

Article

# Adaptive QM/MM for Molecular Dynamics Simulations: 5. On the Energy-Conserved Permuted Adaptive-Partitioning Schemes

Adam W. Duster , Chun-Hung Wang  and Hai Lin \* 

Department of Chemistry, University of Colorado Denver, Denver, CO 80217, USA;  
ADAM.DUSTER@ucdenver.edu (A.W.D.); nakada0155@gmail.com (C.-H.W.)

\* Correspondence: hai.lin@ucdenver.edu; Tel.: +1-303-315-7656

-----  
This paper is published as part of a thematic issue of *Molecules* on “Combined Quantum Mechanical and Molecular Mechanical Methods and Simulations”.

([http://www.mdpi.com/journal/molecules/special\\_issues/QM](http://www.mdpi.com/journal/molecules/special_issues/QM))

-----  
Received: 24 July 2018; Accepted: 24 August 2018; Published: 28 August 2018



**Abstract:** In combined quantum-mechanical/molecular-mechanical (QM/MM) dynamics simulations, the adaptive-partitioning (AP) schemes reclassify atoms on-the-fly as QM or MM in a smooth manner. This yields a mobile QM subsystem with contents that are continuously updated as needed. Here, we tailor the Hamiltonian adaptive many-body correction (HAMBC) proposed by Boreboom et al. [*J. Chem. Theory Comput.* **2016**, *12*, 3441] to the permuted AP (PAP) scheme. The treatments lead to the HAMBC-PAP method (HPAP), which both conserves energy and produces accurate solvation structures in the test of “water-in-water” model system.

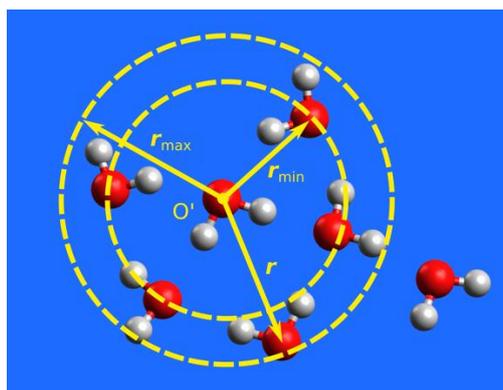
**Keywords:** combined QM/MM; molecular dynamics simulations; energy conservation; radial distribution function; adaptive partitioning

## 1. Introduction

Molecular dynamics (MD) simulations of diffusive systems, such as the diffusion of a solute (a solvated ion or molecule) through solvent, has been a challenging task for multiscale methods, especially for combined quantum-mechanics/molecular-mechanics (QM/MM) methods [1–26]. In conventional QM/MM methodology, atoms are designated as QM or MM at the beginning of a simulation and do not change these identities throughout a simulation. This causes difficulties when solvent molecules are exchanged between the solute’s solvation shells and the bulk solution, which may occur frequently. Adaptive QM/MM mitigates these problems by reclassifying the atoms as QM or MM on-the-fly based on their positions, assuring that the solute and its solvation shells are always described at the QM level of theory [25,27–52]. As a result, the QM and MM partitions in adaptive QM/MM are dynamically updated as needed, in contrast to the static partitions in conventional QM/MM.

One adaptive QM/MM algorithm is the permuted adaptive partitioning (PAP) scheme [29,32], which uses distance-based criteria for the QM and MM partitioning (Figure 1). In PAP, the QM zone, also called the active zone, consists of the solute and all molecules within a preset cutoff distance  $r_{\min}$  from the solute. A group-based prescription is often adopted, where a whole molecule is treated as an entity, and the distance from the solute  $r$  is measured using the center of mass or a representative atom of the entire molecule. The description for whole molecules can also be applied to molecular fragments, such as functional groups [32]. The MM zone, also known as the environmental zone,

consists of molecules with  $r > r_{\max}$ , where  $r_{\max}$  is another preset cutoff distance. Between the QM and MM zones is the buffer zone ( $r_{\max} \geq r \geq r_{\min}$ ), and the molecules in the buffer zone are often called the buffer groups. The energy of the system and the gradients of all (QM, buffer, and MM) atoms are smoothly interpolated when molecules or functional groups migrate into, across, or out of the buffer zone. This is accomplished by expressing the QM/MM potential as a weighted sum of many-body contributions that vary continuously and smoothly as the buffer groups change their positions. The PAP method conserves energy and momentum, and it has been found to yield superior numerical stabilities in MD simulations [29,32].



**Figure 1.** Adaptive QM/MM exemplified by a “water-in-water” model. One selected water molecule (oxygen atom labeled by O') serves as the active-zone center. This molecule and its immediate surrounding water molecules within a distance ( $r < r_{\min}$ ) are treated at the QM level of theory, while those at remote distances ( $r > r_{\max}$ ) at the MM level. The water molecules at intermediate distances ( $r_{\max} \geq r \geq r_{\min}$ ) are in the buffer zone and have mixed QM and MM characters. The energy of the system and the gradients of all (QM, buffer, and MM) atoms are smoothly interpolated when molecules migrate into, across, or out of the buffer zone.

A challenge in PAP (and other distance-based adaptive QM/MM methods) concerns the gradients due to the smoothing functions employed in the interpolation procedure, which, if not negligible, may cause artefacts in the MD simulations [30,35,42,46]. These forces, which are sometimes called transition forces, are proportional to the difference between the QM and MM energies at the current geometry. More specifically, the energy difference is the energy released (or absorbed) when a buffer group is switched from MM to QM while holding the QM or MM classifications of the other groups as well as all atomic coordinates fixed [42]. These transition forces are therefore associated with the difference in chemical potential between the QM and MM descriptions of the given buffer group. These transition forces drive the molecules moving towards the QM or MM zones, even in the absence of the interpolated gradients between the QM and MM potential derivatives, which are considered the “real” or “physical” forces.

In principle, the effects due to these transition forces can be eliminated or minimized by carefully aligning the QM and MM potentials [29]. However, it is difficult to align multi-dimensional potentials, and a simple and general algorithm to for potential alignment has not yet been developed. An alternative solution is the modified PAP (mPAP) scheme, where external forces are applied to cancel out these artificial forces [35]. Mathematically, the transition forces are simply deleted. Conceptually, this means that the chemical potentials are equalized at every step for the system [46]. It has been demonstrated that mPAP yields reasonably accurate structures and dynamics in MD simulations [35,40,43]. The downside is that, because of the involvement of the external forces, the scheme no longer has a Hamiltonian formulation and therefore cannot be used for studies where Hamiltonian systems are required [43].

Recently, Boreboom et al. [42] proposed the Hamiltonian adaptive many-body correction (HAMBC) for sorted adaptive-partitioning (SAP) QM/MM simulations. Inspired by the works of the molecular-mechanics/course-grained (MM/CG) community, especially the Hamiltonian adaptive resolution scheme (H-AdResS) by Potestio et al. [53] the HAMBC method includes per-molecule-based correction terms to the SAP QM/MM Hamiltonian. By design, the gradients of the correction terms should cancel out the *average* transition forces due to the smoothing functions; the cancellation is not necessarily exact at any single time step. The HAMBC was demonstrated in test calculations of a “water-in-water” model using dual MM levels, where a selected water molecule as the solute and its solvation shell are treated by one MM force field model and the bulk solvent by another MM force field model [42]. (Due to their high efficiencies, dual-MM test calculations have frequently been employed in testing adaptive QM/MM schemes, e.g., in the development of the original PAP scheme [29].) It is encouraging to find that HAMBC was able to produce correct solvation structures for the selected solute water while conserving energy in SAP simulations [42].

In this paper, we report the tailoring of the HAMBC treatment to the much more sophisticated PAP method. In the HAMBC by Boreboom et al. [42] the per-molecule correction term is a function of the fractional “QM character” for a solvent molecule, which is the sum of the weights of the contributing partitions that describe this solvent molecule at the QM level. Because in general there is no analytical function for the correction term, the correction must be calculated through thermodynamic integration over the selected variable. In PAP, it is more convenient to use the value of the smoothing function than the QM character for the given buffer group. We have previously shown that this QM character is equivalent to the value of the smoothing function for a given buffer group in a fully expanded PAP potential [46]. However, the many-body expansion of the PAP potential is often truncated to reduce computational costs, and in a truncated PAP potential, the value of the smoothing function no longer equals the QM character. In this work, we demonstrate that the correction can be taken as a function of the value of the smoothing function even when the PAP potential is truncated.

