Flexible boundary layer using exchange for embedding theories. I. Theory and implementation

Zhuofan Shen[1, 2] and William J. Glover[1, 2, 3, a)]
[1)] *NYU Shanghai, 1555 Century Ave., Shanghai, 200122, China*
[2)] *NYU-ECNU Center for Computational Chemistry at NYU Shanghai, 3663 Zhongshan Road North, Shanghai, 200062, China*
[3)] *Department of Chemistry, New York University, New York, NY, 10003, USA*

Embedding theory is a powerful computational chemistry approach to exploring the electronic structure and dynamics of complex systems, with QM/MM being the prime example. A challenge arises when trying to apply embedding methodology to systems with diffusible particles, *e.g.* solvents, if some of them must be included in the QM region, for example in the description of solvent-supported electronic states or reactions involving proton transfer or charge-transfer-to-solvent: without a special treatment, inter-diffusion of QM and MM particles will lead eventually to a loss of QM/MM separation. We have developed a new method called Flexible Boundary Layer using Exchange (FlexiBLE) that solves the problem by adding a biasing potential to the system that closely maintains QM/MM separation. The method rigorously preserves ensemble averages by leveraging their invariance to exchange of identical particles. With a careful choice of the biasing potential, and the use of a tree algorithm to include only important QM and MM exchanges, we find the method has an MM-forcefield-like computational cost and thus adds negligible overhead to a QM/MM simulation. Furthermore, we show that molecular dynamics with the FlexiBLE bias conserves total energy and remarkably, dynamical quantities in the QM region are unaffected by the applied bias. FlexiBLE thus widens the range of chemistry that can be studied with embedding theory.

I. INTRODUCTION

A common strategy to modelling reactions in complex systems is to use an embedding approach, in which a chemically active region is treated with a high level of theory (*e.g. ab initio* electronic structure) and the remainder of the system is treated at a lower level of theory. The most widely used embedding method, and the subject of our first application, is QM/MM, which invokes a Quantum Mechanical (QM) treatment of the active region with the remainder of the system treated with Molecular Mechanics (MM) forcefields;[1–22] however, QM-in-QM embedding has also seen significant interest in recent years.[23–26]

QM/MM has found broad applications ranging from enzymology, structural biology, materials science, and spectroscopy, among others.[21,27–33] Another important use of QM/MM is in describing solution-phase chemistry,[29,34,35] since a solvent environment lends itself naturally to an MM description, with the solute treated at a QM level. This approach can be combined powerfully with *ab initio* molecular dynamics for a first-principles description of reactivity in complex systems.[6,29,32,36,37] However, a problem arises when applying embedding theory to a system of diffusible particles, such as a solvent, if some of them are treated at the QM level: a QM/MM boundary must be made between identical molecules, and without special techniques, the QM and MM molecules will inter-diffuse, leading to a loss of QM/MM partitioning. Of course, one straightforward way to avoid this issue is to treat the solvent entirely at the MM level; however, this is not always possible, for example if the chemical process of interest involves proton or charge transfer with the solvent. Other solvated systems that pose a challenge for traditional QM/MM approaches are solvent-supported electronic states, a prime example of which are solvated electrons, which are excess electrons embedded in a liquid solvent.[38–41]

There are two broad classes of approaches to solving the boundary issue in QM/MM with diffusible particles: Adaptive QM/MM and Constrained QM/MM. In the former approach, the treatment of particles changes dynamically between QM and MM descriptions as the molecule traverses the boundary. This is usually accomplished by an interpolation of QM and MM energies and/or forces between the regions, and several such methods have been developed, differing in how the boundary is defined and how the interpolation is achieved.[3,7–9,11,15,17,19,22] Adaptive methods have the advantages that the number of QM molecules need not be conserved and diffusional dynamics are captured. However, it is known that current adaptive QM/MM methods suffer from structural and dynamical artefacts at the boundary due to a mismatch of QM and MM interactions, which must somehow be corrected.[22] High quality MM forcefields, such as MBPol, show promise at overcoming this QM/MM mismatch issue.[42]

In contrast to Adaptive QM/MM, Constrained QM/MM methods use a fixed definition of QM and MM atoms and apply some form of bias to the system to maintain their separation. Several methods fall under this category, including the method we develop below. To our knowledge, the first Constrained QM/MM method was Flexible Inner Region Ensemble Separator (FIRES), introduced by Rowley and Roux.[10] Based on a rigorous

[a)] Electronic mail: william.glover@nyu.edu

partitioning of the system's configurational integral into an inner and outer region,[43] FIRES adds a half-harmonic repulsion potential to MM molecules at the boundary, whose location dynamically adjusts based on the outermost QM molecule. The magnitude of bias required to maintain QM/MM separation and preserve ensemble-averaged structural quantities was however found to be large, and the authors cautioned against using FIRES with a large QM region or for the prediction of dynamical quantities.[10] Shiga and Masia developed the Boundary based on Exchange Symmetry Theory (BEST) method which uses a carefully chosen bias potential to achieve QM/MM separation while preserving ensemble average quantities.[13,14] However, similar to FIRES, a large magnitude of bias was required to maintain QM/MM separation, necessitating the use of small timesteps of 0.25 fs. These authors later developed Quasi-boundary based on Exchange Symmetry Theory (QUEST), which corrects for violations of exchange symmetry due to differences in the QM and MM potentials, at the expense of additional QM calculations on the exchanged particle configurations.[44] They found these corrections improved the predictions of dynamical properties for a toy system compared to BEST; however, a QM/MM implemention of QUEST has not yet been described. The Boundary Constraint with Correction (BCC) method, developed by Takahashi and co-workers, applies a post-processing correction to constrained QM/MM simulations to remove effects of the bias from equilibrium properties, when the bias potential itself is not constructed to preserve equilibrium properties;[20] however, no correction for dynamical properties was developed.

Based on the current status of constrained QM/MM methods, it is evident that there is room for improvement, in particular with respect to reducing the magnitude of required bias so that longer timesteps can be used and dynamical quantities are minimally affected. We propose that any new constrained QM/MM method should ideally posses the following five qualities: 1) closely maintain QM/MM separation (some small violations of separation, such as found in FIRES and BEST are acceptable). 2) Rigorously preserve ensemble averaged quantities, at least in the limit that the QM and MM potentials are identical. 3) Not introduce potential energy surface discontinuities or large bias forces so that molecular dynamics can be propagated with a 1.0 fs timestep for liquid water while conserving total energy. 4) Introduce bias forces that are localized only to molecules close to the QM/MM boundary, thereby leaving dynamical quantities in the inner QM region unperturbed, at least on sub-diffusional timescales. 5) Contribute negligible computational overhead compared to the cost of QM/MM electronic structure calculations.

Building on the formalism introduced in BEST,[13] in this work we develop the Flexible Boundary Layer using Exchange (FlexiBLE) method. The key idea is a construction of the biasing force that acts, in principle, on every QM and MM pair and all their possible combinatorial exchanges, allowing for ensemble averages to be rigorously maintained, even with some degree of QM/MM mixing. With a careful choice of biasing potential, however, the biasing forces can be truncated outside a narrow boundary layer of QM and MM molecules, and the surviving terms can be efficiently enumerated with a tree algorithm. As a result, the influence of the bias is highly localized to the boundary region. Using one-electron mixed/quantum classical (MQC) simulations of the aqueous solvated electron, $e^-_{(\mathrm{aq})}$,[45,46] as a benchmark, we demonstrate that FlexiBLE satisfies the five qualities set forth above. In the companion paper, we use FlexiBLE to build an *ab initio* many-electron QM/MM model of $e^-_{(\mathrm{aq})}$ to explore its structural, electronic, and dynamical properties.

The remainder of the paper is as follows. In section II A, we briefly review the formalism of FIRES to provide context for other constrained QM/MM methods. In section II B, we then describe BEST and its relation to FIRES. Our FlexiBLE method is introduced in section II C. In section III, we benchmark the methods on equilibrium and dynamical properties of a MQC description of the hydrated electron. Finally, conclusions are drawn in Section IV.

II. THEORY

A. FIRES

To provide the context for our development of FlexiBLE, we first briefly review the FIRES method of Rowley and Roux,[10] since we shall see connections to the BEST and FlexiBLE methods discussed later. We start by considering the classical configurational integral of a homogeneous solvent of N molecules:

$$Z = \frac{1}{N!}\int d\mathbf{R}_1 \int d\mathbf{R}_2 \cdots \int d\mathbf{R}_N \exp(-\beta V(\mathbf{R})), \quad (1)$$

where $V(\mathbf{R})$ is the system's total potential energy, $\beta = 1/k_{\mathrm{B}}T$ is the inverse of the temperature multiplied by Boltzmann's constant, $\mathbf{R}_i$ is a configuration of solvent i's atomic positions, and $\mathbf{R} = (\cdots, \mathbf{R}_i, \cdots, \mathbf{R}_j, \cdots)$ collects the entire system's atomic positions. The $N!$ denominator accounts for the indistinguishability of the solvent molecules. Additional integrals over solute degrees of freedom can be included without loss of generality.[10]

Next, the solvent molecules are partitioned into "inner" and "outer" regions, where the inner solvents are defined as the n closest molecules to a particular coordinate, which could be a solute atom, or a collective coordinate such as the system's center of mass. The remaining $N - n$ solvent molecules are considered to be in the outer region (see Fig. 1 of Ref. 10). With this partitioning, and building on formalism introduced in the mixed explicit-implicit simulation method called Spherical Solvent Boundary Potential,[43] Rowley and Roux showed

that the configurational integral could be re-expressed as:

$$Z = \frac{1}{n!}\int d\mathbf{R}_1 \int d\mathbf{R}_2 \cdots \int d\mathbf{R}_n \frac{1}{(N-n)!} \int' d\mathbf{R}_{n+1} \cdots \int' d\mathbf{R}_N \exp(-\beta V(\mathbf{R})), \quad (2)$$

where the prime indicates that the integral is performed over the region of space outside the outermost of the n inner solvent molecules. This implies that a simulation can sample Eq. 2 by enforcing the constraint that the inner and outer solvent molecules are not allowed to exchange and swap. The exact equivalence between Eq. 1 and 2 then guarantees that correct thermodynamic averages will be obtained from such a constrained simulation. The potential energy function can then be computed by treating the inner solvent molecules with QM electronic structure, and the outer solvent molecules with MM forcefields. Eq. 2 provides the theoretical support for computing ensemble averages while maintaining QM/MM separation.[10]

To render practical the sampling of Eq. 2 in a Molecular Dynamics (MD) simulation, Rowley and Roux added to the system Hamiltonian a repulsive half-harmonic restraining potential:

$$V^{\mathrm{FIRES}} = \frac{1}{2}k^{\mathrm{FIRES}} \sum_{i=n+1}^{N} (R_{\mathrm{inner}} - r_i)^2 \Theta(R_{\mathrm{inner}} - r_i), \quad (3)$$

where r_i is the radial distance of outer solvent molecule i from the center of the inner region (taken with respect to one of solvent i's atoms, or its center of mass), k^{FIRES} is the force constant of the half-harmonic restraint, $\Theta(x)$ is the Heaviside step function, and R_{inner} represents the radial distance of the outermost inner solvent molecule:

$$R_{\mathrm{inner}} = \max(r_1, r_2, \cdots, r_n). \quad (4)$$

The potential of Eq. 3 allows the constraint of no mixing between inner and outer solvent molecules to be (approximately) satisfied in an MD simulation,[10] without the complication of enforcing hard-wall reflecting boundary conditions between the outer and inner molecules.[47]

FIRES has some very attractive features: the theoretical foundation provided by Eq. 2 is particularly transparent, and the restraining potential of Eq. 3 is straightforward to implement in an MD package. However, Eqs. 3 and 4 introduce two possible issues: 1) for finite k^{FIRES} and non-zero temperature, the outer (MM) solvent particles will, to some extent, cross the boundary defined by the outermost inner (QM) solvent particle, since they experience only a linear restoring force upon entering the QM region. This (small amount of) QM/MM exchange then violates Eq. 2, and Jónsson and co-workers showed very recently that this gives rise to artifacts at the QM/MM boundary,[47] explaining similar structural problems observed earlier by Bulo and co-workers.[12] 2) Eq. 4 introduces derivative discontinuities in the potential when the outermost QM particle's identity changes simultaneously with an MM particle being inside the QM region. Although one should expect this to be a relatively rare occurrence, the introduction of potential energy derivative discontinuities means that FIRES will violate energy conservation. The extent of the problem will increase with the surface area of the QM/MM boundary, and therefore with the number of QM particles.

