

BREQM: Boundary Region Embedding Quantum Mechanics/Molecular Mechanics

Willis O’Leary
Mentor: William A. Goddard

March 21, 2018

Summary

Quantum Mechanics / Molecular Mechanics, or QM/MM, molecular dynamic simulations can help determine the complex and mysterious mechanisms of heterogeneous water-splitting catalysts. QM/MM is well suited for these systems since they are naturally split into a QM region (the catalyst) and an MM region (the solvent). However, conventional methods for describing the critical interactions between the QM and MM regions often fall short. In response, we propose and implement a new methodology: boundary region embedding QM/MM (BREQM). In BREQM, we introduce boundary regions through which the QM and MM regions can influence one another. Compared to conventional QM/MM, BREQM can more easily and consistently model water-splitting catalysts. We believe that mechanistic information gleaned from future BREQM simulations will lead to better water-splitting catalysts.

Introduction

Computational chemists are focusing more on extremely large, complicated systems. With this move, Quantum Mechanics / Molecular Mechanics (QM/MM) molecular dynamics (MD) simulations are rising to the forefront of the field. In this technique, first proposed in 1976 by Warshel and Levitt, the interatomic forces in a small, important part of a system are derived using quantum mechanics (QM), typically a flavor of density functional theory (DFT), while the interatomic forces in the rest of the system are derived using classical molecular mechanics (MM), typically a force field. Put simply, QM/MM splits a system into a **QM region (QMR)** and an **MM region (MMR)**. All the interatomic interactions *within* and *between* these two regions are combined and integrated to simulate dynamics.

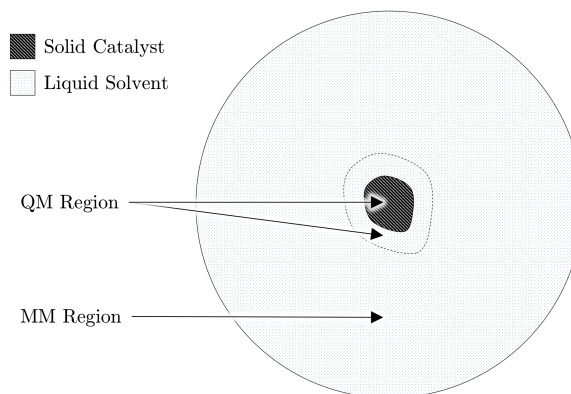


Figure 1: 2D representation of a heterogeneous catalyst model, in which a solid catalyst particle is surrounded by a thick sphere of explicit solvent. The QM region envelops the reaction zone, namely the catalyst and the solvent directly above the catalyst. The MM region envelops the bulk solvent.

QM methods are computationally demanding and scale poorly. For example, DFT is limited to picosecond MD and scales as $\mathcal{O}(N_e^3)$, where N_e is the number of electrons (Bowler et al., 2012). However, QM is necessary to describe the most important, least understood regions of most systems. On the other hand, MM methods are classical, and thus are fast and very scaleable. Force fields are regularly used to perform nanosecond MD and many packages achieve $\mathcal{O}(N)$ scaling, where N is the number of atoms. MM’s fast, approximate nature make it well suited for describing large, non-critical (yet influential) parts of a system.

QM/MM was originally proposed to study enzymes but can be applied equally effectively to man-made heterogeneous catalysts (Figure 1). Of particular importance are aqueous electrocatalysts, which drive water-splitting technologies. The high costs and low activities of these catalysts limit water split-

ting and therefore present a significant barrier to the widespread adoption of hydrogen fuel. To surmount the problems related to cost and activity, computational chemists are working to better understand the atomistic mechanisms of these catalysts. It is hoped that an understanding of the catalytic mechanisms will lead to the next generation of water-splitting catalysts.

Most studies on water-splitting catalysts are done using implicit solvent models. These mathematical descriptions, although effective at reducing system size to the point at which full QM treatment is practical, are too coarse. Specifically, implicit solvent models miss mechanistic steps that are mediated by multiple solvent molecules. For example, proton transfer to catalyst surfaces can be mediated by passing a proton from the bulk solution across multiple water molecules (Cheng et al., 2015). Applied effectively, QM/MM can simulate the solvent explicitly and thus can capture such a mechanism.

Although excellent methods exist for modeling QM and MM regions separately, **describing the interaction between the MMR and QMR remains a fundamental challenge**. In an ideal description, the interatomic forces between the MMR and QMR, as well as the polarization of the QM wavefunction by the MMR, would be captured. Two standard “embedding” methodologies have emerged that achieve this:

- **Mechanical Embedding** Interatomic forces between the QMR and MMR are described using a classical force field (Lin et al., 2006). This technique does not capture polarization and requires a force field for all atoms in the system. If any sort of reaction occurs across the QMR/MMR interface, that force field must be reactive. Developing such a reactive force field, even for a small set of elements and molecules, is prohibitively difficult.
- **Electrostatic Embedding** In the QM calculation, point charges are inserted into the system that simulate the electrostatic structure of the MM region (Lin et al., 2006). These point charges polarize the calculated wavefunction. However, performing electrostatic embedding requires extensive modification of QM packages, slows down QM calculations, and reduces the overall scalability of QM/MM. The methodology itself is also highly dependent on the charge equilibration model used in the MMR.