## 2. Materials and Methods

### 2.1. The PAP Algorithm

In PAP, the QM/MM potential is expressed using a many-body expansion: [29]

$$\begin{aligned}
 V = & V^A + \sum_{i=1}^N P_i (V_i^A - V^A) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N P_i P_j \left( V_{ij}^A - \left( V^A + \sum_{r=i,j} (V_r^A - V^A) \right) \right) \\
 & + \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^N P_i P_j P_k \left( V_{ijk}^A - \left( V^A + \sum_{r=i,j,k} (V_r^A - V^A) \right) \right) \\
 & + \sum_{(p,q)=(i,j),(i,k),(j,k)} \left( V_{p,q}^A - \left( V^A + \sum_{r=p,q} (V_r^A - V^A) \right) \right) + \dots
 \end{aligned} \tag{1a}$$

Or more compactly, [32]

$$V = V^A \prod_{i=1}^N (1 - P_i) + \sum_{i=1}^N V_i^A P_i \prod_{j \neq i} (1 - P_j) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N V_{ij}^A P_i P_j \prod_{k \neq j \neq i} (1 - P_k) + \dots \tag{1b}$$

$$= V^A \sigma_0 + \sum_{i=1}^N V_i^A \sigma_i + \sum_{i=1}^{N-1} \sum_{j=i+1}^N V_{ij}^A \sigma_{i,j} + \dots \tag{1c}$$

Here,  $V^A$  is the QM/MM energy of the partitions with no buffer group treated at the QM level,  $V_i^A$  with the  $i$ -th buffer group treated at the QM level,  $V_{ij}^A$  with the  $i$ -th and  $j$ -th buffer groups treated at the QM level,  $\dots$   $V_{1,2,\dots,N}^A$  with all  $N$  buffer groups at the QM level,  $P_i$  is the smoothing function for the  $i$ -th buffer group, and the weights  $\sigma_0, \sigma_i, \sigma_{i,j} \dots$  are assigned to the partitions  $V^A, V_i^A, V_{ij}^A \dots$

respectively. Note that  $N$  can vary from one time step to another. We have chosen the smoothing function  $P_i$  to be a fifth-order spline:

$$P(\alpha_i) = \begin{cases} 0 & \alpha_i > 1 \\ -6\alpha_i^5 + 15\alpha_i^4 - 10\alpha_i^3 + 1, & 1 \geq \alpha_i \geq 0 \\ 1, & \alpha_i < 0 \end{cases} \quad (2)$$

where  $\alpha_i$  is a reduced distance for the  $i$ -th buffer group,

$$\alpha_i = \frac{r_i - r_{\min}}{r_{\max} - r_{\min}} \quad (3)$$

where the distance  $r_i = |\mathbf{r}_i| = |\mathbf{x}_i - \mathbf{x}_A|$  is between the buffer group at position  $\mathbf{x}_i$  and the solute at position  $\mathbf{x}_A$ , and  $r_{\max} \geq r_i \geq r_{\min}$ . The smoothing function  $P_i$  is continuous and differentiable over the domain  $[0, 1]$ , and corresponds to all possible  $r_i$  that a buffer group can take in the buffer zone. The PAP scheme is a Hamiltonian algorithm and conserves energy.

According to Equation (1), a buffer group is treated by QM in some partitions and by MM in the others, in line with its dual QM-MM characteristics. This contrasts with the groups in the QM and MM zones, which remain as QM and MM, respectively, in all possible partitions. (Note that the QM and MM zones are also called the active and environmental zones, respectively.) All derivatives of the PAP potential energy with respect to the coordinates vary smoothly up to the same order for which  $P_i$  varies continuously. The full expansion of the PAP potential requires  $2^N$  calculations. However, the potential is typically truncated and only the terms from a subset of lower-orders are calculated to increase computational efficiency.

If we denote the included partitions by  $S$ , be it a subset or the total of all the partitions at a given time step, the PAP potential can be written as:

$$V = \sum_{s \in S} V_s \sigma_s \quad (4)$$

The derivative with respect to any atomic Cartesian coordinate component  $x_t$  is:

$$-\frac{\partial V}{\partial x_t} = \sum_{s \in S} \left( -\sigma_s \frac{\partial V}{\partial x_t} - V_s \frac{\partial \sigma_s}{\partial x_t} \right) \quad (5)$$

The first term ( $F_{\text{phys}} = -\sum_{s \in S} \sigma_s \frac{\partial V}{\partial x_t}$ ), the negative of the sum of the smoothed gradients, represents the interpolated physical forces between the QM and MM forces. The second term ( $F_{\text{trans}} = -\sum_{s \in S} V_s \frac{\partial \sigma_s}{\partial x_t}$ ) corresponds to the gradients due to the smoothing functions and is the so-called transition forces. The transition forces are associated with the difference in chemical potentials between the QM and MM regions [42]. These transition forces are pairwise, acting on both the solute molecule at the QM-zone center and the buffer groups, and they cause structural artifacts if non-negligible. This “transition-force problem” is universal in distance-based adaptive QM/MM algorithms and is not limited to PAP [46]. Various treatments have been implemented to eliminate the problem or minimize its effects. For example, in the non-Hamiltonian mPAP method [35], these transition forces are deleted, and as a result, the remaining forces acting on the atoms no longer correspond to the potential defined in Equation (1). The mPAP scheme has been shown to yield very accurate structural and dynamics properties in a number of tests [35,40,41].

## 2.2. HAMBC Expression for PAP

To minimize the effects due to the transition forces while maintaining energy conservation, Boreboom et al. [42] recently proposed the HAMBC for SAP QM/MM, which also suffers from the transition force problem. In HAMBC-SAP, an energy correction term is added to the SAP potential

for every buffer group. The correction term is assumed to depend only on the type of the buffer group (e.g., water versus ethanol) and the so-called “QM character,” which is the sum of the weights of the contributing partitions that describe this buffer group at the QM level. For PAP, it would be more convenient to select the smoothing function  $P_i$  than to choose the QM character. Both  $P_i$  and the QM character are metrics related to the buffer group's position within the buffer zone; the QM character is equivalent to  $P_i$  in a fully expanded PAP potential but differs otherwise [46]. The theory and implementations for HAMBC-PAP (HPAP for short) are described below.

### 2.2.1. A Mean-Field Treatment of the Individual Group Corrections

As in Ref. [42], we assume here that the total correction is a sum of individual group corrections over all  $N$  buffer groups and that the individual  $i$ -th group correction ( $H_i^C(r_i)$ ) only depends on the type of species of the group and on the group's distance  $r_i$  to the active-zone center. Because  $P_i$  decreases monotonously from 1 to 0 as  $r_i$  increases from  $r_{\min}$  to  $r_{\max}$ , the correction term  $H_i^C(r_i)$  can also be expressed as a function of the dimensionless  $P_i$ , i.e.,  $H_i^C(P_i)$ . That the corrections to the buffer groups are independent of each other implies that the interactions between the groups are treated in a “mean-field” manner as the correlations between the groups are neglected. Thus, many-body effects are treated in an average fashion. For simplicity, we only consider one species, the solvent. Under this circumstance, the new HPAP potential is:

$$V^{\text{HPAP}} = \sum_{s \in S} V_s \sigma_s - \sum_{i=1}^N H_i^C(P_i) = \sum_{s \in S} V_s \sigma_s - \sum_{i=1}^N H^C(P_i) \quad (6)$$

The subscript  $i$  in  $H_i^C(P_i)$  is dropped because all corrections are identical for the same species. The corresponding forces are given by:

$$-\frac{\partial V^{\text{HPAP}}}{\partial x_t} = -\sum_{s \in S} \left( \sigma_s \frac{\partial V}{\partial x_t} + V_s \sum_{i=1}^N \frac{\partial \sigma_s}{\partial P_i} \frac{dP_i}{dr_i} \frac{\partial r_i}{\partial x_t} \right) + \sum_{i=1}^N \frac{dH^C(P_i)}{dP_i} \frac{dP_i}{dr_i} \frac{\partial r_i}{\partial x_t} \quad (7)$$

The correction force due to the derivative of the individual correction energy ( $F_{\text{corr}} = -\frac{dH^C(P_i)}{dP_i} \frac{dP_i}{dr_i} \frac{\partial r_i}{\partial x_t}$ ) is designed such that, at any given distance  $r_i$  that separates the solute molecule and the  $i$ -th buffer group, it should cancel the average transition forces acting on the buffer group:

$$\left\langle -\sum_{s \in S} V_s \frac{\partial \sigma_s}{\partial P_i} \frac{dP_i}{dr_i} \frac{\partial r_i}{\partial x_t} \right\rangle + \frac{dH^C(P_i)}{dP_i} \frac{dP_i}{dr_i} \frac{\partial r_i}{\partial x_t} = 0 \quad (8a)$$

The requirement is enforced for any given  $r_i$  and  $x_t$ . Therefore, it follows that

$$\left\langle -\sum_{s \in S} V_s \frac{\partial \sigma_s}{\partial P_i} \right\rangle + \frac{dH^C(P_i)}{dP_i} = 0 \quad (8b)$$

The relevant  $x_t$  are the coordinates of the buffer group. Based on Newton's Third Law of Motion, such cancellation will automatically be realized at the active-zone center.