Both problems of violation of QM/MM separation and derivative discontinuities can be minimized by choosing a sufficiently large value of k^{FIRES}; however, what "sufficiently large" means is expected to be system and QM size specific. Furthermore, increasing k^{FIRES} comes at the cost of introducing larger QM/MM restoring forces, necessitating a smaller MD timestep. Whether a balance can be found, with a k^{FIRES} large enough to avoid boundary artifacts, but small enough to allow a 1.0-fs timestep is an open question. Unfortunately, for the hydrated electron system described below, we were unable to find such a balance. This motivated us to consider other constrained QM/MM methods.

As discussed above, structural artifacts observed in FIRES arise from violations of the formal constraint of no exchanges between QM and MM particles (Eq. 2). One approach to avoid artifacts is to then rigorously enforce QM/MM separation by imposing boundary conditions between the QM and MM particles to prevent them from exchanging. Shortly after the initial submission of our manuscript, a method that achieves this was published: the scattering-adapted (SA-FIRES) approach of Jónsson and co-workers.[47] This method replaces the half-harmonic repulsion potential of FIRES with binary elastic collisions between QM and MM particles at the boundary, using a variable timestep algorithm. Hard-wall reflective boundary conditions were also shown to work but required a larger number of timesteps.[47] SA-FIRES is a very promising approach to solving the QM/MM boundary issue; however, the requirement of a variable timestep introduces significant complexity in the MD integrator and would furthermore add computational expense when combined with QM/MM ab initio molecular dynamics (AIMD) that requires electronic structure calculations at each timestep. Incorporating a multi-timestep integrator in SA-FIRES might ameliorate the additional computational expense.

A simpler approach to avoiding artifacts at the QM/MM boundary is to allow some exchanges between QM and MM solvent molecules, but in a way that closely maintains QM/MM separation while leaving ensemble averages rigorously unchanged. This is precisely what the BEST method seeks to achieve,[13] so we consider it next.

B. BEST

The starting point of BEST is the recognition that an ensemble average is invariant to exchanges of like particles. Thus, by treating QM and MM solvent particles as identical (necessarily an approximation unless the MM potential perfectly mimics the QM potential), one can perform a weighted ensemble average wherein configurations that have MM and QM mixing (exchanges) are included with a low weight, whereas configurations that have MM and QM separation are included with high weight. In particular, consider the canonical ensemble average of observable X:

$$\langle X \rangle = \frac{\int X(\mathbf{R}) \exp(-\beta V(\mathbf{R}))\, d\mathbf{R}}{\int \exp(-\beta V(\mathbf{R}))\, d\mathbf{R}}. \tag{5}$$

In Eq. 5, configurations with an exchange of QM and MM particles are averaged with equal weights. We could instead add their contributions to the ensemble average with different weights by introducing a normalized bias function with the following properties:

$$f_{ijk\cdots} \equiv f(\cdots, \mathbf{R}_i, \cdots, \mathbf{R}_j, \cdots, \mathbf{R}_k, \cdots) \geq 0 \tag{6}$$

$$\sum_L \hat{P}_L(f_{ijk\cdots}) = 1, \tag{7}$$

where $\hat{P}_L$ is an operator that permutes the indices $ijk\cdots$ and the second equality above ensures that f is normalized after summing over all possible QM and MM exchanges. Shiga and Masia showed that with a bias function that satisfies Eqs. 6 and 7, and using the permutational symmetry of X and V, the ensemble average is formally unaffected by the inclusion of the bias function:[13]

$$\langle X \rangle = \frac{\int X \exp(-\beta V) f_{ijk\cdots}\, d\mathbf{R}}{\int \exp(-\beta V) f_{ijk\cdots}\, d\mathbf{R}} \tag{8}$$

$$= \frac{\int X \exp(-\beta \tilde{V})\, d\mathbf{R}}{\int \exp(-\beta \tilde{V})\, d\mathbf{R}} \tag{9}$$

where the second equality results from incorporating the bias function with an added bias potential:

$$\tilde{V}_{ijk\cdots} = V + V^{\text{bias}}_{ijk\cdots} \tag{10}$$

$$V^{\text{bias}}_{ijk\cdots} = -k_{\text{B}} T \log f_{ijk\cdots} \tag{11}$$

To ensure normalization, the bias function can be written as:

$$f_{ijk\cdots} = \frac{h_{ijk\cdots}}{\sum_L \hat{P}_L(h_{ijk\cdots})}, \tag{12}$$

where $h_{ijk\cdots}$ is an unnormalized penalty function, and the denominator of Eq. 12 is a normalization factor that sums the penalty function over all possible QM and MM particle exchanges, resulting in $(N_{\text{QM}} + N_{\text{MM}})!/(N_{\text{QM}}! N_{\text{MM}}!)$ unique terms.

Eq. 9 holds true for any bias function satisfying Eqs. 6 and 12, and this provides a great deal of flexibility in the choice of penalty, h, in Eq. 12, which we will take advantage of below in our formulation of FlexiBLE. To maintain QM/MM separation, h should be non-zero for configurations with QM/MM separation and vanishingly small for configurations with QM/MM mixing. In this work, following FIRES[10] and BEST,[13] we take the QM region to be a sphere centered around the origin of the system, with MM particles outside the sphere. Generalizations to other QM geometries are possible.[13] Perfect QM/MM separation is then achieved when the radial distance from the origin to any QM particle is always less than the radial distance from the origin to any MM particle. Shiga and Masia proposed the following penalty function to bias the ensemble towards QM/MM separation:[13]

$$h_{ij\cdots,kl\cdots} = \prod_{m=i,j,\cdots}^{N_{QM}} \prod_{n=k,l,\cdots}^{N_{MM}} g_{mn} \tag{13}$$

$$g_{ik} = \begin{cases} 1, & x_i < x_k \\ \exp(-\alpha(x_i - x_k)), & x_i \geq x_k, \end{cases} \tag{14}$$

where x_i and x_k are the radial distances from the origin of the QM and MM particles respectively, and g_{ik} is a pair function of QM and MM distances that is unity for QM particles closer to the origin than MM particles (thus favoring QM/MM separation), and decays exponentially to zero with the distance of any QM particle further from the origin than an MM particle (thus biasing against QM/MM mixing). The exponent parameter, α, controls the rate of decay of bias with QM-MM distance: larger values of α achieve better QM/MM separation, although as we show below, at the expense of requiring a small MD timestep.

While Eq. 13 formally applies a biasing force to every QM and MM particle, and therefore one might worry that their dynamics are strongly influenced by the bias, the product-of-exponentials form of the penalty function means that for a large enough α, the dominant bias forces will be imposed only on a single pair corresponding to the furthest QM and closest MM particles. This motivated Shiga and Masia to develop a Single Exchange (SE) approximation to the bias function, wherein Eq. 13 involves only the pair function between the furthest QM and closest MM particles, which after normalization via Eq. 12 leads to the following bias function:[13,14]

$$f^{\text{SE}}_{i,k} = \frac{g_{ik}}{g_{ik} + g_{ki}}$$

$$= \frac{1}{1 + \exp(+\alpha(x_i - x_k))}, \tag{15}$$

where i and k index the furthest QM and closest MM particles respectively.

The pair function and resulting bias potential of BEST in the SE approximation are shown as the black and red curves respectively in Fig. 1(a). Here it is seen that in the

limit of large QM and MM separation, the bias potential tends to a linear increasing function of distance, meaning a constant bias force is applied to the furthest QM and closest MM particles to restore QM/MM separation. The magnitude of the restoring force is proportional to α, which explains why BEST in the SE approximation can fail to maintain QM/MM separation for small α.[13]

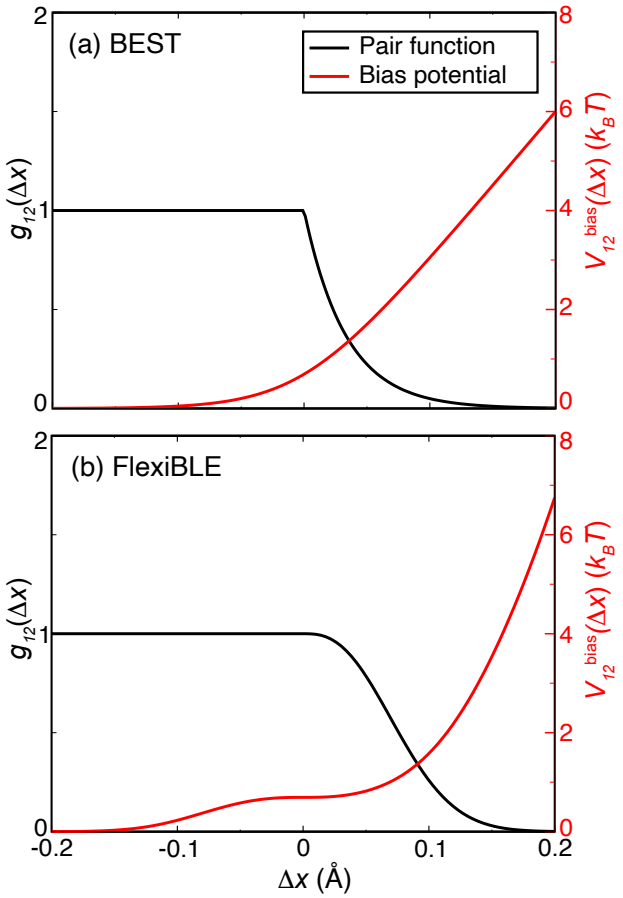

FIG. 1. Pair functions (black curves, left axis) and bias potentials (red curves, right axis) for a single QM and MM pair as a function of their displacement difference, $\Delta x = x_i - x_k$. Panel (a) BEST (Eq. 14) with $\alpha = 30$ Å^{-1}, panel (b) FlexiBLE (Eq. 18) with $\alpha = 15$ Å^{-1}.

The SE approximation has the advantage that it avoids the formally factorially scaling number of exchanges required by Eq. 12, and the approximation can be made better by increasing the value of α in Eq. 15. α can thus be related to $k^{\rm FIRES}$ of Eq. 3. In fact, in the limits $\alpha \to \infty$ and $k^{\rm FIRES} \to \infty$, SE-BEST and FIRES (and the FlexiBLE method we develop below) become equivalent to SA-FIRES,[47] and rigorously achieve QM/MM separation. However, unlike FIRES, SE-BEST allows a finite α to be used while rigorously preserving ensemble averages, if α is large enough. As in FIRES, a balance must be found: we would like α to be small enough to allow a 1.0-fs timestep, but large enough to make the SE approximation sufficiently accurate. The balance, if achievable, is again expected to be system and QM-size specific. Unfortunately, for our hydrated electron system described below, we were unable to find such a balance.

Another issue with SE-BEST is that while Eq. 15 is an analytic function for a given QM and MM pair, since it is defined in terms of the furthest QM and closest MM particles, whose identities can change discontinuously in the course of a dynamics simulation, similar to the issue with FIRES discussed in Section II A, $f^{\rm SE}$ is a non-analytic function of the particles' coordinates, and therefore using it as a biasing function in molecular dynamics leads to energy non-conservation, as we show below.

Instead of increasing α at the expense of a small timestep, another approach to improving accuracy in BEST with a small α would be to go beyond the SE approximation and account for multiple QM and MM exchanges in the denominator of Eq. 12. A step up in complexity from the SE approximation is the Double Exchange (DE) approximation, which considers pair functions involving the furthest two QM and closest two MM particles.[13] Then, following Eq. 13, the penalty function becomes the product of four pair functions:

$$h_{ij,kl} = g_{ik} \times g_{il} \times g_{jk} \times g_{jl}, \tag{16}$$

and the bias function is normalized by a sum of $(2+2)!/(2! \times 2!) = 6$ terms:

$$f^{\rm DE}_{ij,kl} = \frac{h_{ij,kl}}{h_{ij,kl} + h_{kj,il} + h_{lj,ki} + h_{ik,jl} + h_{il,kj} + h_{kl,ij}}, \tag{17}$$

where ij and kl index the two furthest QM and two closest MM particles respectively. While the DE approximation might be expected to improve over the SE approximation, we find below that it exhibits poorer energy conservation than the SE approximation. This results from the derivative discontinuity of the pair function in Eq. 14 at $x_i = x_k$, shown as the kink in the black curve in Fig. 1(a). This derivative discontinuity happens to cancel in the SE approximation (Eq. 15), but not when multiple exchanges are included.

