant $\tau \sim 25$

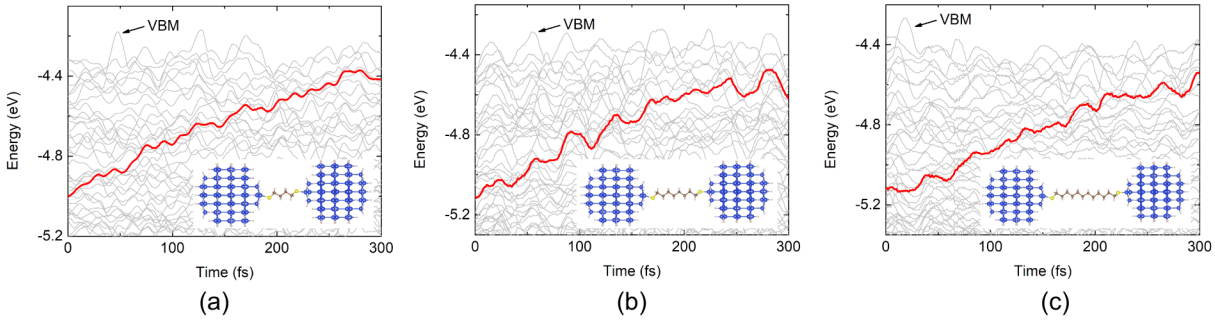

FIG. 2. (a) The average energy and excess energy of the hot hole as a function of time for the n=4 QD pair. Dashed line is exponential fitting. (b)-(c) The same as (a) but for n=8 and n=12 QD pairs, respectively.

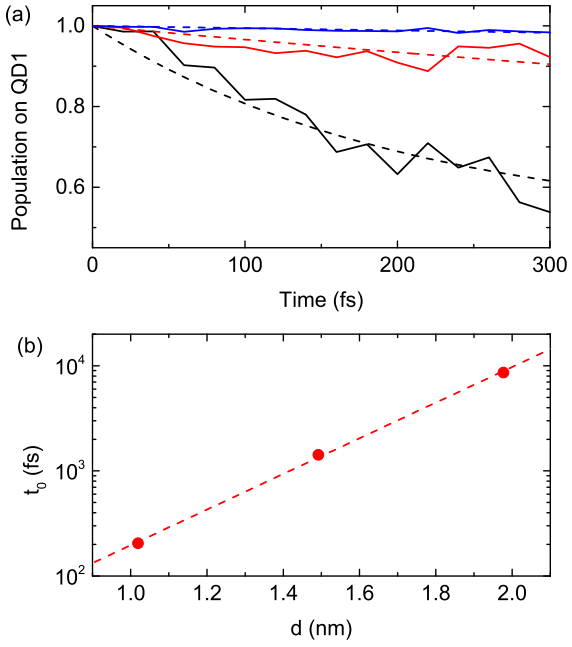

FIG. 3. (a) The hot-hole population on QD1 as a function of time for the three QD-pair systems. (b) The charge transfer time scale t_0 as a function of QD spacing d. Dashed line is exponential fitting.

fs. Thus, 25 fs is the typical time scale for the wavefunction decoherence in the QD studied here. In many other systems the decoherence time scale is similar.[38,39]

Next we investigate the hot carrier cooling and transfer in Si QD-pairs. In the constructed systems, two 87-atom Si QDs (QD1 and QD2) are connected by a -S-$(CH_2)_n$-S- ligand. The distance between the QDs is thus controlled by the C chain length n. Three cases with different QD spacing were studied, namely n=4, 8, and 12 as shown in Fig. 2. Because the interaction between QD1 and QD2 is relatively weak, most of the eigen states in the QD pairs are localized in one particular QD. At t=0, QD1 is excited by placing a hole in a state which is completely localized inside QD1 and ∼0.7 eV below VBM. The charge cooling and transfer are then calculated by P-Matrix. The calculated $E_{\rm ave}$ are shown in Fig. 2a-2c. The fitted relaxation times are: $T_r^{n=4}$=179 fs, $T_r^{n=8}$=200 fs, and $T_r^{n=12}$=180 fs. The energy decay behaviors are quite similar to the case of a single QD, and the relaxation time shows weak dependence on the QD spacing.