We compute the individual correction derivative  $\frac{dH^C(P_i)}{dP_i}$  as the average energy difference between a pair of partitions where the *only* variation is that the  $i$ -th group at  $r_i$  is treated at QM or MM:

$$\frac{dH^C(P_i)}{dP_i} = \langle V_{i=\text{QM}}(r_i) - V_{i=\text{MM}}(r_i) \rangle = \langle \Delta E_i(r_i) \rangle = \langle \Delta E_i(P_i) \rangle \quad (9)$$

Here, the energies of the partition pair are denoted  $V_{i=\text{QM}}(r_i)$  and  $V_{i=\text{MM}}(r_i)$ , respectively, and  $\langle \rangle$  indicates the corresponding average. As usual, the zeros of the potentials  $V_{i=\text{QM}}$  and  $V_{i=\text{MM}}$  are chosen such that these energies represent the net interaction energies [46]. At any given time step of the simulation, the number of partitions depends on the truncation order of mPAP potential,

and usually there are multiple such pairs of partitions for the  $i$ -th group at  $r_i$ , each with different buffer groups treated by QM. Thus, the average is over the partition pairs for the buffer group at  $r_i$  in a sampled geometry and over all sampled geometries while keeping the buffer group at  $r_i$  in the mPAP simulations. By varying  $r_i$  continuously, one obtains the complete profile of  $\langle \Delta E_i \rangle$ . Again, because  $P_i$  decreases monotonously from 1 to 0 as  $r_i$  increases from  $r_{\min}$  to  $r_{\max}$ , the average energy difference  $\langle \Delta E_i \rangle$ , which depends solely on  $r_i$ , can also be expressed as a function of the dimensionless  $P_i$ .

### 2.2.2. Calculations of $\langle \Delta E_i(P_i) \rangle$ and $H^C(P_i)$

First, we consider one partition pair for a sampled geometry at a given time step. Let us assume that the partition of  $V_{i=\text{QM}}(r_i)$  has a QM subsystem consists of the active-zone "A", the  $i$ -th group, and  $(m - 1)$  other groups selected from the  $N$  buffer groups. Accordingly, the other partition of  $V_{i=\text{MM}}(r_i)$  will feature a QM subsystem consisting of the active-zone "A" and the other  $(m - 1)$  buffer groups, since the  $i$ -th group is treated by MM. These two partitions correspond to the  $m$ -th and  $(m - 1)$ -th order terms in the mPAP potential, respectively. For the sake of brevity, we refer to this partition pair as of the  $m$ -th order. Without losing generality, we label these  $m$  buffer groups by  $1, 2 \dots m$ . The energy difference for this partition pair is

$$\Delta E_i^{(m)}(P_i) = \Delta E_i^{(m)}(r_i) = V_{1,2,\dots,i-1,i+1,\dots,m}^A(r_i) - V_{1,2,\dots,i-1,i+1,\dots,m}^A(r_i) \quad (10)$$

Next, we combine all such partition pairs corresponding to a given  $m$ -th order term in the mPAP Hamiltonian, where there are  $\binom{N-1}{m-1} = \frac{(N-1)!}{(m-1)!(N-m)!}$  such pairs, up to the  $p$ -th order at which the mPAP Hamiltonian is truncated, all for the sampled geometry at the given time step. That is, we take the average of  $\Delta E_i^{(m)}$  over partition pairs for a given  $m$  and then over  $m$  for  $m = 1, 2, \dots, p$ . This gives the average energy difference for a given sampled geometry:

$$\overline{\Delta E_i} = \frac{\Delta E_i^{(1)} + \sum_{\text{all pairs}} \Delta E_i^{(2)} + \sum_{\text{all pairs}} \Delta E_i^{(3)} + \dots}{\binom{N-1}{0} + \binom{N-1}{1} + \binom{N-1}{2} + \dots} \quad (11)$$

We then average  $\overline{\Delta E_i}$  over all sampled geometries in the simulations where the  $i$ -th group is located at  $r_i$ . This gives  $\langle \Delta E_i \rangle$  for the corresponding  $P_i$ . The complete curve of  $\langle \Delta E_i(P_i) \rangle$  is obtained by varying  $P_i$  continuously from 0 to 1. For better statistics, we also average over all buffer groups of the same type in the simulations.

In general, there is no analytic form for  $H^C(P_i)$ , which must be obtained through thermodynamic integration:

$$H^C(P_i) = \int_0^{P_i} \langle \Delta E_i(P'_i) \rangle dP'_i \quad (12)$$

where  $P'_i$  is the dummy variable for integration.

The final correction is applied to each group in the system according to the following scheme:

$$H^C(P_i) = \begin{cases} H^C(0), & r_i > r_{\max} \\ H^C(P_i), & r_{\max} \geq r_i \geq r_{\min} \\ H^C(1), & r_i < r_{\min} \end{cases} \quad (13)$$

We note that  $H^C(0) \neq H^C(1)$  in general.

### 2.2.3. Many-Body Contributions to $\Delta E_i^{(m)}$

To see how many-body interactions are incorporated into the energy difference for a partition pair, we use the standard formula of many-body expansion for a system made of  $n$  monomers:

$$V_{1,2,\dots,n} = \sum_{a=1}^n \varepsilon_a + \sum_{a=1}^{n-1} \sum_{b=a+1}^n \varepsilon_{ab} + \sum_{a=1}^{n-2} \sum_{b=a+1}^{n-1} \sum_{c=b+1}^n \varepsilon_{abc} + \dots \quad (14)$$

where  $\varepsilon_a$  is the energy of an isolated monomer,  $\varepsilon_{ab}$  is the energy of a dimer minus the energies of the monomers it comprises (i.e., it is the pairwise interaction energy), etc. Applying Equation (14) to  $V_{1,2,\dots,i-1,i,i+1,\dots,m}^A(r_i)$  and  $V_{1,2,\dots,i-1,i+1,\dots,m}^A(r_i)$ , and after canceling identical terms, one has

$$\Delta E_i^{(m)} = (\Delta \varepsilon_i + \Delta \varepsilon_{Ai}) + \sum_{j \neq i}^m (\Delta \varepsilon_{ji} + \Delta \varepsilon_{Aji}) + \sum_{j \neq i}^m \sum_{k \neq j \neq i}^m (\Delta \varepsilon_{jki} + \Delta \varepsilon_{Ajki}) + \dots \quad (15)$$

where

$$\Delta \varepsilon_i = \varepsilon_{i=QM} - \varepsilon_{i=MM} \quad (16a)$$

$$\Delta \varepsilon_{Ai} = \varepsilon_{A,i=QM} - \varepsilon_{A=QM}; i=MM \quad (16b)$$

$$\Delta \varepsilon_{ji} = \varepsilon_{j,i=QM} - \varepsilon_{j=QM}; i=MM \quad (16c)$$

$$\Delta \varepsilon_{Aji} = \varepsilon_{A,j,i=QM} - \varepsilon_{A,k=QM}; i=MM \quad (16d)$$

$$\Delta \varepsilon_{jki} = \varepsilon_{j,k,i=QM} - \varepsilon_{j,k=QM}; i=MM \quad (16e)$$

$$\Delta \varepsilon_{Ajki} = \varepsilon_{A,j,k,i=QM} - \varepsilon_{A,j,k=QM}; i=MM \quad (16f)$$

...