The issues of derivative discontinuities in the BEST bias potential motivated us to explore alternative functional forms for the pair potential, and led to the development of the FlexiBLE method described in the next section.

C. FlexiBLE

Our FlexiBLE method involves two key differences to BEST. The first modification we make is to the pair function, g_{ik}, to ensure it applies an increasing bias force with QM and MM distance while also having continuous first and second derivatives. This guarantees that QM/MM separation is (approximately) achieved for any choice of $\alpha > 0$ and that total energy is conserved in MD simulations. The second modification is to include all non-negligible permuted penalty functions, $h_{ij\cdots,kl\cdots}$,

by truncating the denominator of the bias function in Eq. 12 using a tree algorithm. By including all important QM/MM exchanges in the bias function, we can use a much smaller value of α than SE-BEST required, without introducing errors in ensemble averages. This allows us to achieve energy-conserving dynamics with a 1.0-fs timestep. We provide details of these developments below.

1. Pair function

As noted above, the BEST pair function, g_{ik}, in Eq. 14 leads to a constant bias force (being proportional to the derivative of the pair function) between a pair of QM and MM particles with large QM and MM separation (see Fig. 1). The magnitude of the bias force that restores QM and MM separation is furthermore proportional to the exponent parameter, α, so if this parameter is too small, QM and MM separation is not maintained (see for example Fig. 4 of Ref. 13). Although this issue can be resolved by choosing a sufficiently large value of α, we find below that the resulting large bias forces require a small timestep of 0.25 fs for a stable MD simulation, and this incurs a significant computational overhead unless a multiple timestep algorithm is used.[14] To allow the use of a smaller α parameter, it would be preferable to have a bias force that increases with QM and MM separation, as in FIRES.[10] This can be achieved by replacing the exponent in Eq. 14 with a function that has a quadratic decay at large QM and MM separation. Although a simple half-Gaussian function for $x_i \geq x_k$ would satisfy this property, the second derivative of such a pair function would be discontinuous at $x_i = x_k$ and this would likely cause convergence problems in geometry optimizations. After some experimentation, we opted for the following rational form of the pair function's exponent:

$$g_{ik}^{\mathrm{FlexiBLE}} = \begin{cases} 1, & x_i < x_k \\ \exp(-\dfrac{\alpha^3 (x_i - x_k)^3}{1 + \alpha(x_i - x_k)}), & x_i \geq x_k \end{cases} \tag{18}$$

The FlexiBLE pair function of Eq. 18 and the resulting bias potential for a pair of QM and MM particles are shown Fig. 1(b), where we see our choice of a rational function exponent yields a smooth pair function (continuous first and second derivatives) while also exhibiting a quadratically increasing bias potential in the limit of large QM and MM separation.

It is interesting to explore the behavior of the bias potential for a pair of QM and MM particles with $\Delta x = x_i - x_k \leq 0$. In this regime, the QM and MM particles are not mixed, yet the bias potential is non-zero. This results from the normalization condition on the bias function (Eq. 12): as QM and MM particles approach each other from $\Delta x < 0$, the bias function with exchanged QM and MM indices becomes non-negligible, so the magnitude of the bias function is reduced from unity. Contributions to the bias potential from QM and MM pairs that have $x_i \leq x_k$ must therefore be considered. FIRES neglects these contributions, and sets the bias potential to zero for any configuration with QM/MM separation.[10] This can be understood as the cause of the structural artifacts at the boundary observed in FIRES when k^{FIRES} is not large enough.[12,47] However, in spite of the requirement in FlexiBLE of including a bias even for configurations with QM/MM separation, the bias potential is seen to decay rapidly to zero for $x_i \ll x_k$ suggesting a truncation of terms involving exchanges between QM and MM particles with large negative displacement differences is possible. Furthermore, the rapid increase of V^{bias} with $x_i \gg x_k$ strongly biases against a large number of QM and MM pairs having $\Delta x > 0$, which always contribute to the bias potential, again suggesting a truncation of terms is possible. Indeed, we show below that the FlexiBLE bias function can be aggressively truncated. Nevertheless, the form of the bias function does permit a small degree of deviation from perfect QM/MM separation in a region around $\Delta x = 0$ close to the QM/MM boundary. This is true for any choice of pair function, including the form used in BEST. The width of this region is controlled by how rapidly the pair function decays with QM/MM separation, and can thus be made negligibly small with a suitably large value of α.

2. Truncating the bias function

The second modification to BEST we make is to include all non-negligible QM and MM exchanges in the calculation of the denominator of the bias function in Eq. 12. Since the total number of exchanges increases factorially as $(N_{\mathrm{QM}}+N_{\mathrm{MM}})!/(N_{\mathrm{QM}}!N_{\mathrm{MM}}!)$, which would introduce a computational bottleneck for QM regions of more than a few particles, to make FlexiBLE practical, we must aggressively truncate the sum over exchanges in Eq. 12. The key to making this possible is a product of pair function form of the penalty function, which we retain from the BEST formalism:

$$h^{\mathrm{FlexiBLE}}_{i''j''\cdots,k''l''\ldots} = \prod_{m=i'',j'',\cdots}^{N_{\mathrm{QM}}^{\mathrm{imp}}} \prod_{n=k'',l'',\cdots}^{N_{\mathrm{MM}}^{\mathrm{imp}}} g^{\mathrm{FlexiBLE}}_{I_m I_n}, \tag{19}$$

where $i'', j'', \cdots$ index *important* QM particles (to be defined below) of which there are $N_{\mathrm{QM}}^{\mathrm{imp}}$, and $k'', l'', \cdots$ index important MM particles of which there are $N_{\mathrm{MM}}^{\mathrm{imp}}$. I_m is the original index of the m-th important particle. Since the bias potential favors QM/MM separation, most of the QM and MM pairs give $g_{mn}^{\mathrm{FlexiBLE}} = 1$ and their contributions therefore do not need to be included in the penalty function. Furthermore, given the exponential form of the pair function in Eq. 18, the magnitude of the penalty function in Eq. 19 is dominated by the smallest value of $g_{mn}^{\mathrm{FlexiBLE}}$, which comes from the pair involving the furthest QM and closest MM particles after a particular exchange pattern. When the distance difference

between the furthest QM and closest MM particle is large following exchange, the value of g_{mn}^{FlexiBLE}, and therefore also h^{FlexiBLE}, for that exchange pattern can be safely neglected from the denominator of Eq. 12.

The observations above imply two levels of truncation in forming the bias function: Firstly, most QM and MM particles do not contribute to the denominator of the bias potential, because they give $g_{mn}^{\text{FlexiBLE}} = 1$ (and therefore could be neglected in the evaluation of Eq. 19) and $g_{mn}^{\text{FlexiBLE}} \approx 0$ when exchanged (and therefore yield a value of h^{FlexiBLE} that can be neglected from the denominator of Eq. 12). Secondly, most exchange patterns in the denominator of Eq. 12 can be neglected because they involve multiple exchanges of QM and MM particles that result in a negligibly small h^{FlexiBLE}. The denominator of the bias function can therefore be aggressively truncated and the FlexiBLE bias function is thus:

$$f^{\text{FlexiBLE}} = \frac{h_{ij\cdots,kl\cdots}^{\text{FlexiBLE}}}{\sum_{L\in\mathbf{L}^{\text{imp}}} \hat{P}_L\left(h_{i''j''\cdots,k''l''\cdots}^{\text{FlexiBLE}}\right)}, \tag{20}$$

where $\mathbf{L}^{\text{imp}}$ is the set of important QM/MM exchanges which lead to a non-negligible contribution to the denominator, *i.e.* for which $\hat{P}_L\left(h_{i''j''\cdots,k''l''\cdots}\right) > h_{\text{thre}}$, where h_{thre} is a threshold parameter. Note: the numerator penalty function is evaluated for all QM and MM pairs, without truncation (indicated by unprimed indices). This is because the numerator could have a value below the threshold parameter, but must not be approximated as zero, or else the bias potential would diverge.

The first type of truncation of the FlexiBLE bias function involves a pre-screening of important QM and MM atoms that have at least one exchange resulting in a non-negligible value of $h^{\text{FlexiBLE}} > h_{\text{thre}}$ in the denominator of Eq. 20. To find the important QM and MM particles we start by reordering their indices based on the particle's distance from the QM origin. Note: this reordering might involve exchanges of QM and MM particles, but as explained below, this is an exchange pattern that will always be retained in the denominator of Eq. 20. Then, as we show in Appendix C, a rigorous upper bound to h^{FlexiBLE} resulting from any exchange pattern (following reordering) that exchanges the p'th particle from QM to MM, and no closer QM particles, is given by the value of the ordered penalty function with a single exchange of the p'th particle with the innermost MM particle (with index k'), *i.e.*:

$$h_{p'}^{\text{QM bound}} = h_{i'j'\cdots k'\cdots,p'l'\cdots}^{\text{FlexiBLE}}, \tag{21}$$

where the primed indices indicate they have been reordered. Likewise, an upper bound to h^{FlexiBLE} resulting from any exchange pattern that exchanges the q'th particle from MM to QM, and no further MM particles, is given by its exchange with the *outermost* QM particle (having index N^{QM}):

$$h_{q'}^{\text{MM bound}} = h_{i'j'\cdots q',k'l'\cdots N^{\text{QM}}\cdots}^{\text{FlexiBLE}}. \tag{22}$$

Furthermore, it is clear that the bounds decay monotonically with increasing distance from the QM/MM boundary, so that important QM and MM atoms can be quickly pre-screened to satisfy $h_{p'}^{\text{QM bound}} > h_{\text{thre}}$ and $h_{q'}^{\text{MM bound}} > h_{\text{thre}}$.

The ability to pre-screen important QM and MM particles demonstrates that the FlexiBLE bias potential incurs non-negligible bias forces only on QM and MM particles within a thin layer at the QM/MM boundary. The boundary layer's radius depends on the instantaneous location of the outermost QM and innermost MM particles (following reordering), and the width of the layer depends on the bounds in Eq. 21 & 22. The boundary layer is therefore *flexible* and allows for density fluctuations in both the QM, MM, and their boundary regions. A schematic of the boundary layer is shown in Fig. 2.

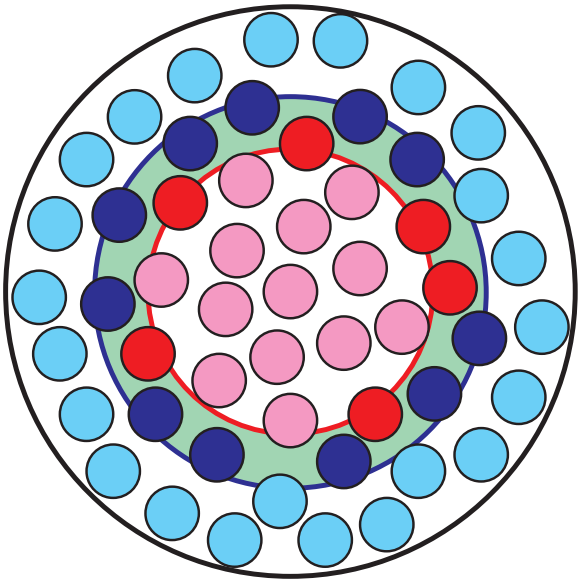

FIG. 2. Schematic of the FlexiBLE boundary layer. Particles are indicated by small circles (QM: red shade, MM: blue shade). Important QM and MM particles that experience bias forces are indicated by dark-shaded colors and form a boundary layer (shaded in green).

While pre-screening the important QM and MM particles is vital to an efficient implementation of FlexiBLE, used alone it does not avoid a steeply scaling computational cost with number of QM particles (confirmed numerically in Fig. 10 below). This can be understood from Fig. 2: as the total number of QM particles grows, so too does the number of important particles in the boundary layer, which scales as the ratio of the QM region's surface area to its volume $(N^{\text{QM}})^{2/3}$ for the spherical QM/MM partitioning of interest. Thus, to reduce the scaling, a second type of truncation must be performed: to neglect small values of the exchanged penalty function, $\hat{P}_L\left(h_{i''j''\cdots,k''l''\cdots}\right)$ in the denominator of Eq. 20, which we achieve with a tree algorithm.