For the charge transfer between the two QDs, one can expect that the QD spacing will have great influence. To analysis this process, the charge density $\rho(r)$ of the hot hole is calculated by $\rho(\mathbf{r}) = \sum_{ij} D_{ij}\phi_i(\mathbf{r})\phi_j(\mathbf{r})^*$, and the population of the hot hole in QD1 and QD2 are then determined by $w_{\rm QD1}$=$\int_{V_{\rm QD1}} \rho(\mathbf{r})dV$ and $w_{\rm QD2}$=$\int_{V_{\rm QD2}} \rho(\mathbf{r})dV$. Note $w_{\rm QD1}$+$w_{\rm QD2}$ always equal to 1. As described above, initially $w_{\rm QD1}$=1 and $w_{\rm QD2}$=0. A decrease in $w_{\rm QD1}$ indicates charge transfer from QD1 to QD2, and vice versa. Fig. 3a shows the $w_{\rm QD1}$ as a function of time for the three QD-pair systems. The curves are also fitted into a exponential decay function $0.5\exp(-t/t_0)+0.5$ (as the two QDs have the same structure, one can expect the average population on either QD is 0.5 when $t \to \infty$), in which t_0 gives a typical time scale of the charge transfer. For n=4, 8, and 12, t_0=205 fs, 1423 fs, and 8649 fs, respectively. Therefore, the charge transfer rate decreases as the QD spacing increases. Experimentally, the charge transfer time between quantum dots is found to be sensitive to the QD size.[58,59] For instance, in CdSe/TiO_2 systems, the charge transfer time decreases by three orders of magnitude (from ∼100 ns to ∼0.1 ns) when the size of CdSe QD decrease from 7.5 nm to 2.4 nm.[58] In our simulation the diameter of the Si QD is only ∼1.3 nm, which is much smaller than 2.4 nm. In this regard, our resulting (sub-)10 ps charge transfer time should be reasonable when compared with experiments[58,59].

As discussed, the hot-hole relaxation time T_r for the Si-QDs is ∼200 fs. Hence, when the QD spacing is large (n=12), the hot-hole in QD1 will lose its excess energy before it transfers to QD2. When the QD spacing decreases, the hot-hole transfer become more efficient. In the case of n=4 (QD spacing is 1 nm), t_0 is quite close to T_r, indicating that excitation in QD1 is able to create significant amount of "hot" carrier in QD2. It is also interesting to explore the relationship between the QD

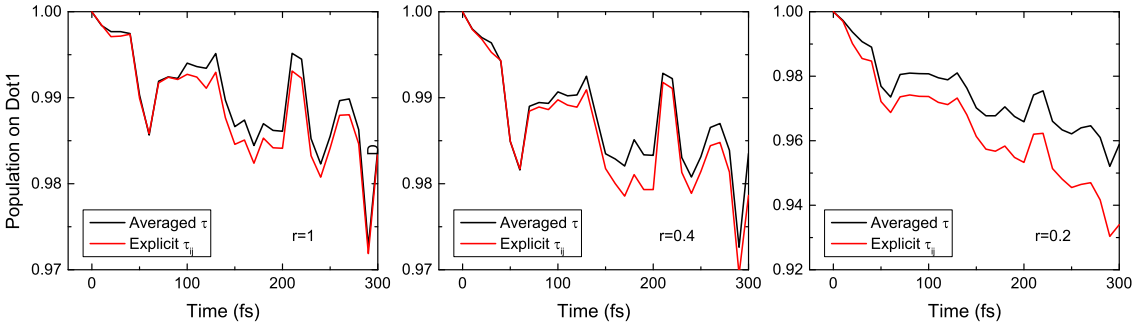

FIG. 4. The charge transfer rate for the n=12 QD-pair case using both the explicit calculated τ_{ij} and the averaged τ, with different reducing factor r applied to the decoherence times.

spacing d and t_0. As presented in Fig. 3b, t_0 scales exponentially with d, and their relationship can be fitted to t_0=A·exp(d/d_0). The fitted A and d_0 are 3.95 fs and 0.26 nm, respectively. The carrier transfer in the QD pair system can be understood by the state coupling when their energies anticrossing each other[7]. The coupling strength decays exponentially as a function of the QD-QD distance as shown in Fig. 3b. The rate of charge transfer is a competition between the transfer rate and the internal QD cooling as lower energy (below VBM) region has higher density of state and also tend to have larger coupling constant (e.g., below the ligand molecule HOMO level).

Finally we'd like to have more discussions on the effect of decoherence time τ_{ij}. As mentioned, one advantage of the P-Matrix is that the decoherence between different pairs of electronic states can be treated independently. For example, in the Si QD-pair systems, there could be two different types of τ_{ij}. When both states i and j are localized in the same QD, their energy fluctuations have similar trends since they are both affected by the atomic vibration of the same QD, especially for the surface atom vibrations which significantly alter the inner QD potentials. Hence the τ_{ij} value in this case could be relatively large. In contrast, when i and j belong to different QDs, their energy fluctuations can be quite different, resulting in a relatively small τ_{ij} according to Eq. 9. In the wavefunction collapsing scheme, for each state there is only one associated average decoherence time, and there is no distinguish between intra- and inter-QD decoherence. Such a effect is especially significant when the average τ is small. To demonstrate this point, we have calculated the charge transfer rate for the n=12 QD-pair case using both the explicit calculated τ_{ij} and a fixed average τ (the rate $1/\tau$ equals to the average rate of $1/\tau_{ij}$). To see how the magnitude of decoherence time affects the trend, we also rescale τ_{ij} and τ using a factor r. The results for different r are shown in Fig. 4. It is seen that although with r=1, the difference using τ_{ij} and averaged τ are small, with r=0.2, the difference becomes significant (in this case the averaged τ is 3.2 fs). Thus, the explicit treatment of decoherence time is important, especially when the decoherence time is short (several femtoseconds).