Here  $\varepsilon_{i=QM}$  and  $\varepsilon_{i=MM}$  are the QM and MM energies of the  $i$ -th group at the QM and MM levels, respectively;  $\varepsilon_{j,i=QM}$  is the pairwise interaction energy between groups  $i$  and  $j$ , both treated at the QM level;  $\varepsilon_{j=QM}; i=MM$  is the pairwise interaction energy between group  $j$  described at the QM level and group  $i$  at the MM level;  $\varepsilon_{j,k,i=QM}$  is the three-body contribution by groups  $i$ ,  $j$ , and  $k$  treated at the QM level;  $\varepsilon_{j,k=QM}; i=MM$  is the three-body contribution by groups  $j$  and  $k$  described at the QM level and group  $i$  at the MM level ... Terms involving the active-zone "A" are defined similarly. The many-body interactions where the  $i$ -th group treated by MM depend on the specific QM/MM embedding model. Note that these many-body interactions never need to be explicitly evaluated; they are prescribed here merely to assist understanding of the many-body interactions in  $\Delta E_i^{(m)}$ . The above analysis indicates that  $\Delta E_i^{(m)}$  is a sum of the differences between the many-body terms in the partition pair involving the  $i$ -th group in the many-body expansions up to the  $m$ -th order.

### 2.2.4. Effects Due to Truncation in mPAP Hamiltonian

If the mPAP potential is truncated at a low order  $p$  and if  $N \gg p$ , which is the case in our current simulations, the average is dominated by the  $p$ -th order  $\Delta E_i^{(p)}$ , which has the largest number of partitions:

$$\langle \Delta E_i \rangle \approx \langle \Delta E_i^{(p)} \rangle \quad (17)$$

Moreover, because all groups are the same type of molecules,

$$\begin{aligned} \langle \Delta E_i^{(p)} \rangle &= (\langle \Delta \varepsilon_i \rangle + \langle \Delta \varepsilon_{Ai} \rangle) + (p-1)(\langle \Delta \varepsilon_{ji} \rangle + \langle \Delta \varepsilon_{Aji} \rangle) \\ &+ \frac{(p-1)(p-2)}{2} (\langle \Delta \varepsilon_{jki} \rangle + \langle \Delta \varepsilon_{Ajki} \rangle) + \dots \end{aligned} \quad (18)$$

That is,  $\langle \Delta E_i \rangle$  depends on the truncation order in the mPAP potential. A special case is dual-MM, where effective two-body potentials are employed, the third and higher order terms are 0 in Equation (14). Consequently,  $\langle \Delta \varepsilon_{\Delta ji} \rangle$ ,  $\langle \Delta \varepsilon_{jki} \rangle$ , and higher-order contributions vanish, leading to

$$\langle \Delta E_i^{(p)} \rangle = (\langle \Delta \varepsilon_i \rangle + \langle \Delta \varepsilon_{\Delta i} \rangle) + (p - 1) \langle \Delta \varepsilon_{ji} \rangle \quad (19)$$

### 2.3. Simulation Details

An Open-MP-based parallel version of HPAP method has been implemented in our in-house code, QMMM [54]. The QMMM code calls an MM program such as NAMD [55] for MM calculations and a QM program such as MNDO [56] for QM calculations, synthesizes the energies and gradients from both, and propagates the trajectory. Here, to reduce the computational expense from QM/MM calculations, we carry out dual-MM (MM'/MM) simulations, as we have done previously in the development of the original PAP algorithm [29] and as Boreboom et al. [42] did in their HAMBC study. Since we are only focusing on the adaptive-boundary treatments, the use of dual-MM test calculations suffices.

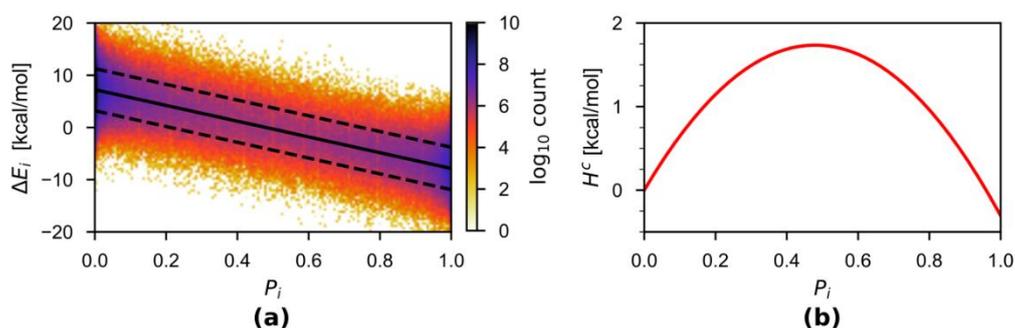
The model system for all calculations is a  $30.25 \times 30.25 \times 30.25 \text{ \AA}^3$  water box that contains 915 water molecules (density = 0.99 g/mL). The water box was prepared by equilibration for 10 ns under the *NVT* ensemble at the single-MM level using the TIP3P/Fs [57,58] water model. For the adaptive-partitioning simulations, each water molecule is an adaptive partitioning group, with the oxygen atom designated as the representative atom for the group. One water oxygen was chosen to be the active-zone center for the simulation, and the buffer and active zone radii were set to  $r_{\max} = 6.4 \text{ \AA}$  and  $r_{\min} = 5.5 \text{ \AA}$  respectively. All adaptive-partitioning potentials were truncated at the 2nd order. For all calculations, periodic boundary conditions with the minimum-image convention were adopted. The cutoff for non-bonded interactions was set to 12  $\text{\AA}$  with smoothing switched on at 11  $\text{\AA}$ . All simulations are under constant volume with a time step of 0.5 fs. If constant temperatures were required, a Nosé-Hoover [59] thermostat of 298.15 K was coupled to the model system with a coupling coefficient of 40 fs.

First, to obtain the correction term  $H^C$ , the model was simulated at constant temperature at the MM'/MM level using the mPAP scheme, where MM' = SPC/Fw [58] and MM = TIP3P/Fs [57,58] (the force field parameters are compared in Table 1). Five 20-ps trajectories were propagated. Only the second half 10 ps of every trajectory was utilized to determine the correction term  $H^C$ , during which the energies of the partitions were recorded. Note that multiple partitions and thus multiple energy differences are available at one single time step, unless there is no or only one buffer group. The combined 50-ps trajectories for the  $H^C$  calculations resulted in a total of 16,791,809 energy differences ( $\Delta E_i^{(1)}$  and  $\Delta E_i^{(2)}$ ). Figure 2 shows the density distribution of  $\Delta E_i(P_i)$ . We found that the 50 ps combined simulation time was sufficient to converge the correction term in the present work (more details in the Results section). These energy differences  $\Delta E_i(P_i)$  were added to 100 bins equally spaced over the domain [0, 1] according to the  $P_i$  value of the buffer group. The average of each bin was then taken and used as input for linear regression to obtain the function  $\Delta E(P_i)$ . We have dropped the subscript  $i$ , because  $\Delta E_i(P_i)$  is identical for all buffer groups here. The linear regression was motivated by the approximately linear nature of the  $\Delta E(P_i)$  curve (more details in the Discussion section). The analytical integral of the regression line was used to obtain  $H^C(P_i)$ . Both  $\Delta E(P_i)$  and  $H^C(P_i)$  were discretized into 10,000 evenly spaced points between the domain [0, 1] and are stored as a look-up table read in by the QMMM source code. The correction for a group was chosen by looking up the  $\Delta E(P_i)$  and  $H^C(P_i)$  values in the table for the next largest  $P_i$  value.

**Table 1.** Force field parameters for the water models employed in this work.

Interaction	Parameter	TIP3P/Fs <sup>1</sup>	SPC/Fw <sup>2</sup>
O-H stretching	$k_b$ [kcal/mol/Å <sup>2</sup> ]	1059.162	1059.162
	$B_0$ [Å]	0.960	1.012
H-O-H bending	$k_\theta$ [kcal/mol/rad <sup>2</sup> ]	68.097	75.900
	$\theta_0$ [deg]	104.500	113.24
Electrostatic	$q_H$ [e]	0.417	0.41
	$q_O$ [e]	−0.834	−0.82
van der Waals (O only)	$R_{\min}/2$ [Å]	1.7682	1.7767
	$\epsilon$ [kcal/mol]	−0.1522	−0.1524

<sup>1</sup> Ref. [58] <sup>2</sup> Refs. [57,58].