We initiate the tree by noting that the one exchange

pattern of QM and MM indices that is guaranteed to have the largest contribution to the denominator of the bias function is the exchange that results in an ordering of QM and MM distances, *i.e.* perfect QM/MM separation, with a penalty function equal to 1. This exchange pattern must always be included, and since the penalty function from all other exchange patterns must be less than or equal to the ordered penalty function, we take this to be the root of the tree structure. The tree is then traversed as follows: taking each node on the previous level as a parent, we generate child nodes on the next level by performing a single exchange between each QM particle and the next immediate outer particle if it is an MM particle. We avoid duplication of children with each other by comparison against a growing list of visited child nodes on each level. Without truncation, the tree would terminate on a single node corresponding to complete MM and QM exchange, although of course this would ordinarily correspond to a negligibly small penalty function, and the tree could have been truncated many levels before this.

We generate the tree in the above way for two reasons: 1) without any truncation, the procedure is guaranteed to visit every unique QM and MM exchange, as proved in Appendix A. 2) The penalty functions for child nodes constructed as a single exchange between a QM and the next outer particle (if MM) are guaranteed to be less than or equal to the penalty function of their parent node, as we prove in Appendix B. Thus, once a parent node is below the penalty function threshold, h_{thre}, no children need to be generated and that branch of the tree can be truncated: this is the key to avoiding strict factorial scaling in the number of denominator terms.

In Fig. 3, we show a schematic of our tree algorithm applied to a hypothetical system of 3 important QM and 3 important MM particles. As in this example, the zeroth level always contains a single root node corresponding to the ordered indices. In addition, the root node always gives rise to a single child on the first level corresponding to an exchange of the outermost QM and innermost MM particle indices (3 and 4 respectively in this example). If the displacement difference between the outermost QM and innermost MM particle was sufficiently large compared to α^{-1}, the tree could be truncated at this level (we do find this to occur occasionally during the MD trajectories described below). In the hypothetical example considered in Fig. 3 however, this is not the case, and the second level thus has two children from exchanging QM particle 2 with MM particle 3, and QM particle 4 with MM particle 5. These nodes then generate three children on level 3 (note, one child is shared between the two parents of level 2). At this level, our algorithm detects that the node corresponding to $h^{\text{FlexiBLE}}_{126,345}$ falls below the threshold (indicated by the change of color to red), and therefore no children are generated from this node. At level 4, only a single child meets the threshold criterion, while at level 5 no children are above the threshold, and therefore the algorithm stops. All surviving nodes are summed to give the denominator of Eq. 20: note, we include parent nodes that were below threshold, since we had already computed their bias function; however, their children are discarded. We provide pseudocode for the tree algorithm in Supplementary Material section S-I. We verify numerically in Section III D below that the number of surviving child nodes exhibits sub-exponential scaling with QM size, with a lower exponent power than without truncation.

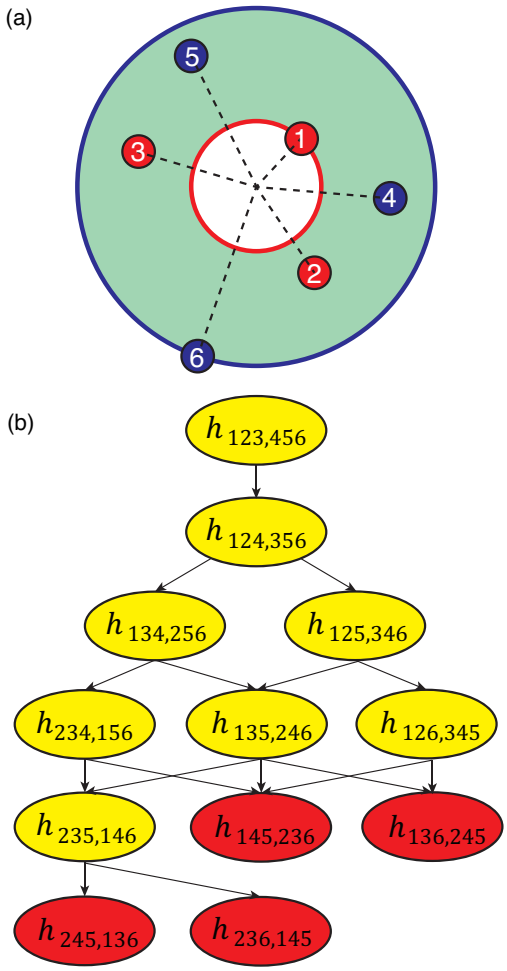

FIG. 3. The FlexiBLE tree algorithm applied to a hypothetical system of 3 important QM and 3 important MM particles, indicated by the schematic shown in panel (a), following the color scheme of Fig. 2. Each node of the tree (ellipses), shown in panel (b), indicates a unique QM/MM exchange pattern (shown by the indices of h). Child nodes are generated by performing a single exchange between each QM particle and the next immediate outer particle if it is an MM particle, indicated by arrows. Duplicate children are not double counted. Yellow nodes indicate the bias function is above the threshold, $h > h_{\text{thre}}$, and further children are generated. Red nodes indicate $h \leq h_{\text{thre}}$ and no children are generated (*i.e.* the branch is truncated). The tree is terminated when all nodes on a level are below threshold, which happens here on level 5. h is summed over all nodes to give the denominator of Eq. 20.

While the tree algorithm described above is an efficient means to sum over important exchanges of the penalty function in the denominator of the FlexiBLE bias function, the relation between the penalty function threshold, h_{thre}, and the error introduced by truncation of the denominator is not immediately clear. In particular, since the penalty function is always positive, truncation necessarily leads to an underestimate of the true denominator. Furthermore, the number of neglected nodes in the tree grows factorially with the number of important QM and MM particles following pre-screening. While our tree algorithm guarantees that each neglected node has a penalty function below the threshold, one might be concerned that an astronomical number of neglected nodes sum to a non-negligible value. To allay this concern, we use an adaptive thresholding, by sweeping through the tree in an iterative fashion, and at each new iteration, tightening the threshold, h_{thre}, to 0.5 of its previous value. The relative change in the denominator of Eq. 20 is monitored between iterations, and when below a convergence parameter, γ, the iterations stop. Given the logarithmic dependence of the bias potential on the bias function (Eq. 11), γ directly provides an estimate of the error in the bias potential due to truncation, in units of $k_{\text{B}}T$. We find a value of $\gamma = 0.001\ k_{\text{B}}T$ yields stable MD trajectories, and convergence is reached typically within two or three FlexiBLE iterations, suggesting that the error in the denominator is typically no worse than an order of magnitude larger than h_{thre}.

Putting all the pieces together, we present pseudocode for FlexiBLE in Algorithm 1 below. Note: analytical bias forces on the QM and MM particles follow straightforwardly from chain-rule derivatives of the pair, penalty, bias, and potential functions (Eqs. 11, 18-20) with respect to particle coordinates. We thus evaluate and accumulate bias force contributions simultaneously with the computation of each surviving h^{FlexiBLE} term. With the pre-screening of important QM and MM pairs, combined with the tree-algorithm enumeration of important QM and MM exchanges, evaluation of the bias potential and forces has an MM-forcefield-like computational cost, and thus adds negligible overhead to a QM/MM simulation.

III. RESULTS AND DISCUSSION

To benchmark the FIRES, BEST, and FlexiBLE methods, we consider the quintessential system exhibiting solvent-supported electronic states: the hydrated electron, $e^-_{(\text{aq})}$, which corresponds to an excess electron embedded in liquid water. We choose an MQC description of the system such that all water molecules are identically treated at an MM level. We can then divide the water molecules into inner and outer regions (MM* and MM respectively), allowing us to verify the accuracy of FlexiBLE without possible errors arising due to a mismatch of QM and MM interactions, and allowing a direct comparison of structural and dynamical properties against results from MQC simulations without FlexiBLE partitioning. The low computational cost of MQC also allows us to reach very large MM* sizes and hundreds of picoseconds of sampling.

A. Computational details

We used a 20 Å radius spherical droplet model of the condensed-phase $e^-_{(\text{aq})}$, containing 1035 water molecules, with an excess electron solvated at the center of the droplet. Although FlexiBLE is fully compatible with Periodic Boundary Conditions (PBC), we chose a spherical droplet model to match the many-electron FlexiBLE-QM/MM model of e^-(aq) that we develop in the companion paper.[48] Comparing our droplet results against previous PBC MQC results shows the impact on both structural and dynamic quantities from using a spherical droplet model to be minimal.

An initial water droplet configuration was generated from a previous PBC MQC simulation[49] of the hydrated electron with the Turi-Borgis (TB) potential[46] by first centering the electron at the origin, tiling space with periodic replicas of the cubic simulation cell, then pruning to form a spherical droplet of the closest 1035 water molecules based on oxygen distance from the origin. This number of water molecules was chosen to be consistent with the output of a SolvateCap command of the tleap program in Amber18 for a water droplet radius of 20 Å.[50] To prevent water evaporation into the surrounding vacuum, a half-harmonic confining potential was applied to the oxygen atoms: $V^{\text{conf}} = 0.5k^{\text{conf}}(r_{\text{O}} - r_{\text{drop}})^2\Theta(r_{\text{O}} - r_{\text{drop}})$, where r_{O} is the distance of the oxygen atom from the origin, taken to be the center of mass of the system, $r_{\text{drop}} = 20$ Å is the droplet radius, $k^{\text{conf}} = 10$ eV/Å^2 is the confining force constant.

To be consistent with previous PBC MQC simulations,[49,51] the SPC/Flex water model was used and the excess electron's wavefunction was discretized in a Fourier-Grid (FG) basis and the one-electron Schrodinger Equation was solved with an iterative diagonalization method (further details below). For Ground-State (GS) simulations, we used $14 \times 14 \times 14$ grid points with a spacing of 1.1034 Å. For excited-state calculations, the grid was extended to $32 \times 32 \times 32$, keeping the grid spacing the same. Since the FG did not span the entire simulation cell, following Ref. 52, we periodically shifted particle positions by an integer number of grid spacings in order to recenter the electron's wavefunction, and avoid FG boundary artefacts. Furthermore, a harmonic restraint with a force constant of $k^{\text{e-COM}} = 10$ eV/Å^2 was placed on the electron centroid using quantum-biased MD[49,53] in order to tether the electron to the center of the droplet, and prevent it from diffusing to the droplet surface. Coulomb and Lennard-Jones pair interactions were not truncated.

For each system (unbiased, FIRES, BEST, FlexiBLE), we computed observables from ten trajectories of 50 ps in

Algorithm 1 Complete FlexiBLE algorithm

$y \leftarrow h^{\text{FlexiBLE}}_{ij\cdots,kl\cdots}$ ▷ Compute numerator of bias function (Eqs. 20 & 18).
$\{i', j', \cdots, k', l', \cdots\} \leftarrow \text{OrderIndices}(\{i, j, \cdots, k, l, \cdots\})$ ▷ Reorder particle indices based on distance from center.
$h_{\text{thre}} \leftarrow \gamma$
for a FlexiBLE iteration **do**
  $\{I_{i''}, I_{j''}, \cdots, I_{k''}, I_{l''}, \cdots\} \leftarrow \text{PreScreen}(\{i', j', \cdots, k', l', \cdots\})$ ▷ With current threshold, h_{thre}, pre-screen to get the N^{imp}_{QM} and N^{imp}_{MM} important QM and MM particle indices using bounds of Eqs. 21 & 22
  $d_a \leftarrow \sum_{L\in\mathbf{L}^{\text{imp}}} \hat{P}_L \left(h^{\text{FlexiBLE}}_{i''j''\cdots,k''l''\cdots}\right)$ ▷ Compute truncated denominator with tree algorithm.
  if $d_a - d_{a-1} > \gamma d_{a-1}$ **then**
    $h_{\text{thre}} \leftarrow \frac{1}{2} h_{\text{thre}}$; Continue
  else
    Break
  end if
end for
$V^{\text{bias}} \leftarrow -k_{\text{B}}T(\log y - \log d_a)$

length. The initial configurations of each trajectory were sampled at 5 ps intervals from another trajectory of 50 ps in length, following an equilibration period of 10 ps. Initial velocities were randomly sampled from the Maxwell-Boltzmann distribution, using a different random number seed for each trajectory. Temperature was maintained at 298 K using the Bussi-Parrinello thermostat[54] with a friction coefficient of 0.5 ps^{-1}. Unless otherwise stated, the velocity Verlet algorithm with a timestep of 0.5 fs was used to propagate MD.[55]

At each MD timestep, the lowest eigenvalue of the one-electron Hamiltonian with the TB potential was solved iteratively using the Generalized Davidson (GD) algorithm,[56] as implemented in SLEPc 3.14.1.[57] Forces on the water molecules from the excess electron were evaluated with the Hellmann-Feynman theorem,[58] which is exact for a fixed FG basis. To analyze energy-gap fluctuations, the lowest five electronic excited states of $e^-_{(\text{aq})}$ were computed in an ex post facto fashion on snapshots taken from the GS MQC trajectories. However, for these excited-state calculations, we found that the GD algorithm occasionally had convergence problems and would miss roots, so we instead used the Krylov-Schur algorithm[59] along with a second-order Chebyshev polynomial spectral transformation of the eigenspectrum, described previously.[52]

Unless otherwise stated, 64 inner water molecules were treated as "QM" particles using either BEST or FlexiBLE; however, since these particles retained MM forcefields, we label them MM* to avoid confusion. MM* and MM particle distances from the origin, needed in Eqs. 14 and Eqs. 18, were computed based on the water oxygen distance from the center of mass of the entire system. This choice maintained translation invariance of the Hamiltonian.