IV. CONCLUSIONS

In conclusion, using the the density matrix representation, we have developed a new NAMD formalism under the NBRA approximation incorporating both decoherence and detailed balance effects. In this formalism, the decoherence between different pairs of electronic states are treated independently. The density matrix is divided into two parts so that one can distinguish the energy-increasing and energy-decreasing adiabatic state transitions. The detailed balance correction is then included by a Boltzmann factor applied to the energy-increasing transitions. This formalism overcomes the lack of decoherence and overestimation of the cooling rate in the ME (and possibly FSSH), it also overcomes the pitfall of the wavefunction collapsing where the decoherence time τ_{ij} cannot be treated independently. Computationally, the new P-Matrix method is also inexpensive. The hot-hole cooling and charge transfer processes in Si QD systems are investigated using the proposed method. The calculated hot carrier relaxation time is consistent with experiments. In the QD-pair systems, the hot-hole cooling time is almost independent with the QD spacing, while the charge transfer rate between QDs decreases exponentially as the QD spacing increases. It is also shown that the explicit treatment of decoherence time is important to accurately predict the charge transfer rate.

ACKNOWLEDGMENTS

This work was supported by the Director, Office of Science, the Office of Basic Energy Sciences (BES), Materials Sciences and Engineering (MSE) Division of the U.S. Department of Energy (DOE) through the theory of material (KC2301) program under Contract No. DE-AC02-05CH11231. It used resources of the National Energy Research Scientific Computing Center and the Oak Ridge Leadership Computing Facility through the INCITE project.