**Figure 2.** (a) Density plot of  $\Delta E(P_i)$  computed from the combined 50-ps simulations. The solid and dashed black lines indicated the  $\langle \Delta E(P_i) \rangle$  and the associated standard deviation, respectively. Linear regression yielded an analytical representation for  $\langle \Delta E(P_i) \rangle = (-15.087P_i + 7.236)$  kcal·mol<sup>−1</sup> with  $R^2 = 0.9997$ . The average standard deviation over the 100 bins is 4.045 kcal·mol<sup>−1</sup>. (b) The energy correction term  $H^C(P_i)$  computed by integration using the analytical representation of  $\langle \Delta E(P_i) \rangle$ .

Second, to examine the solvation structures obtained through the HAMBC correction, we carried out adaptive-partitioning simulations at constant temperature using three schemes: PAP, mPAP, and HPAP. For each scheme, five trajectories were propagated using the final geometries and velocities from the five mPAP simulations used to compute  $H^C$ . Each trajectory was propagated for 10 ps. The combined 50 ps trajectories were employed to calculate the pairwise radial distribution function for a given scheme.

Finally, to examine the degree of energy conservation, we also performed NVE simulations for all three schemes. One trajectory for each method was propagated for 25 ps with 0.5 fs time steps.

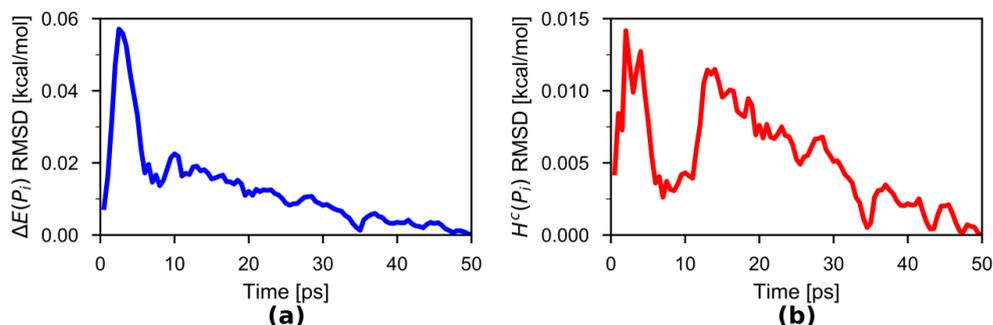
### 3. Results

#### 3.1. HAMBC Correction Term

Figure 2a plots the energy difference over the thermodynamic variable  $\Delta E(P_i)$  obtained from the combined 50-ps trajectories, together with the linear regression results. The almost perfect linear fit is incidental due to the similarity in some parameters of the two employed force fields (see Section 4 for detailed discussion). In general, the linearity is very approximate or does not hold. Nevertheless, by taking advantage of the analytical representation of  $\Delta E(P_i)$  from the linear regression, we estimated the correction term  $H^C(P_i)$ , which is depicted in Figure 2b.

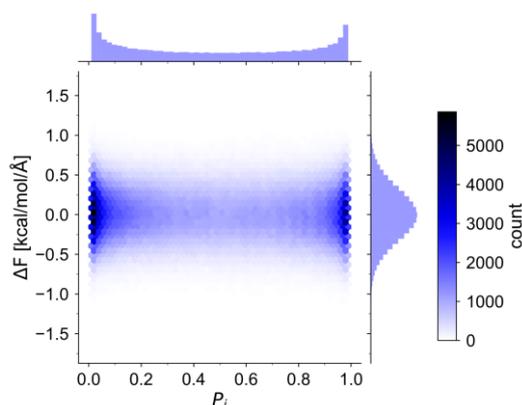
To explore the length of simulation needed to generate the HPAP correction, we calculated the root mean squared deviation (RMSD) of  $\Delta E(P_i)$  and  $H^C(P_i)$  as functions of simulation time, taking as reference the final curves from the accumulated 50-ps simulations (Figure 3). The RMSD values of  $\Delta E(P_i)$  and  $H^C(P_i)$  were calculated over the 100 bins that are equally distributed bins along the  $P_i$

domain  $[0, 1]$ . It can be seen that both RMSDs drop consistently after  $\sim 15$  ps. The function  $\Delta E(P_i)$  converged to below  $0.01 \text{ kcal mol}^{-1}$  after 30 ps of combined simulations. Because integration reduces the fluctuations in  $\Delta E(P_i)$ , the RMSD values of  $H^C(P_i)$  are even smaller than those of  $\Delta E(P_i)$ .



**Figure 3.** Root mean squared deviations (RMSDs) of (a)  $\Delta E(P_i)$  and (b)  $H^C(P_i)$ , respectively, as functions of simulation time, taking their final curves from the combined 50-ps simulations as references.

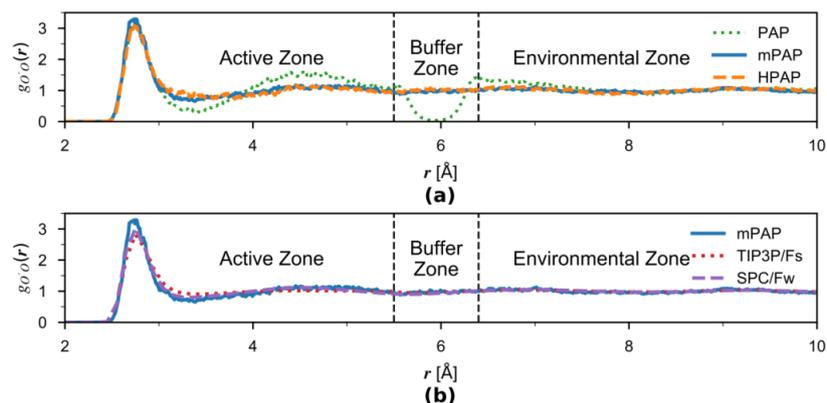
A necessary requirement for the correction term is that the correction force  $F_{\text{corr}}$  must, on average, cancel out the transition force  $F_{\text{trans}}$ . This fact is exemplified in the distribution of the difference between these two forces ( $\Delta F = F_{\text{trans}} - F_{\text{corr}}$ ) in the final HPAP simulations (Figure 4). It can be seen that  $\Delta F$  is distributed approximately normally around  $0 \text{ kcal}\cdot\text{mol}^{-1}\text{\AA}^{-1}$  across the entire domain of  $P_i$ . The calculated average  $\Delta F$  over all  $P_i$  is  $0.0007 \text{ kcal}\cdot\text{mol}^{-1}\text{\AA}^{-1}$ , with a standard deviation of  $0.4 \text{ kcal}\cdot\text{mol}^{-1}\text{\AA}^{-1}$ , in excellent agreement with the cancellation requirement.



**Figure 4.** Bivariate distribution of the difference between the transition forces and the correction forces (defined as  $\Delta F = F_{\text{trans}} - F_{\text{corr}}$ ) in the HPAP simulations.

### 3.2. Solvation Structures

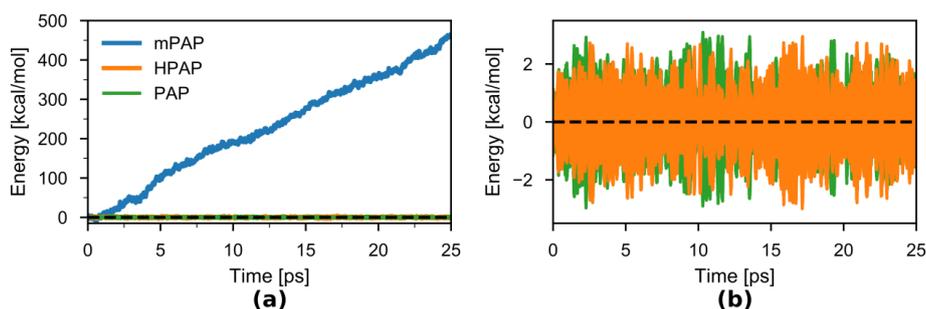
We examined the solvation structure around the solute water by computing the pairwise radial distribution function (RDF)  $g_{O'O}(r)$  between its oxygen ( $O'$ , serving as the center of the active zone) and the surrounding oxygen (O) atoms for the original PAP, mPAP, and HPAP simulations under constant temperature (Figure 5a). The original PAP results show an erroneous dip around  $r = 5.95 \text{ \AA}$  which corresponds to the center of the buffer zone. This is caused by the transition forces pushing water out of the buffer zone. However, this artifact is eliminated by inclusion of the correction term in HPAP, for which the RDF closely matches the mPAP reference data, indicating that the HPAP method is able to produce accurate solvation structures. The results here are in line with what were found by Boreboom et al. [42] although different adaptive schemes are employed.