B. Structural properties

1. FIRES

We start by exploring how accurately FIRES reproduces ensemble-averaged structural properties of $e^-_{(\text{aq})}$. The quantity of interest is the Radial Distribution Function (RDF), $g(r)$, of the water atoms relative to the excess electron's centroid position. Fig. 4(a) plots the e^--oxygen RDF from FIRES simulations with the half-harmonic force constant, k^{FIRES}, varied from 23.1 to 1000 eV/Å^2. For the highest force constants, we found that a timestep of 0.25 fs was needed to achieve energy conservation during MD propagation, so we used this timestep for all FIRES simulations. Results without a MM*/MM constraining potential are labelled "Full system" and are shown as purple circles.

From Fig. 4, we see overall good agreement between FIRES and full system results except for a noticeable blip of increased water density at a radius of 7.7 Å, corresponding to the MM*/MM transition region. The observation of structural artifacts at the MM*/MM boundary matches previous findings in Refs. 12 and 47. It is interesting to note that Rowley and Roux did not observe such structural problems in their simulations, even with $k^{\text{FIRES}} = 500$ kcal/mol/Å^2,[10] which is close to our smallest chosen force constant of $k^{\text{FIRES}} = 23.1$ eV/Å^2. The likely origin of this different behavior is that our MM* size is much larger than the QM size used by Rowley and Roux. Bulo and co-workers observed similar behavior, noting that the performance of FIRES is density dependent, and placing the QM/MM boundary in a region of low density can mask structural problems.[12]

As expected based on the analysis in Section II A, the deviations between FIRES and the full-system results are seen to decrease with increasing k^{FIRES} (see inset to panel (a)). However, evidence of a small increase in water density at the boundary is seen even for $k^{\text{FIRES}} = 1000$ eV/Å^2. Unfortunately, the violation in energy conservation during MD propagation using this value of force constant is already quite large when a 1.0-fs timestep

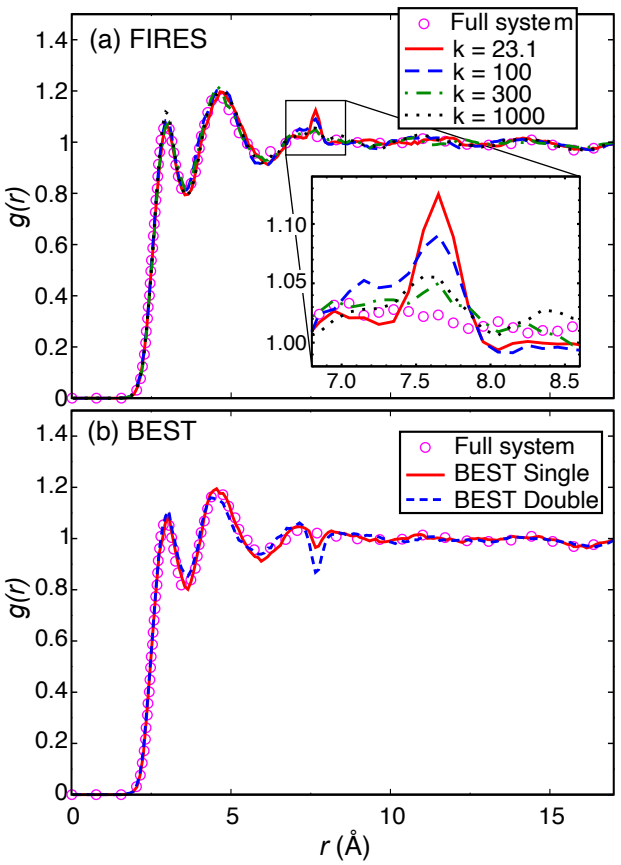

FIG. 4. RDF of e^--oxygen distances from constrained MM*/MM simulations with 64 inner MM* water molecules. Panel (a): FIRES using different force constants, k, in units of eV/Å^2. The inset shows a blow up of the MM*/MM transition region. Panel (b): BEST using the single exchange or double exchange approximations (solid red and dashed blue curves respectively) with $\alpha = 189$ Å^{-1}. Unconstrained "Full system" results are indicated as purple circles.

is used (0.125 eV/ps, see Fig. S1), meaning that k^{FIRES} cannot be increased beyond 1000 eV/Å^2. This motivated us to consider other constrained QM/MM methods.

2. BEST

Fig. 4(b) plots the RDF for e^--oxygen distances from BEST simulations in the Single (SE) and Double (DE) approximations (solid red and dashed blue curve respectively). Here we used a high value of the exponent parameter, $\alpha = 189$ Å^{-1}, to match the previous BEST work.[13,14] With this choice, we found that a timestep of 0.25 fs was needed for stable MD propagation.

Similar to FIRES, we see overall good agreement between BEST and Full system results except for noticeable artefacts at the MM*/MM boundary region at 7.6 Å. Interestingly, BEST-DE performs worse than SE. Evidence that derivative discontinuities in the BEST potential (discussed in Section II B) are the source of the problem is seen in Fig. 5, which reveals noticeable drifts in the total energy of ~ 0.8 eV/ps for BEST-SE (red curve) and worsening to ~ 1.7 eV/ps for BEST-DE (blue curve). We note that these problems were not observed in the original BEST work,[13,14] which however used much smaller QM (or MM*) regions of 12 or fewer water molecules, compared to 64 in this work. In Fig. S2, we verify that BEST-SE conserves energy and reproduces full system results with MM* = 4. The derivative discontinuity problems therefore seem to become appreciable only for large QM regions. This makes sense, since a large QM region will experience frequent exchanges between the furthest QM and second furthest QM particle, leading to derivative discontinuities resulting from the SE approximation. The DE approximation also apparently fails to ameliorate this problem, since the pair function it relies on (Eq. 14) itself has a derivative discontinuity, which is not perfectly cancelled out, unlike in the SE approximation.

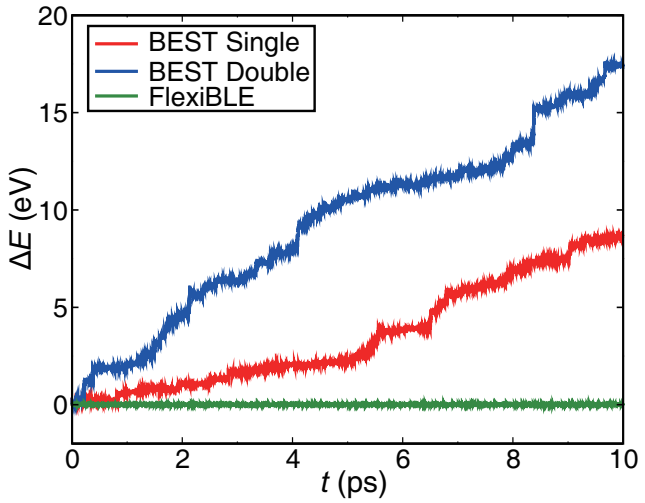

FIG. 5. Total energy drift along a 10-ps trajectory for BEST-SE, BEST-DE, and FlexiBLE simulations (red, blue, and green curves respectively) in the NVE ensemble with a timestep of 0.25 fs. The BEST simulations used $\alpha = 189$ Å^{-1}, and FlexiBLE used $\alpha = 15$ Å^{-1}.

The lack of energy conservation in BEST leads to a temperature gradient, with the MM* region heating up relative to the MM region, as shown in Fig. 6. This is despite the use of a thermostat, and the resulting non-equilibrium condition explains why ensemble quantities like the RDF in Fig. 4 do not agree with full system results. It is possible that the temperature gradients in BEST could be ameliorated with a more aggressive thermostating, in particular by associating an independent thermostat with each particle, and increasing the friction coefficient from our choice of 0.5 ps^{-1}. However, such an approach would strongly affect dynamical quantities, which we wish to preserve. The lack of energy conservation in both FIRES and BEST motivated us to develop the FlexiBLE method, described in Section II C.

We consider its performance next.

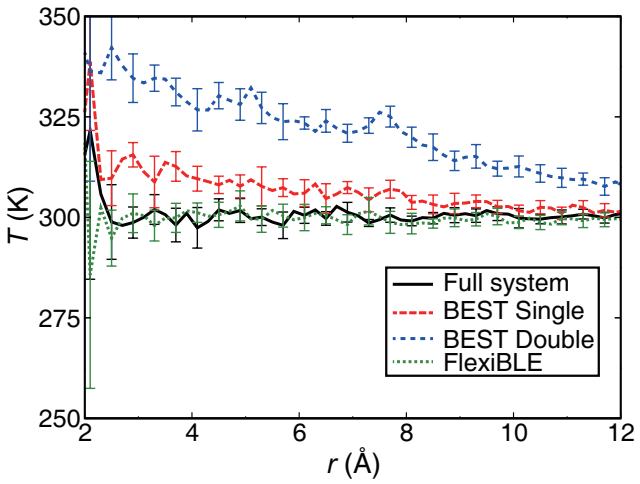

FIG. 6. Temperature distribution for BEST-SE, BEST-DE and FlexiBLE simulations (red long dashed, blue dashed, and green dotted curves respectively) as a function of distance from the excess electron. Full system results are shown as the solid black curve. A noticeable temperature gradient is observed for the BEST simulations, while FlexiBLE maintains thermal equilibrium, in agreement with the full system results.

3. FlexiBLE

Before considering FlexiBLE's ability to reproduce structural properties, we first confirmed that it perfectly conserves energy during MD propagation (green curve of Fig. 5) and does not exhibit temperature gradients (green curve of Fig. 6). We then computed e^--water RDFs from FlexiBLE simulations using $\alpha = 5$ Å^{-1}. This small exponent allows for a stable MD propagation over a long timescale with a timestep of 1.0 fs, which we confirm from a series of NVE simulations, with results shown in Fig. S4. RDFs are shown in Fig. 7, where we see perfect agreement between FlexiBLE (solid black curves) and Full system results (green circles) for both e^--O and e^--H RDF in panels (a) and (b) respectively.

Aside from the total RDF, there is also a question of how well MM*/MM separation is maintained with a small bias exponent. In particular, it was seen in BEST that QM/MM separation broke down unless a bias exponent of $\alpha \geq 19$ Å^{-1} was used.[13] Fig. 7 shows a decomposition of the RDF into MM* and MM contributions (solid red and blue shaded regions respectively), which demonstrates that FlexiBLE successfully maintains MM*/MM separation, even when a small $\alpha = 5$ Å^{-1} is used. The small amount of overlap between MM* and MM contributions to the e^--O RDF in panel (a) largely results from density fluctuations that dynamically move the location of the MM*/MM boundary, resulting in a smearing out of their distributions in the ensemble average. The MM*

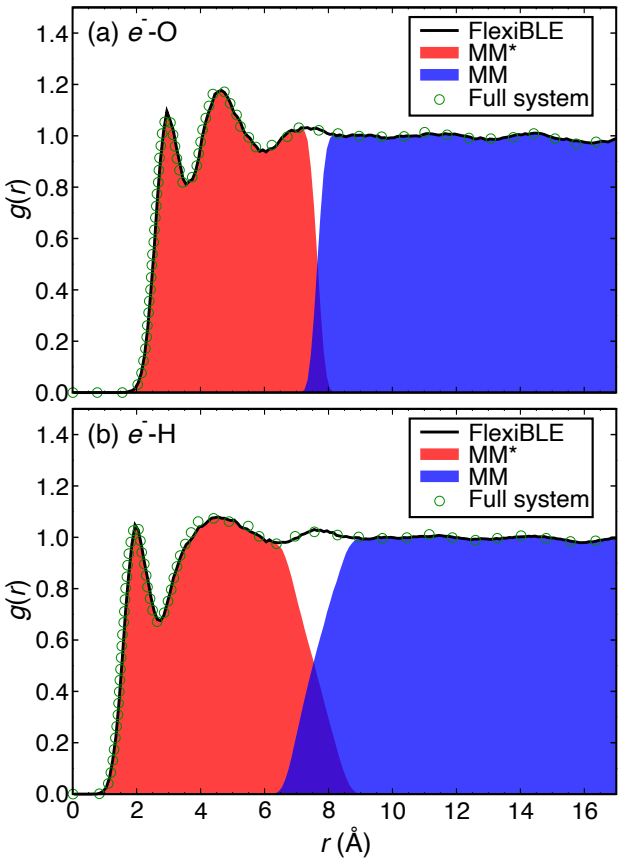

FIG. 7. RDF of $e^-_{(aq)}$ from FlexiBLE simulations with $\alpha = 5$ Å^{-1} (solid black curves) compared to unbiased Full system simulations (green circles). Panel (a): e^--Oxygen distributions. Panel (b): e^--Hydrogen distributions. Also shown is a breakdown into MM* (red) and MM (blue) contributions. FlexiBLE reproduces full system results and successfully maintains MM*/MM separation.

and MM e^--H distributions in panel (b) display more overlap than the e^--O distributions since the bias potential is applied to the oxygen atom only, *i.e.* some MM water molecules can rotate their hydrogen atoms to be within the MM* region. This is expected and desirable behavior: applying orientational constraints on MM particles would lead to an artificial structuring at the boundary, so we apply boundary constraints only on a single atom (or virtual site) of each MM* and MM molecule.