* lwwang@lbl.gov

[1] B. F. E. Curchod and T. J. Martinez, Chem. Rev. **118**, 3305 (2018).
[2] R. Crespo-Otero and M. Barbatti, Chem. Rev. **118**, 7026 (2018).
[3] S. V. Kilina, D. S. Kilin, and O. V. Prezhdo, ACS Nano **3**, 93 (2009).
[4] M. E. Madjet, G. R. Berdiyorov, F. El-Mellouhi, F. H. Alharbi, A. V. Akimov, and S. Kais, J. Phys. Chem. Lett. **8**, 4439 (2017).
[5] C. Zhao, Q. Zheng, J. Wu, and J. Zhao, Phys. Rev. B **96**, 134308 (2017).
[6] R. Long, J. Liu, and O. V. Prezhdo, J. Am. Chem. Soc. **138**, 3884 (2016).
[7] J. Ren, N. Vukmirović, and L.-W. Wang, Phys. Rev. B **87**, 205117 (2013).
[8] W. R. Duncan, W. M. Stier, and O. V. Prezhdo, J. Am. Chem. Soc. **127**, 7941 (2005).
[9] Q. Zheng, W. A. Saidi, Y. Xie, Z. Lan, O. V. Prezhdo, H. Petek, and J. Zhao, Nano Lett. **17**, 6435 (2017).
[10] S. Giannini, A. Carof, and J. Blumberger, J. Phys. Chem. Lett. **9**, 3116 (2018).
[11] R. Kapral and G. Ciccotti, J. Chem. Phys. **110**, 8919 (1999).
[12] Q. Shi and E. Geva, J. Chem. Phys. **121**, 3393 (2004).
[13] E. Mulvihill, A. Schubert, X. Sun, B. D. Dunietz, and E. Geva, J. Chem. Phys. **150**, 034101 (2019).
[14] A. Kelly and T. E. Markland, J. Chem. Phys. **139**, 014104 (2013).
[15] A. Kelly, N. Brackbill, and T. E. Markland, J. Chem. Phys. **142**, 094110 (2015).
[16] R. Dann, A. Levy, and R. Kosloff, Phys. Rev. A **98**, 052129 (2018).
[17] A. McLachlan, Mol. Phys. **8**, 39 (1964).
[18] D. A. Micha, J. Chem. Phys. **78**, 7138 (1983).
[19] H.-D. Meyer and W. H. Miller, J. Chem. Phys. **72**, 2272 (1980).
[20] Z. Wang, S.-S. Li, and L.-W. Wang, Phys. Rev. Lett. **114**, 063004 (2015).
[21] J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).
[22] J. C. Tully and R. K. Preston, J. Chem. Phys. **55**, 562 (1971).
[23] K. Drukker, J. Comput. Phys. **153**, 225 (1999).
[24] J. C. Tully, J. Chem. Phys. **137**, 22A301 (2012).
[25] P. V. Parandekar and J. C. Tully, J. Chem. Phys. **122**, 094102 (2005).
[26] P. V. Parandekar and J. C. Tully, J. Chem. Theory Comput. **2**, 229 (2006).
[27] A. Bastida, C. Cruz, J. Zuniga, A. Requena, and B. Miguel, Chem. Phys. Lett. **417**, 53 (2006), ISSN 0009-2614.
[28] S. J. Cotton and W. H. Miller, J. Phys. Chem. A **117**, 7190 (2013).
[29] N. Bellonzi, A. Jain, and J. E. Subotnik, J. Chem. Phys. **144**, 154110 (2016).
[30] J. R. Schmidt, P. V. Parandekar, and J. C. Tully, J. Chem. Phys. **129**, 044104 (2008).
[31] A. V. Akimov, R. Long, and O. V. Prezhdo, J. Chem. Phys. **140**, 194107 (2014).
[32] C. Zhu, A. W. Jasper, and D. G. Truhlar, J. Chem. Phys. **120**, 5543 (2004).
[33] G. Granucci and M. Persico, J. Chem. Phys. **126**, 134114 (2007).
[34] T. Nelson, S. Fernandez-Alberti, A. E. Roitberg, and S. Tretiak, J. Chem. Phys. **138**, 224111 (2013).
[35] B. R. Landry and J. E. Subotnik, J. Chem. Phys. **137**, 22A513 (2012).
[36] H. M. Jaeger, S. Fischer, and O. V. Prezhdo, J. Chem. Phys. **137**, 22A545 (2012).
[37] M. J. Bedard-Hearn, R. E. Larsen, and B. J. Schwartz, J. Chem. Phys. **123**, 234106 (2005).
[38] R. Long and O. V. Prezhdo, Nano Lett. **16**, 1996 (2016).
[39] W. Li, J. Liu, F.-Q. Bai, H.-X. Zhang, and O. V. Prezhdo, ACS Energy Lett. **2**, 1270 (2017).
[40] L. Li, R. Long, T. Bertolini, and O. V. Prezhdo, Nano Lett. **17**, 7962 (2017).
[41] R. Long, W. Fang, and A. V. Akimov, J. Phys. Chem. Lett. **7**, 653 (2016).
[42] Z. Zhang, W.-H. Fang, M. V. Tokina, R. Long, and O. V. Prezhdo, Nano Lett. **18**, 2459 (2018).
[43] A. V. Akimov and O. V. Prezhdo, J. Chem. Theory Comput. **9**, 4959 (2013).
[44] K. F. Wong and P. J. Rossky, J. Chem. Phys. **116**, 8429 (2002).
[45] W. Jia, Z. Cao, L. Wang, J. Fu, X. Chi, W. Gao, and L.-W. Wang, Comput. Phys. Commun. **184**, 9 (2013).
[46] W. Jia, J. Fu, Z. Cao, L. Wang, X. Chi, W. Gao, and L.-W. Wang, J. Comput. Phys. **251**, 102 (2013).
[47] A. V. Akimov and O. V. Prezhdo, J. Chem. Theory Comput. **10**, 789 (2014).
[48] M. Fuchs and M. Scheffler, Comput. Phys. Commun. **119**, 67 (1999).
[49] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[50] Y. Lin and A. V. Akimov, J. Phys. Chem. A **120**, 9028 (2016).
[51] A. J. Sabbah and D. M. Riffe, Phys. Rev. B **66**, 165217 (2002).
[52] T. Ichibayashi and K. Tanimura, Phys. Rev. Lett. **102**, 087403 (2009).
[53] K. G. Reeves, A. Schleife, A. A. Correa, and Y. Kanai, Nano Lett. **15**, 6429 (2015).
[54] J. Li, Y.-M. Niquet, and C. Delerue, Phys. Rev. B **95**, 205401 (2017).
[55] J. C. Wong, L. Li, and Y. Kanai, J. Phys. Chem. C **122**, 29526 (2018).
[56] A. V. Akimov and O. V. Prezhdo, J. Phys. Chem. Lett. **4**, 3857 (2013).
[57] E. Pradhan, R. J. Magyar, and A. V. Akimov, Phys. Chem. Chem. Phys. **18**, 32466 (2016).
[58] I. Robel, M. Kuno, and P. V. Kamat, J. Am. Chem. Soc. **129**, 4136 (2007).
[59] D. A. Hines, R. P. Forrest, S. A. Corcelli, and P. V. Kamat, J. Phys. Chem. B **119**, 7439 (2015).