**Figure 5.** (a) Comparison of radial distribution functions (RDF) between AP simulations. The water oxygen atom  $O'$  serves as the active-zone center, and  $O$  is the surrounding water oxygen atom. The results are illustrated by a green dotted line for the PAP, orange dashed line for the HPAP, and blue solid line for the reference mPAP methods, respectively. The buffer zone boundaries at  $r_{\max} = 6.4 \text{ \AA}$  and  $r_{\min} = 5.5 \text{ \AA}$  are indicated by the two black dashed vertical lines. (b) Comparisons of RDF between the mPAP and the two single-level MM simulations.

For comparison, we also present in Figure 5b the RDF computed from the mPAP and the two single-level MM simulations. All three RDF curves are overall quite similar, although the TIP3P/Fs water is slightly less structural than SPC/Fw. The mPAP RDF better resembles the SPC/Fw curve in the active-zone, as it is supposed to be. Interestingly, despite the similarity between SPC/Fw and TIP/Fs, a significantly distorted RDF was produced by PAP in the buffer-zone region, as observed in Figure 5a. This demonstrates the sensitivity of the RDF to the boundary treatment in the adaptive simulations and underscores the importance of proper treatments implemented in mPAP and HPAP.

Next, we checked the degree of energy conservation in the *NVE* simulations (Figure 6a,b). As expected, there was a large energy drift during the duration of the mPAP simulation of  $17 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{ps}^{-1}$ . On the other hand, both the original PAP and the newly introduced HPAP show negligible drifts in energy ( $0.004$  and  $0.015 \text{ kcal}\cdot(\text{mol}^{-1}\cdot\text{ps}^{-1})$ , respectively) over the duration of the 25-ps simulations. The root mean squared deviations of the energy over the timeseries were almost the same for HPAP ( $\sim 0.82 \text{ kcal}\cdot\text{mol}^{-1}$ ) and PAP ( $\sim 0.76 \text{ kcal}\cdot\text{mol}^{-1}$ ).



**Figure 6.** (a) Total energy as the sum of kinetic and potential energies of the model system in the 25-ps *NVE* simulations using the PAP (green), mPAP (blue), and HPAP (orange) methods. The many-body expansions of the potentials were all truncated at the 2nd order. (b) PAP and HPAP Total energies in greater detail. The zero of energy was chosen to be the mean energy of the first 250 fs.

#### 4. Discussion

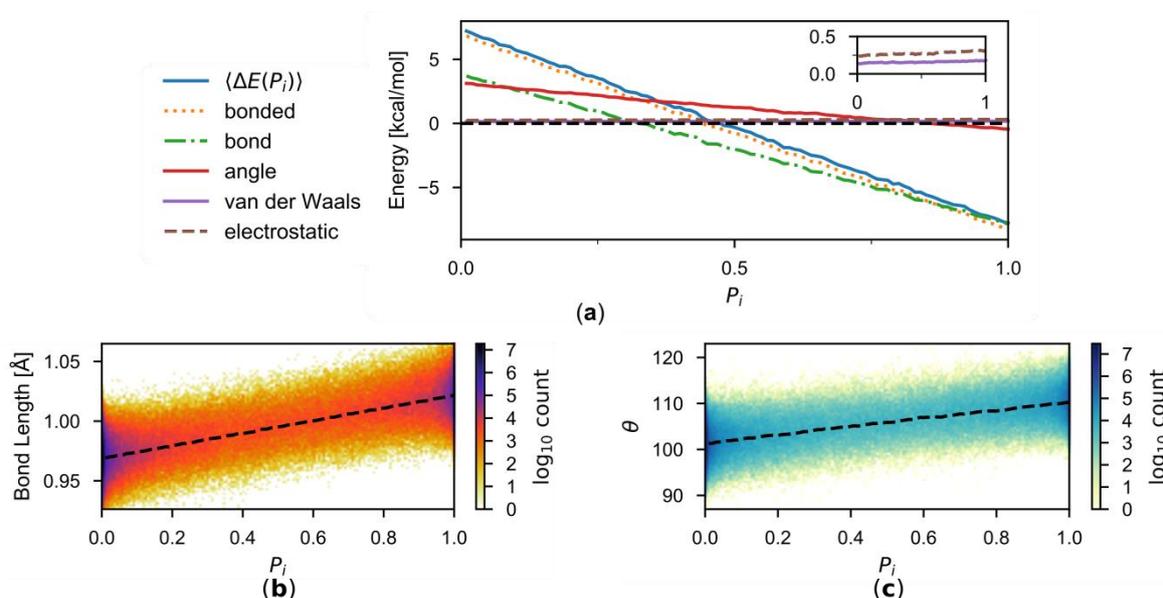
It is interesting to note that the correction  $\langle \Delta E(P_i) \rangle$  is approximately a linear function of  $P_i$  in the present work. To understand the origin of this approximate linearity, let us consider a given  $i$ -th

buffer group. When the description of this buffer groups is switched between the two employed MM force fields, SPC/Fw and TIP3P/Fs, there will be changes in the non-bonded (van der Waals and electrostatic) interactions through which this buffer group interacts with active-zone groups, environmental-zone groups, and the other buffer groups. Also changing are the intramolecular bonded (O-H bond and H-O-H angle) interactions within this buffer group. The decomposition of  $\langle \Delta E_i(P_i) \rangle$  is depicted in Figure 7a according to

$$\langle \Delta E_i \rangle = \langle V_{i=\text{SPC/Fw}} - V_{i=\text{TIP3P/Fs}} \rangle \quad (20a)$$

$$= \langle \Delta E_{i,\text{bonded}} \rangle + \langle \Delta E_{i,\text{nonbonded}} \rangle = \langle \Delta E_{i,\text{bond}} \rangle + \langle \Delta E_{i,\text{angle}} \rangle + \langle \Delta E_{i,\text{vdw}} \rangle + \langle \Delta E_{i,\text{elec}} \rangle \quad (20b)$$

where  $i$  denotes the  $i$ -th buffer group.



**Figure 7.** (a) Decomposition of  $\langle \Delta E(P_i) \rangle$  into various energy components. The inset shows the van der Waals and electrostatic components in greater detail. (b) The distribution of O-H bond as functions of  $P_i$ . The average is indicated by the dashed curve. (c) Same as (b), but for H-O-H angle.

The bonded energy terms are the dominant contributors to  $\Delta E_i$ . This is not surprising because of the similarity in the nonbonded interaction parameters (Table 1). Both the O-H bond and H-O-H angle energies are sizeable, but the O-H bond energy is overall more significant. The difference in the O-H bond energies between the two descriptions are

$$\Delta E_{\text{bond}} = k_1(x - x_{01})^2 - k_2(x - x_{02})^2 \quad (21)$$

where  $x$  is the instantaneous O-H bond length, and  $k_1$  and  $k_2$  are the force constants parameters and  $x_{01}$  and  $x_{02}$  the equilibrium geometric parameters of the two water models, respectively. If  $k_1 = k_2 = k$ , as it is for the two water models employed here, one has

$$\Delta E_{\text{bond}} = -2k(x_{01} - x_{02})x + k(x_{01}^2 - x_{02}^2) \quad (22)$$

This means that the energy difference will vary linearly with respect to the instantaneous bond length. Figure 7b suggests smooth structural changes of the water molecule in the buffer zone between the TIP3P/Fs model at  $P_i = 0$  and the SPC/Fw model at  $P_i = 1$  and confirms that the average instantaneous O-H bond length in the simulations is indeed approximately linear with respect to  $P_i$ .

The situation is similar in the H-O-H angle energy (Figure 7c), where  $k_1 \approx k_2 = k$ , and the linearity also roughly holds. As a result,  $\langle \Delta E(P_i) \rangle$  is approximately linear with respect to  $P_i$ .