FlexiBLE is able to maintain MM*/MM separation even with a small bias exponent, due to our construction of a bias potential that has an asymptotically quadratic repulsion as particles cross the boundary (see Fig. 1). It should be noted, however, that FlexiBLE does permit a degree of MM*/MM mixing at the boundary, the amount of which depends on the choice of bias exponent, α. This mixing can be quantified by r_{mix}, the difference between the radii of the outermost MM* and innermost MM molecules, the distribution of which is shown in Fig. 8.

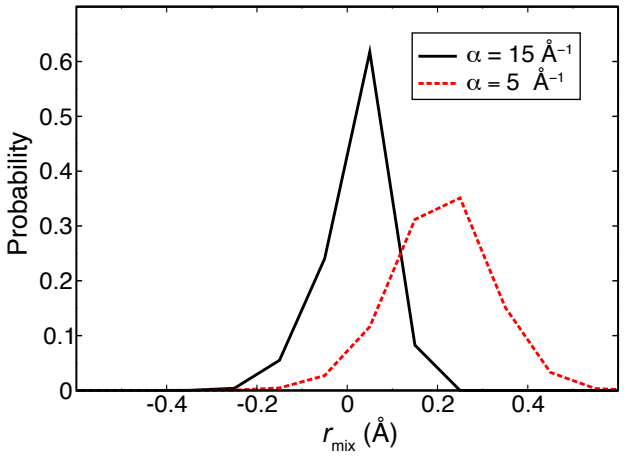

FIG. 8. MM*/MM mixing in FlexiBLE, measured by r_{mix}, the difference between the radii of the outermost MM* and innermost MM molecules. Distributions of r_{mix} are shown for two choices of bias exponent, $\alpha = 5\text{Å}^{-1}$ (dotted red curve) and $\alpha = 15\text{Å}^{-1}$ (solid black curve). $r_{\mathrm{mix}} \geq 0$ indicates MM*/MM mixing, $r_{\mathrm{mix}} < 0$ indicates MM*/MM separation.

For $\alpha = 5$ Å^{-1} (dotted red curve of Fig. 8), r_{mix} has an average value of 0.2 Å and remains below 0.6 Å, which is less than an OH bond length. We view this as an acceptable amount of MM*/MM mixing because the rotational freedom of water places MM hydrogen atoms within the MM* region (and vice versa) anyway (see Fig. 7). Furthermore, the definition of r_{mix} involves differences in radial distances between the outermost MM* and innermost MM molecules, which are not necessarily neighboring each other. The MM* molecules are still mostly surrounded by other MM* molecules. As a result, we do not expect the observed mixing in FlexiBLE to introduce problematic over-polarization of QM molecules when it is applied to actual QM/MM simulations, at least no worse than seen in regular QM/MM simulations. Indeed, in Part II, we see that structural properties of liquid water and the hydrated electron are free of artifacts in FlexiBLE-QM/MM AIMD simulations.[48] With that said, it is straightforward to reduce the amount of QM/MM mixing, if so desired, by increasing α, at the expense of needing a smaller timestep. We demonstrate this in Fig. 8 for $\alpha = 15$ Å^{-1} (solid black curve), which required a timestep of 0.5 fs.

C. Dynamical properties

Having seen that FlexiBLE successfully reproduces structural properties of $e^-_{(\mathrm{aq})}$ compared to simulations without a bias, we next consider dynamical properties. Our focus is on Time Autocorrelation Functions (TACF), which although are equilibrium dynamical properties, require a correct description of not only the system's configurational space distribution, but also its time evolution, making TACFs much more challenging to reproduce than time-independent ensemble averages like the RDF. Of particular interest is the energy gap autocorrelation function, since this reports the dynamical couplings between the solvent and QM solute, and can be used to predict the absorption spectrum of $e^-_{(\mathrm{aq})}$, including nuclear quantum effects.[60] The gap correlation function for state i is defined as

$$C_i(t) = \frac{\langle \delta E_i(t) \cdot \delta E_i(0) \rangle}{\langle \delta E_i(0)^2 \rangle}, \tag{23}$$

where $\delta E_i(t) = (E_i(t) - E_0(t)) - \langle E_i - E_0 \rangle$, and E_i is the adiabatic energy of state i (0 indexes the ground state). We computed energy gaps along the MQC trajectories. A timestep of 0.5 fs was found necessary to properly resolve the influence of the OH stretch on the correlation functions in the frequency domain.

Fig. 9 shows $C(t)$ for the first and fifth excited states (panels a and b respectively). Full system results (green dot dashed curves) agree with previously published PBC MQC simulation results using the TB potential,[60] where we see that both state 1 and 5 exhibit rapid decorrelation, reaching ∼10% of their initial values by $t = 500$ fs. Fourier transforms of the correlation functions reveal that the energy gaps of either state are predominantly coupled to translational and librational motions of water; however, the gap to state 5 exhibits stronger coupling to higher frequency vibrational modes.[60] The difference between the gap correlation functions for state 1 and 5 results from the cavity localized nature of the s-like ground state and p-like state 1 (see orbital plots in the insets to Fig. 9), so that their energies are modulated in a fairly parallel fashion by first-solvent-shell intramolecular motion. On the other hand, having d-like symmetry, and therefore a larger centrifugal repulsion from the cavity, state 5 is much more delocalized and weakly coupled to first-solvent-shell motions, such that its energy gap fluctuations are dominated by the ground-state energy, which is more strongly coupled to vibrations of the first solvent shell.

We now test FlexiBLE's ability to reproduce dynamical quantities. Since a 0.5 fs timestep was used for the Full system calculations, we used the same timestep and chose a FlexiBLE bias exponent of $\alpha = 15$ Å^{-1}. We first consider a very small MM* region of 2 water molecules. This choice places the MM*/MM partition in the first solvent shell of $e^-_{(\mathrm{aq})}$ and would be expected to alter its dynamics. Indeed, we see noticeable deviations between the correlation functions from FlexiBLE 2MM* (dotted black curve) and full-system results (dot-dashed green curve), although the differences are modest. In particular, the gap correlation functions for both state 1 and state 5 decay slower for FlexiBLE 2MM*, presumably because first solvent shell translations and librations are hindered by the biasing potential. Remarkably however, upon extending FlexiBLE to include 16 MM* molecules (two solvation shells), the computed gap correlation functions (solid red curves) are indistinguishable from full

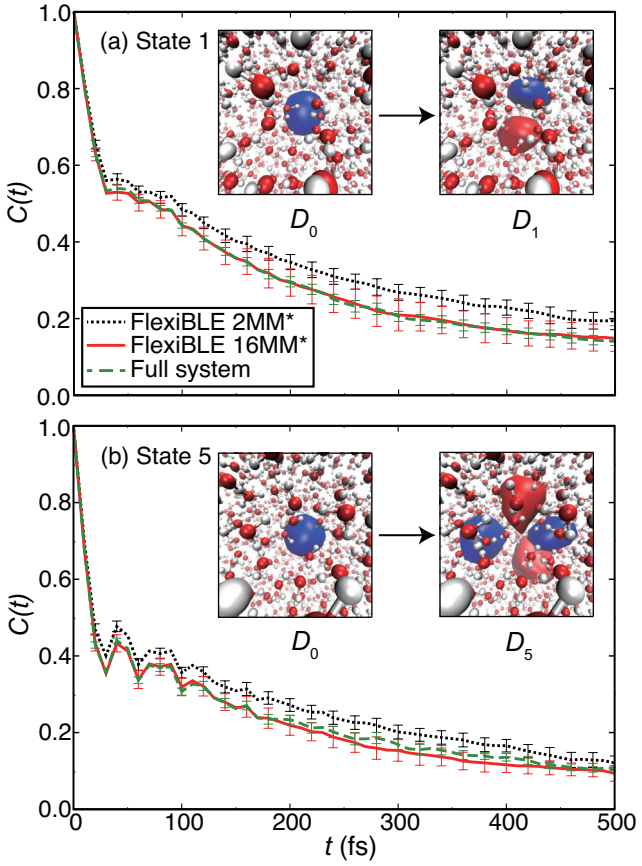

FIG. 9. Gap Correlation Function, $C(t)$, for $e^-_{(aq)}$ from MQC simulations with FlexiBLE MM*/MM partitioning. Panel (a): ground to State 1 energy gap correlation function. Panel (b): ground to State 5 energy gap correlation function. Molecular graphics in the insets show isosurfaces of the relevant orbitals (enclosing 0.667 e of charge density) involved in the electronic excitations. Results are shown with two different sized MM* regions: 2 MM* (dotted black curves) and 16 MM* (solid red curves), with the latter in excellent agreement with full system results (dot-dashed green curves).

system results. Similar agreement is found also for larger MM* regions. Thus, FlexiBLE is seen to preserve dynamical properties of the inner QM region. Similar findings for the FIRES method were observed by Bulo and co-workers.[12]

The electronic structure of $e^-_{(aq)}$ is strongly coupled to solvent motions (being a solvent-supported species), particularly via translational and librational modes that should be most strongly affected by the FlexiBLE potential, so the fact that we do not observe any detectable influence of the bias potential on the dynamics of $e^-_{(aq)}$ gives us confidence the technology could be used to study the dynamics of many other solutes at a FlexiBLE-QM/MM level, providing a sufficiently large QM region is chosen.

D. Computational scaling

We finally consider the computational scaling of FlexiBLE. Without truncation, the cost of the method is expected to scale as the number of penalty function terms in the denominator of Eq. 20, N_{terms}, so we consider this quantity first. As described in Section II C, there are two stages of truncation: first a pre-screening of important MM* and MM particles, then a truncation of exchanges between the surviving important MM* and MM particles in the denominator of Eq. 20. We considered the scaling of N_{terms} following both of these truncations by performing a series of $e^-_{(aq)}$ simulations with the number of MM* waters varied from 4 to 384. An exponent parameter of $\alpha = 15$ Å^{-1} was used for all simulation. 50 ps of dynamics was propagated and N_{terms} was averaged along the trajectory. All simulations were performed on a single core of an Intel Xeon E5-2630 v4 2.20GHz chip.

In Fig. 10 we plot how N_{terms} scales with the size of the QM region, taken here to be the number of MM* water molecules. The black curve shows the total number of exchanges after a pre-screening of important MM* and MM particles. As expected based on theoretical grounds (see Supplementary Material section S-II), following pre-screening, N_{terms} scales as $2^{a(N^{\text{QM}})^{\frac{2}{3}}}$, where a is a free parameter. The relatively high power of $(N^{\text{QM}})^{\frac{2}{3}}$ means that for MM* regions of greater than ∼200 QM particles, the FlexiBLE boundary potential evaluation would become the most computationally expensive part of the MQC calculation. This motivates the truncation of QM/MM exchanges with our tree algorithm (dashed red curve), which we see dramatically reduces N_{terms} and lowers their scaling to $\sim 2^{b(N^{\text{QM}})^{0.53}}$, with b a free parameter.

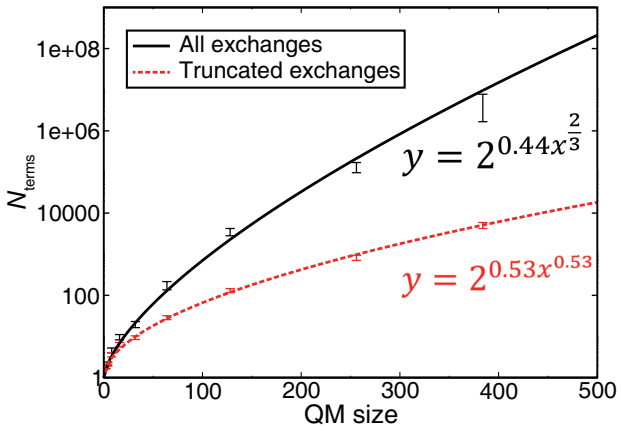

FIG. 10. Scaling of number of FlexiBLE denominator terms (Eq. 20), N_{terms} with QM size. The effects of two types of truncation are shown: the black curve shows the number of terms following a pre-screening of important QM and MM particles. The dashed red curve shows the number of terms following truncation of the denominator according to the FlexiBLE tree algorithm. Equations of best fit are shown.

The computational cost of FlexiBLE is explored in Fig. 11, which plots the average CPU time per timestep spent on FlexiBLE and its breakdown into different operations. The first operation is to sort QM and MM indices (red squares): this is seen to have a low cost that is constant with QM size, since the cost of sorting depends on the total number of particles, $N_{\mathrm{QM}} + N_{\mathrm{MM}}$, which is fixed in our simulations. The next operation is the pre-screening of important QM and MM labels (black circles). This also has a low cost that fits to a sublinear power law with QM size. The evaluation of the FlexiBLE numerator penalty function in Eq. 20 (purple upward triangles) also has a negligible cost that does not scale with QM size. The two operations that dominate a FlexiBLE calculation are thus the evaluation of denominator penalty function terms (turquoise downward triangles), with a scaling that matches Fig. 10, and the bookkeeping associated with maintaining a growing child-node history to avoid their duplication (blue diamonds). Since both of these operations happen in the inner loop of the FlexiBLE algorithm, they exhibit similar sub-exponential scaling. As we see, it is the computations associated with the child history that dominate the total cost of FlexiBLE.

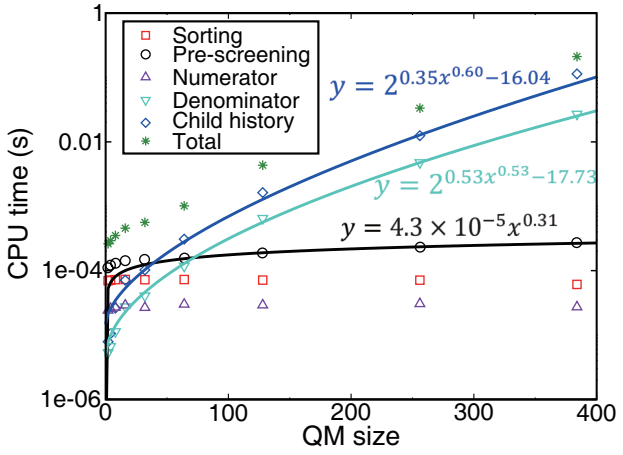

FIG. 11. Computational cost of FlexiBLE with varying QM size, and its breakdown into different contributions, in order of operation. CPU times report averages per timestep. Red squares: cost associated with sorting particle indices by distance from center. Black circles: pre-screening of important QM and MM indices. Purple upward triangles: numerator of Eq. 20. Turquoise downward triangles: denominator of Eq. 20. Blue diamonds: bookkeeping of child history to avoid duplication of denominator terms.

That the bookkeeping associated with avoiding child node duplication in the FlexiBLE algorithm should dominate the calculation is a testament to how significantly our truncation algorithm has reduced the number of surviving denominator penalty function terms: without truncation, they would overwhelm the computational cost. Maintaining a child history is the remaining expensive operation since at every level of the tree, each new child must be compared against a growing list of previously visited children on that level, and the number of children grows steeply with the depth of the tree, even with truncation. We have made this operation as efficient as possible by using a binary representation of the QM and MM labels of each node. As a result, the observed overall sub-exponential scaling of FlexiBLE has a sufficiently small prefactor that the method has a CPU time of <1 s even for QM sizes of ∼400 atoms, which is certainly negligible compared to the cost of *ab initio* electronic structure calculations on QM regions of this size. Nevertheless, the sub-exponential scaling of FlexiBLE, compared to the polynomial scaling of electronic structure, means that a crossover point will occur for QM regions of sufficiently large size, after which the computational cost of FlexiBLE would dominate. The crossover point depends on the level of electronic structure, but appears to be in the thousands of QM particles, for which AIMD is already infeasible, without considering the cost of FlexiBLE, and this will likely remain true for the foreseeable future. Thus, we consider FlexiBLE to be practical for some time to come.

It is reasonable to ask whether the complexity of FlexiBLE presented above is necessary, and if a simpler single or double exchange approximation, as in BEST, could work. We show in Supplementary Material section S-VIII that a double exchange approximation, FlexiBLE-DE, is able to reproduce RDFs, but only in the limit of large bias exponent $\alpha \geq 40$ Å^{-1}, where the approximation of exchanges between only two important MM* and two important MM molecules is accurate. This comes at the cost of requiring a 0.25-fs timestep, so we do not recommend this approximation.

IV. CONCLUSIONS

In this paper, we developed a new constrained QM/MM method called FlexiBLE, which allows for a QM/MM partitioning between identical diffusible particles such as a solvent or gas. We demonstrated that the method has the five qualities of a good constrained QM/MM method set out in the introduction: 1) it closely maintains QM/MM separation. 2) It rigorously preserves ensemble averaged quantities. 3) It exhibits energy conserving MD propagation in liquid water with a 1.0-fs timestep. 4) It introduces bias forces that are highly localized to the QM/MM boundary, leaving dynamical quantities in the inner QM region unperturbed. 5) It comes with negligible computational overhead compared to the cost of QM/MM electronic structure calculations.

We expect that FlexiBLE embedding will be broadly useful in simulating static and dynamic equilibrium properties of complex systems. As a first showcase of the method, in the companion paper,[48] we apply FlexiBLE to the dynamics and electronic structure of $e^-_{(\mathrm{aq})}$ at a many-electron QM/MM level. Although the initial focus of the method has been on solvated systems, the bias function can be easily modified to describe surface ge-

ometries, allowing a description of heterogeneous interfaces. Furthermore, since the FlexiBLE bias potential is a function of nuclear coordinates only, and has an MM-forcefield-like computational cost, the approach is entirely compatible with other embedding schemes, such as polarizable,[61–64] density,[23,24] and mean-field embedding theories.[25] We thus believe that constrained QM/MM partitioning, as in FlexiBLE, is a promising alternative to adaptive QM/MM.

SUPPLEMENTARY MATERIAL

See the supplementary material for further details on our tree algorithm, the computational scaling of FlexiBLE, energy conservation for FIRES, a comparison of BEST with small versus large QM regions, BEST versus FlexiBLE energy conservation, energy conservation for FlexiBLE with a small bias exponent, dynamical properties from FlexiBLE with a large QM region, and structural properties from FlexiBLE with a double-exchange approximation.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China Young Scientist Fund (Grant No. 21603145), the Science and Technology Commission of Shanghai Municipality Foreign Experts Program (Grant No. 21WZ2503600), the NYU-ECNU Center for Computational Chemistry, and start-up funds from NYU Shanghai. We thank Siyao Guo for invaluable assistance with the proof in Appendix A.

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material. Molecular coordinates are available from the corresponding author upon reasonable request.

Appendix A: Proof that tree visits all QM/MM exchanges

Here we prove that all QM/MM exchanges are visited by the tree algorithm described in Section II C. We start by summarizing our algorithm in the language of graph theory.

- With the set of important particle indices, $U := \{1, \ldots, N_{\mathrm{QM}}^{\mathrm{imp}} + N_{\mathrm{MM}}^{\mathrm{imp}}\}$, we consider the directed graph $G = (\{h_{S_0}, \ldots, h_S, \ldots\}, E)$, where $S_0 := \{1, \ldots, N_{\mathrm{QM}}^{\mathrm{imp}}\}$ and for every set $S \subseteq [U]$ of size $N_{\mathrm{QM}}^{\mathrm{imp}}$, we define the node h_S to represent $h_{S,U\setminus S}^{\mathrm{FlexiBLE}}$.
- There exists a directed edge from h_S to $h_{S'}$ if S' is obtained by removing some $a \in S$ from S and adding the corresponding $a+1 \notin S$ into S.

Claim. *There exists a direct path from h_{S_0} to every node in the graph.*

Proof. For every h_S, let d_S denote the difference between the sum of elements in S and the sum of elements in S_0 (*i.e.* d_S is the level of the tree). We will prove the claim inductively according to d_S. The base is true because S must be S_0 itself when $d_S = 0$.

Assume that for integer $n \geq 0$ there exists a directed path from h_{S_0} to any h_S with $d_S = n$. Now we prove that $h_{S'}$ can be reached for any S' with $d_{S'} = n+1$. It suffices to show that there exists a directed edge between S' and S for some S with $d_S = n$. Then we can have a path from S_0 to S' by extending the path from S_0 to S via this edge.

Let $a_1 < \cdots < a_{N_{\mathrm{QM}}^{\mathrm{imp}}}$ denote sorted elements in S' and $b_1 < \cdots b_{N_{\mathrm{MM}}^{\mathrm{imp}}}$ denote sorted elements in $U \setminus S'$. Observe that $b_1 < a_{N_{\mathrm{QM}}^{\mathrm{imp}}}$ (otherwise S' can only be S_0 and $d_{S'} = 0$). Then there must exist i, j such that $a_i = b_j + 1$.[65] Hence, by swapping a_i and b_j, we will obtain the desired S with $d_S = n$. □

Appendix B: Proof of tree hierarchy

To simplify the proofs of inequalities involving the FlexiBLE penalty function, h^{FlexiBLE} (Eq. 19), we first introduce a new notation to identify the QM and MM particle indices that the penalty function depends on. To start, we assume that all QM and MM particles have been ordered by distance from the center of the QM region, as discussed in Section II C. A QM/MM list used to build h^{FlexiBLE} is then defined as $[\mathrm{AA}\cdots\mathrm{ABB}\cdots\mathrm{B}]$ where A and B are QM and MM particle labels respectively, and the location of the label in the list indicates the particle index. Perfect QM/MM separation corresponds to the situation where all A labels are on the left of B labels. Each unique list therefore corresponds to a unique h^{FlexiBLE} function. For example, a list of [AABABB] corresponds to $h_{124,356}^{\mathrm{FlexiBLE}}$.

By the definitions of Eq. 18 and 19, the root of the FlexiBLE tree, $h_{\mathrm{root}}^{\mathrm{FlexiBLE}}$, which has perfect QM/MM separation, is:

$$h_{\mathrm{root}}^{\mathrm{FlexiBLE}} = h[\mathrm{AA}\cdots\mathrm{ABB}\cdots\mathrm{B}] = 1. \tag{B1}$$

$h_{\mathrm{root}}^{\mathrm{FlexiBLE}}$ is the largest term in the denominator because all of its pair functions, g_{mn}, are 1.

According to the tree algorithm discussed in Section II C, we define a child node by performing on its parent a single exchange between a QM particle and the next immediate outer particle if it is an MM particle. Thus, the only child generated from the root node is $h[\mathrm{AA}\cdots\mathrm{ABABB}\cdots\mathrm{B}]$, with a value of $g_{N_{\mathrm{QM}},N_{\mathrm{QM}}+1} \leq 1$.

Thus we see that the first-level node of the tree is guaranteed to be equal to or less than the zeroth-level node.