Of course, the above analysis can only be conducted when the interactions are pairwise, which is true here for the dual-MM simulations. However, because pairwise interactions are usually predominant (e.g., accounting for ~80% of total interaction energies in water [60]), the approximate linearity may still hold when models with higher-order many-body potentials are employed, albeit to a less extent, unless the two employed potentials differ significantly from each other. In reality, the potentials should agree with each other reasonably well in the buffer region, otherwise it is likely that at least one potential is very inaccurate and should not be used at all.

We note that while the forces by the correction term cancel the average transition forces, the cancellation is not exact at every step. Inexact transient cancellations lead to “residue” forces, whose effects on the simulation may or may not lead to erroneous solvation structures, which probably varies from case to case. While further investigations will be needed to explore this in the future, it is conceivable that narrow, symmetric distributions of the residue forces can help to minimize their effects, which is the case in this work. Therefore, it is prudent to match the potentials as closely as possible in the buffer zone. At this point, we note that Jiang et al. [61] are developing MM water models specifically designed for adaptive QM/MM simulations.

Overall, the results presented here demonstrate that the HPAP can both yield accurate solvation structures and conserve energy in NVE simulations. The progress thus fills a gap in the PAP algorithms. The successful applications of the HAMBC treatment to both SAP by Boreboom et al. [42] and PAP in this work suggest that other distance-based adaptive QM/MM schemes may also benefit from this treatment. We expect that the new HPAP algorithm will be useful to many applications where simulations of Hamiltonian systems are required. Future studies are especially encouraged to investigate the cases where multiple types of buffer groups (e.g., water and ions) are present and to explore the treatments for inhomogeneous systems (e.g., ion channels).

**Author Contributions:** Conceptualization, H.L.; Methodology, A.W.D. and H.L.; Software, A.W.D. and C.-H.W.; Validation, A.W.D. and C.-H.W.; Formal Analysis, A.W.D. and C.-H.W.; Writing-Original Draft Preparation, A.W.D.; Writing-Review & Editing, A.W.D., C.-H.W., and H.L.; Supervision, H.L.; Project Administration, H.L.; Funding Acquisition, H.L.

**Funding:** This work is supported by the NSF (CHE-1564349), Camille & Henry Dreyfus Foundation (TH-14-028), and NVIDIA Corporation. This work used XSEDE under grant CHE-140070, supported by NSF grant number ACI-1053575, and NERSC under grant m2495.

**Acknowledgments:** We thank Danielle Miller and Christina Garza for helpful discussion.

**Conflicts of Interest:** The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249. [[CrossRef](#)]
2. Singh, U.C.; Kollmann, P.A. A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the  $\text{CH}_3\text{Cl} + \text{Cl}^-$  exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* **1986**, *7*, 718–730. [[CrossRef](#)]
3. Field, M.J.; Bash, P.A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700–733. [[CrossRef](#)]
4. Gao, J. Methods and applications of combined quantum mechanical and molecular mechanical potentials. *Rev. Comput. Chem.* **1996**, *7*, 119–185.
5. Monard, G.; Merz, K.M., Jr. Combined quantum mechanical/molecular mechanical methodologies applied to biomolecular systems. *Acc. Chem. Res.* **1999**, *32*, 904–911. [[CrossRef](#)]
6. Hammes-Schiffer, S. Theoretical perspectives on proton-coupled electron transfer reactions. *Acc. Chem. Res.* **2000**, *34*, 273–281. [[CrossRef](#)]

7. Sherwood, P. Hybrid quantum mechanics/molecular mechanics approaches. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; John von Neumann-Institut: Jülich, German, 2000; Volume 3, pp. 285–305.
8. Gao, J.; Truhlar, D.G. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505. [[CrossRef](#)] [[PubMed](#)]
9. Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; et al. Development of effective quantum mechanical/molecular mechanical (QM/MM) methods for complex biological process. *J. Phys. Chem. B* **2006**, *110*, 6458–6469. [[CrossRef](#)] [[PubMed](#)]
10. Lin, H.; Truhlar, D.G. QM/MM: What have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185–199. [[CrossRef](#)]
11. Senn, H.M.; Thiel, W. QM/MM methods for biological systems. *Top. Curr. Chem.* **2007**, *268*, 173–290.
12. Hu, H.; Yang, W. Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601. [[CrossRef](#)] [[PubMed](#)]
13. Sherwood, P.; Brooks, B.R.; Sansom, M.S.P. Multiscale methods for macromolecular simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 630–640. [[CrossRef](#)] [[PubMed](#)]
14. Sabin, J.R.; Brändas, E. *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Academic Press: San Diego, CA, USA, 2010; pp. 1–416.
15. Ferrer, S.; Ruiz-Pernia, J.; Marti, S.; Moliner, V.; Tunon, I.; Bertran, J.; Andres, J. Hybrid schemes based on quantum mechanics/molecular mechanics simulations: Goals to success, problems, and perspectives. In *Advances in Protein Chemistry and Structural Biology, Vol. 85: Computational Chemistry Methods in Structural Biology*; Christov, C., Ed.; Elsevier Academic Press Inc.: San Diego, CA, USA, 2011; Volume 85, pp. 81–142.
16. Menikarachchi, L.C.; Gascon, J.A. QM/MM approaches in medicinal chemistry research. *Curr. Top. Med. Chem.* **2010**, *10*, 46–54. [[CrossRef](#)] [[PubMed](#)]
17. Wallrapp, F.H.; Guallar, V. Mixed quantum mechanics and molecular mechanics methods: Looking inside proteins. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 315–322. [[CrossRef](#)]
18. Woodcock, H.L.; Miller, B.T.; Hodoscek, M.; Okur, A.; Larkin, J.D.; Ponder, J.W.; Brooks, B.R. MSCALE: A general utility for multiscale modeling. *J. Chem. Theory Comput.* **2011**, *7*, 1208–1219. [[CrossRef](#)] [[PubMed](#)]
19. Chung, L.W.; Hirao, H.; Li, X.; Morokuma, K. The ONIOM method: Its foundation and applications to metalloenzymes and photobiology. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 327–350. [[CrossRef](#)]
20. Lonsdale, R.; Harvey, J.N.; Mulholland, A.J. A practical guide to modelling enzyme-catalysed reactions. *Chem. Soc. Rev.* **2012**, *41*, 3025–3038. [[CrossRef](#)] [[PubMed](#)]
21. Monari, A.; Rivail, J.-L.; Assfeld, X. Theoretical modeling of large molecular systems. Advances in the local self consistent field method for mixed quantum mechanics/molecular mechanics calculations. *Acc. Chem. Res.* **2012**, *46*, 596–603. [[CrossRef](#)] [[PubMed](#)]
22. Wu, R.; Cao, Z.; Zhang, Y. Computational simulations of zinc enzyme: Challenges and recent advances. *Prog. Chem.* **2012**, *24*, 1175–1184.
23. Mennucci, B. Modeling environment effects on spectroscopies through QM/classical models. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6583–6594. [[CrossRef](#)] [[PubMed](#)]
24. Meier, K.; Choutko, A.; Dolenc, J.; Eichenberger, A.P.; Riniker, S.; van Gunsteren, W.F. Multi-resolution simulation of biomolecular systems: A review of methodological issues. *Angew. Chem. Int. Ed.* **2013**, *52*, 2820–2834. [[CrossRef](#)] [[PubMed](#)]
25. Pezeshki, S.; Lin, H. Recent developments in QM/MM methods towards open-boundary multi-scale simulations. *Mol. Simulat.* **2015**, *41*, 168–189. [[CrossRef](#)]
26. Duarte, F.; Amrein, B.A.; Blaha-Nelson, D.; Kamerlin, S.C.L. Recent advances in QM/MM free energy calculations using reference potentials. *BBA-Gen. Subj.* **2015**, *1850*, 954–965. [[CrossRef](#)] [[PubMed](#)]
27. Kerdcharoen, T.; Liedl, K.R.; Rode, B.M. A QM/MM simulation method applied to the solution of Li<sup>+</sup> in liquid ammonia. *Chem. Phys.* **1996**, *211*, 313–323. [[CrossRef](#)]
28. Kerdcharoen, T.; Morokuma, K. ONIOM-XS: An extension of the ONIOM method for molecular simulation in condensed phase. *Chem. Phys. Lett.* **2002**, *355*, 257–262. [[CrossRef](#)]
29. Heyden, A.; Lin, H.; Truhlar, D.G. Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for multiscale simulations. *J. Phys. Chem. B* **2007**, *111*, 2231–2241. [[CrossRef](#)] [[PubMed](#)]

30. Bulo, R.E.; Ensing, B.; Sikkema, J.; Visscher, L. Toward a practical method for adaptive QM/MM simulations. *J. Chem. Theory Comput.* **2009**, *5*, 2212–2221. [[CrossRef](#)] [[PubMed](#)]
31. Nielsen, S.O.; Bulo, R.E.; Moore, P.B.; Ensing, B. Recent progress in adaptive multiscale molecular dynamics simulations of soft matter. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12401–12414. [[CrossRef](#)] [[PubMed](#)]
32. Pezeshki, S.; Lin, H. Adaptive-partitioning redistributed charge and dipole schemes for QM/MM dynamics simulations: On-the-fly relocation of boundaries that pass through covalent bonds. *J. Chem. Theory Comput.* **2011**, *7*, 3625–3634. [[CrossRef](#)] [[PubMed](#)]
33. Takenaka, N.; Kitamura, Y.; Koyano, Y.; Nagaoka, M. The number-adaptive multiscale QM/MM molecular dynamics simulation: Application to liquid water. *Chem. Phys. Lett.* **2012**, *524*, 56–61. [[CrossRef](#)]
34. Bulo, R.E.; Michel, C.; Fleurat-Lessard, P.; Sautet, P. Multiscale modeling of chemistry in water: Are we there yet? *J. Chem. Theory Comput.* **2013**, *9*, 5567–5577. [[CrossRef](#)] [[PubMed](#)]
35. Pezeshki, S.; Davis, C.; Heyden, A.; Lin, H. Adaptive-partitioning QM/MM dynamics simulations: 3. Solvent molecules entering and leaving protein binding sites. *J. Chem. Theory Comput.* **2014**, *10*, 4765–4776. [[PubMed](#)]
36. Waller, M.P.; Kumbhar, S.; Yang, J. A density-based adaptive quantum mechanical/molecular mechanical method. *ChemPhysChem* **2014**, *15*, 3218–3225. [[CrossRef](#)] [[PubMed](#)]
37. Watanabe, H.C.; Kubař, T.; Elstner, M. Size-consistent multipartitioning QM/MM: A stable and efficient adaptive QM/MM method. *J. Chem. Theory Comput.* **2014**, *10*, 4242–4252. [[CrossRef](#)] [[PubMed](#)]
38. Böckmann, M.; Doltsinis, N.L.; Marx, D. Adaptive switching of interaction potentials in the time domain: An extended lagrangian approach tailored to transmute force field to QM/MM simulations and back. *J. Chem. Theory Comput.* **2015**, *11*, 2429–2439. [[CrossRef](#)] [[PubMed](#)]
39. Jiang, T.; Boereboom, J.M.; Michel, C.; Fleurat-Lessard, P.; Bulo, R.E. Proton transfer in aqueous solution: Exploring the boundaries of adaptive QM/MM. In *Quantum Modeling of Complex Molecular Systems*; Rivail, J.-L., Ruiz-Lopez, M., Assfeld, X., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 51–91.
40. Pezeshki, S.; Lin, H. Adaptive-partitioning QM/MM for molecular dynamics simulations: 4. Proton hopping in bulk water. *J. Chem. Theory Comput.* **2015**, *11*, 2398–2411. [[CrossRef](#)] [[PubMed](#)]
41. Pezeshki, S.; Lin, H. Recent developments in adaptive QM/MM. In *Quantum Modeling of Complex Molecular Systems*; Rivail, J.-L., Ruiz-Lopez, M., Assfeld, X., Eds.; Springer: New York, NY, USA, 2015; pp. 93–113.
42. Boereboom, J.M.; Potestio, R.; Donadio, D.; Bulo, R.E. Toward Hamiltonian adaptive QM/MM: Accurate solvent structures using many-body potentials. *J. Chem. Theory Comput.* **2016**, *12*, 3441–3448. [[CrossRef](#)] [[PubMed](#)]
43. Duster, A.; Garza, C.; Lin, H. Adaptive partitioning QM/MM dynamics simulations for substrate uptake, product release, and solvent exchange. In *Computational Approaches for Studying Enzyme Mechanism*; Voth, G.A., Ed.; Elsevier: Cambridge, UK, 2016; pp. 342–358.
44. Zheng, M.; Waller, M.P. Adaptive quantum mechanics/molecular mechanics methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2016**, *6*, 369–385. [[CrossRef](#)]
45. Dohm, S.; Spohr, E.; Korth, M. Developing adaptive QM/MM computer simulations for electrochemistry. *J. Comput. Chem.* **2017**, *38*, 51–58. [[CrossRef](#)] [[PubMed](#)]
46. Duster, A.W.; Wang, C.-H.; Garza, C.M.; Miller, D.E.; Lin, H. Adaptive quantum/molecular mechanics: What have we learned, where are we, and where do we go from here? *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1310. [[CrossRef](#)]
47. Field, M.J. An algorithm for adaptive QC/MM simulations. *J. Chem. Theory Comput.* **2017**, *13*, 2342–2351. [[CrossRef](#)] [[PubMed](#)]
48. Zheng, M.; Kuriappan, J.A.; Waller, M.P. Toward more efficient density-based adaptive QM/MM methods. *Int. J. Quantum Chem.* **2017**, *117*, e25336. [[CrossRef](#)]
49. Boereboom, J.M.; Fleurat-Lessard, P.; Bulo, R.E. Explicit solvation matters: Performance of QM/MM solvation models in nucleophilic addition. *J. Chem. Theory Comput.* **2018**, *14*, 1841–1852. [[CrossRef](#)] [[PubMed](#)]
50. Hofer, T.S.; Hünenberger, P.H. Absolute proton hydration free energy, surface potential of water, and redox potential of the hydrogen electrode from first principles: QM/MM MD free-energy simulations of sodium and potassium hydration. *J. Chem. Phys.* **2018**, *148*, 222814. [[CrossRef](#)] [[PubMed](#)]
51. Zheng, M.; Waller, M.P. Yoink: An interaction-based partitioning API. *J. Comput. Chem.* **2018**, *39*, 799–806. [[CrossRef](#)] [[PubMed](#)]

52. Watanabe, H. Improvement of performance, stability and continuity by modified size-consistent multipartitioning quantum mechanical/molecular mechanical method. *Molecules* **2018**, *23*, 1882. [[CrossRef](#)] [[PubMed](#)]
53. Potestio, R.; Fritsch, S.; Español, P.; Delgado-Buscalioni, R.; Kremer, K.; Everaers, R.; Donadio, D. Hamiltonian adaptive resolution simulation for molecular liquids. *Phys. Rev. Lett.* **2013**, *110*, 108301. [[CrossRef](#)] [[PubMed](#)]
54. Lin, H.; Zhang, Y.; Pezeshki, S.; Duster, A.; Truhlar, D.G. *QMMM*, Version 2.2.8.CO; University of Minnesota: Minneapolis, MN, USA, 2017.
55. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [[CrossRef](#)] [[PubMed](#)]
56. Thiel, W. *MNDO2005*, Version 7.0; Max-Planck-Institut für Kohlenforschung: Mülheim an der Ruhr, Germany, 2005.
57. Schmitt, U.W.; Voth, G.A. The computer simulation of proton transport in water. *J. Chem. Phys.* **1999**, *111*, 9361–9381. [[CrossRef](#)]
58. Wu, Y.; Tepper, H.L.; Voth, G.A. Flexible simple point-charge water model with improved liquid-state properties. *J. Chem. Phys.* **2006**, *124*, 024503. [[CrossRef](#)] [[PubMed](#)]
59. Hoover, W.G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697. [[CrossRef](#)]
60. Cisneros, G.A.; Wikfeldt, K.T.; Ojamäe, L.; Lu, J.; Xu, Y.; Torabifard, H.; Bartók, A.P.; Csányi, G.; Molinero, V.; Paesani, F. Modeling molecular interactions in water: From pairwise to many-body potential energy functions. *Chem. Rev.* **2016**, *116*, 7501–7528. [[CrossRef](#)] [[PubMed](#)]
61. Jiang, T.; Simko, S.; Bulo, R.E. Accurate QM/MM simulation of aqueous solutions with tailored MM models. *J. Chem. Theory Comput.* **2018**, *14*, 3943–3954. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).