To show that the hierarchical nature of the tree holds generally for any parent and child, consider a parent node of $[\cdots \mathrm{AB} \cdots]$, where A is the ith particle and B is the $(i+1)$th particle. Suppose that there are N B labels to the left of i with indices $\{n_1, n_2, ..., n_N\}$ and M A labels to the right of $i+1$, with indices $\{m_1, m_2, ..., m_M\}$. Note: there can be an arbitrary number of A labels to the left of i and an arbitrary number of B labels to the right of $i+1$. Upon exchange of i and $i+1$ labels, the relation between the value of the child node and parent node is:

$$h^{\mathrm{FlexiBLE}}_{\mathrm{child}} = h^{\mathrm{FlexiBLE}}_{\mathrm{parent}} \times g_{i+1,i} \times \frac{g_{n_1,i+1}}{g_{n_1,i}} \times \frac{g_{n_2,i+1}}{g_{n_2,i}} \times \cdots \times \frac{g_{n_N,i+1}}{g_{n_N,i}} \times \frac{g_{i,m_1}}{g_{i+1,m_1}} \times \frac{g_{i,m_2}}{g_{i+1,m_2}} \times \cdots \times \frac{g_{i,m_M}}{g_{i+1,m_M}}, \quad \text{(B2)}$$

where we have retained only non-unit terms. Since g_{mn} is monotonically decreasing from 1 to 0, we have the following relations for any exchange:

$$\begin{aligned} g_{i+1,i} &\leq 1, \\ g_{n_a,i+1} &\leq g_{n_a,i}, \forall\, n_a, \\ g_{i,m_b} &\leq g_{i+1,m_b}, \forall\, m_b, \end{aligned} \quad \text{(B3)}$$

which together with Eq. B2 proves that any child node has a value equal to or less than its parent:

$$h^{\mathrm{FlexiBLE}}_{\mathrm{child}} \leq h^{\mathrm{FlexiBLE}}_{\mathrm{parent}}. \quad \text{(B4)}$$

Appendix C: Proof of FlexiBLE penalty function bounds

As discussed in section II C, the number of h^{FlexiBLE} terms can be aggressively truncated by including only the important QM/MM particles that always contribute at least one pair function g^{FlexiBLE} (Eq. 18) that is neither exactly 1 nor results in a penalty function below the truncation threshold, h_{thre}. This allows a pre-screening of particles, and reduces the size of system from (N_{QM},N_{MM}) to ($N^{\mathrm{Imp}}_{\mathrm{QM}}$,$N^{\mathrm{Imp}}_{\mathrm{MM}}$) based on the upper bounds to h^{FlexiBLE} in Eqs. 21 and 22.

To see how pre-screening works, consider the specific case of 4 QM particles and 4 MM particles. The denominator of Eq. 20 involves a sum over h^{FlexiBLE} terms resulting from all possible QM and MM exchanges. However, if every exchange producing a QM label at the eighth particle results in h^{FlexiBLE} below the threshold, h_{thre}, these terms are truncated, and the only surviving terms all have an MM label at the eighth particle position. Following Eq. 19, every surviving pair function, g, involving particle 8 is exactly 1 and does not alter h^{FlexiBLE}, thus particle 8 can be excluded entirely in the construction of h^{FlexiBLE}, and $N^{\mathrm{Imp}}_{\mathrm{MM}}$ can be reduced by one.

To efficiently pre-screen particle indices, we seek an upper bound to h^{FlexiBLE} following any exchange pattern that leaves the outermost QM label at the particle position under consideration. In the 4 QM, 4 MM example, it is clear that the largest h^{FlexiBLE} value is h[AAABBBBA] (using the notation introduced in Appendix B). Any further exchanges (indicated below with "X" labels) will lead to a smaller h, according to the tree hierarchy relation in Appendix B. Therefore, starting from perfect QM/MM separation, the single exchange between the outermost QM label (particle 4) and the eighth particle label (MM) serves as the upper bound for pre-screening:

$$\begin{aligned} h[\mathrm{XXXXXXXA}] \leq h[\mathrm{AAABBBBA}] = \\ g_{84} \times g_{85} \times g_{86} \times g_{87} = h^{\text{MM bound}}_{8}. \end{aligned} \quad \text{(C1)}$$

Generalizing to arbitrary N_{QM} and N_{MM}, the MM bound for particle q' is:

$$\begin{aligned} h[\underbrace{\mathrm{X}\cdots\mathrm{X}}_{q'-1\geq N_{\mathrm{QM}}}\ \mathrm{AB}\cdots\mathrm{B}] \leq h[\underbrace{\mathrm{A}\cdots\mathrm{A}}_{N_{\mathrm{QM}}-1}\ \underbrace{\mathrm{B}\cdots\mathrm{B}}_{q'-N_{\mathrm{QM}}-1}\ \mathrm{AB}\cdots\mathrm{B}] = \\ \prod_{i=N_{\mathrm{QM}}}^{q'-1} g_{q',i} = h^{\text{MM bound}}_{q'}. \end{aligned} \quad \text{(C2)}$$

This proves Eq. 22. The proof of Eq. 21 follows similarly.

REFERENCES

[1] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[2] J. Gao, "Methods and applications of combined quantum mechanical and molecular mechanical potentials," in *Reviews in Computational Chemistry* (John Wiley & Sons, Ltd, 1996) pp. 119–185.
[3] T. Kerdcharoen and K. Morokuma, Chem. Phys. Lett. **355**, 257 (2002).
[4] H. Lin and D. G. Truhlar, Theor. Chem. Acc. **117**, 185 (2006).
[5] S. C. L. Kamerlin, M. Haranczyk, and A. Warshel, J. Phys. Chem. B **113**, 1253 (2009).
[6] H. M. Senn and W. Thiel, Angew. Chem. Int. Ed. **48**, 1198 (2009).
[7] A. Heyden, H. Lin, and D. G. Truhlar, J. Phys. Chem. B **111**, 2231 (2007).
[8] R. E. Bulo, B. Ensing, J. Sikkema, and L. Visscher, J. Chem. Theory Comput. **5**, 2212 (2009).
[9] N. Bernstein, C. Várnai, I. Solt, S. A. Winfield, M. C. Payne, I. Simon, M. Fuxreiter, and G. Csányi, Phys. Chem. Chem. Phys. **14**, 646 (2012).
[10] C. N. Rowley and B. Roux, J. Chem. Theory Comput. **8**, 3526 (2012).
[11] N. Takenaka, Y. Kitamura, Y. Koyano, and M. Nagaoka, Chem. Phys. Lett. **524**, 56 (2012).

[12]R. E. Bulo, C. Michel, P. Fleurat-Lessard, and P. Sautet, J. Chem. Theory Comput. **9**, 5567 (2013).
[13]M. Shiga and M. Masia, J. Chem. Phys. **139**, 044120 (2013).
[14]M. Shiga and M. Masia, J. Chem. Phys. **139**, 144103 (2013).
[15]M. P. Waller, S. Kumbhar, and J. Yang, ChemPhysChem **15**, 3218 (2014).
[16]A. Warshel, Angew. Chem. Int. Ed. **53**, 10020 (2014).
[17]H. C. Watanabe, T. Kubař, and M. Elstner, J. Chem. Theory Comput. **10**, 4242 (2014).
[18]X. Lu, D. Fang, S. Ito, Y. Okamoto, V. Ovchinnikov, and Q. Cui, Mol. Sim. **42**, 1056 (2016).
[19]M. J. Field, J. Chem. Theory Comput. **13**, 2342 (2017).
[20]H. Takahashi, H. Kambe, and A. Morita, J. Chem. Phys. **148**, 134119 (2018).
[21]U. Morzan, D. J. Alonso de Armiño, N. Foglia, F. Ramírez, M. C. González Lebrero, D. A. Scherlis, and D. A. Estrin, Chem. Rev. **118**, 4071 (2018).
[22]H. C. Watanabe and Q. Cui, J. Chem. Theory Comput. **15**, 3917 (2019).
[23]F. Libisch, C. Huang, and E. A. Carter, Acc. Chem. Res. **47**, 2768 (2014).
[24]C. R. Jacob and J. Neugebauer, Wiley Interdiscip. Rev. Comput. Mol. Sci. **4**, 325 (2014).
[25]M. E. Fornace, J. Lee, K. Miyamoto, F. R. Manby, and T. F. Miller, J. Chem. Theory Comput. **11**, 568 (2015).
[26]V. V. Rybkin, J. Chem. Theory Comput. **17**, 3995 (2021).
[27]P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer, and C. Lennartz, J. Mol. Struct. THEOCHEM **632**, 1 (2003).
[28]Y. Zhang, Theor. Chem. Acc. **116**, 43 (2006).
[29]H. Hu and W. Yang, J. Mol. Struct. THEOCHEM **898**, 17 (2009).
[30]E. M. Sproviero, M. B. Newcomer, J. A. Gascón, E. R. Batista, G. W. Brudvig, and V. S. Batista, Photosynth. Res. **102**, 455 (2009).
[31]S. Difley, L.-P. Wang, S. Yeganeh, S. R. Yost, and T. V. Voorhis, Acc. Chem. Res. **43**, 995 (2010).
[32]E. Brunk and U. Rothlisberger, Chem. Rev. **115**, 6217 (2015).
[33]T. J. Zuehlsdorff and C. M. Isborn, Int. J. Quantum Chem. **119**, e25719 (2019).
[34]M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, Chem. Rev. **112**, 632 (2012).
[35]M. S. Gordon, Q. A. Smith, P. Xu, and L. V. Slipchenko, Annu. Rev. Phys. Chem. **64**, 553 (2013).
[36]K. Bolton, W. L. Hase, and C. Doubleday, J. Phys. Chem. B **103**, 3691 (1999).
[37]C. M. Isborn, A. W. Götz, M. A. Clark, R. C. Walker, and T. J. Martínez, J. Chem. Theory Comput. **8**, 5092 (2012).
[38]J. R. Casey, A. Kahros, and B. J. Schwartz, J. Phys. Chem. B **117**, 14173 (2013).
[39]L. Turi and P. J. Rossky, Chem. Rev. **112**, 5641 (2012).
[40]J. M. Herbert, Phys. Chem. Chem. Phys. **21**, 20538 (2019).
[41]X. Chen and S. E. Bradforth, Annu. Rev. Phys. Chem. **59**, 203 (2008).
[42]E. Lambros, F. Lipparini, G. A. Cisneros, and F. Paesani, J. Chem. Theory Comput. **16**, 7462 (2020).
[43]D. Beglov and B. Roux, J. Chem. Phys. **100**, 9050 (1994).
[44]M. Shiga and M. Masia, Mol. Sim. **41**, 827 (2015).
[45]E. J. Hart and J. W. Boag, J. Amer. Chem. Soc. **84**, 4090 (1962).
[46]L. Turi and D. Borgis, J. Chem. Phys. **117**, 6186 (2002).
[47]B. Kirchhoff, E. O. Jónsson, A. O. Dohn, T. Jacob, and H. Jónsson, J. Chem. Theory Comput. **17**, 5863 (2021).
[48]Z. Shen, S. Peng, and W. J. Glover, J. Chem. Phys. **X** (2021).
[49]W. J. Glover and B. J. Schwartz, J. Chem. Theory Comput. **16**, 1263 (2020).
[50]D. Case, I. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. Cheatham, V. W. D. Cruzeiro, T. Darden, R. Duke, D. Ghoreishi, M. Gilson, H. Gohlke, A. Götz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, and P. A. Kollman, *Amber 2018* (University of California, San Francisco, 2018).
[51]W. J. Glover and B. J. Schwartz, J. Chem. Theory Comput. **12**, 5117 (2016).
[52]W. J. Glover, R. E. Larsen, and B. J. Schwartz, J. Chem. Phys. **129**, 164505 (2008).
[53]W. J. Glover, J. R. Casey, and B. J. Schwartz, J. Chem. Theory Comput. **10**, 4661 (2014).
[54]G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).
[55]W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).
[56]K. Neymeyr, Linear Algebra Its Appl. **322**, 61 (2001).
[57]V. Hernandez, J. E. Roman, and V. Vidal, ACM Trans. Math. Software **31**, 351 (2005).
[58]R. P. Feynman, Phys. Rev. **56**, 340 (1939).
[59]G. W. Stewart, SIAM J. Matrix Anal. Appl. **23**, 601 (2002).
[60]L. Turi, G. Hantal, P. J. Rossky, and D. Borgis, J. Chem. Phys. **131**, 024119 (2009).
[61]M. A. Thompson and G. K. Schenter, J. Phys. Chem. **99**, 6374 (1995).
[62]J. Gao, J. Comp. Chem. **18**, 1061 (1997).
[63]J. Kongsted and B. Mennucci, J. Phys. Chem. A **111**, 9890 (2007).
[64]M. A. Hagras and W. J. Glover, J. Chem. Theory Comput. **14**, 2137 (2018).
[65]Consider QM/MM strings according to the above sequences. $b_1 < a_{N_{\mathrm{QM}}^{\mathrm{imp}}}$ implies that there exists some MM index on the left of some QM index. Then between these indices there must exist a consecutive MM-QM pair, otherwise the string can only be of MM indices, which is unphysical.