The issues with mechanical and electrostatic embedding techniques described above motivate our de-

velopment of a new embedding methodology – boundary region embedding – which we believe lies at a happy medium between the two conventional embedding approaches. Furthermore, we believe that boundary region embedding QM/MM (BREQM) MD simulations will prove highly useful when studying aqueous electrocatalysts.

Methodology

In BREQM, we introduce two additional regions, which we hereby refer to as **boundaries**. These boundaries separate the QMR and MMR. We dub the region in contact with the MMR and QMR the **MM boundary (MMB)** and the **QM boundary (QMB)**, respectively. The QMR and QMB make up the **QM part** of the system, while the MMR and MMB make up the **MM part**.

The interatomic interactions considered by BREQM, compared to those considered in mechanical embedding QM/MM, are diagrammed in Figure 2:

- Interactions *within* the QM and MM parts of the system remain identical to the mechanical embedding approach.
- The MM part of the system interacts only with the QMB, not the entire QM part; these interactions are described at the QM level, and only effect the positions and velocities of atoms in the QM part.
- Similarly, the QM part of the system interacts only with the MMB, not the entire MM part; these interactions are described at the MM level, and only effect the positions and velocities of atoms in the MM part.

Within the context of a heterogeneous catalyst model, the QMR envelops the catalyst itself. The QMB envelops the first monolayer of solvent above the catalyst, while the MMB envelops the second. The MMR envelops the rest of the solvent.

Discussion

Our BREQM methodology affords several benefits over traditional mechanical embedding QM/MM. First, the forces in the QM part of the system are derived solely at the QM level. This is not the case in mechanical embedding, which mixes MM- and QM-derived forces. Thus, BREQM cleanly separates the methods. In addition, by separating the QMR and

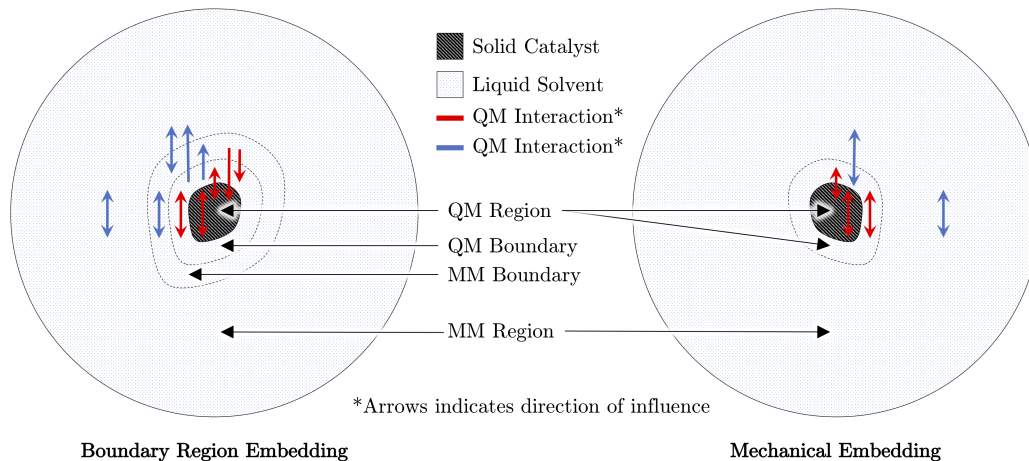


Figure 2: Interaction diagram for boundary region embedding (left) versus mechanical embedding (right) in a heterogeneous catalyst model.

MMR with boundary regions, BREQM eliminates the need to develop all-atom, reactive force fields for all elements in the system. Instead, only a reactive force field for the solvent is required. Thus, BREQM removes one of the most prohibitive aspects of applying QM/MM to heterogeneous catalysts.

BREQM also allows atoms in the MM part to polarize the QM part’s wavefunction while avoiding the challenges of conventional electrostatic embedding. Although this polarization is only partial, since only the MMB is included in the QM calculation, BREQM can theoretically utilize any QM software and method. Since no complicated software modification is required, boundary region embedding is more user-friendly and flexible than electrostatic embedding.

Despite these clear benefits, BREQM remains imperfect. As a consequence of cleanly separating the QM and MM methods, BREQM introduces a force discontinuity between the QMB and MMB. Since the QM- and MM-derived forces across this boundary are almost certainly unequal, the method is in violation of Newton’s third law! Although this fact poses a theoretical problem, it has little consequence in practice. The forces across the region are fairly close and conservation of energy is maintained by thermostating anyway (see Implementation). In the future, we plan on mixing the QM and MM-derived forces across the boundary in order to resolve the force discontinuity (Mones et al., 2015).

In its current form, BREQM ignores the very long-range forces between the MMR and QMR. This approximation is reasonable, since these regions are separated by several monolayers of solvent. However,

this approximation is not actually necessary; very long-range forces can be calculated easily and accurately using an all-atom, non-reactive force field. We plan on including such a force field in the next iteration. With this force field, BREQM will capture all interactions considered by mechanical embedding.

All in all, BREQM effectively describes the interatomic forces and polarization between the QM and MM parts of a system. Our methodology is attractively simple, and we believe that it will prove useful when studying aqueous heterogeneous catalysts. We hope that accurate and comprehensive BREQM studies on these catalysts will pave the way towards cheaper water-splitting and hydrogen fuel.

Implementation

Software

We implemented BREQM in Python. Our code interfaced with the Vienna Ab initio Simulation Package (VASP) (Kress et al., 1993; Kress et al., 1994; Kress et al., 1996; Kress et al., 1996), which we used to perform QM calculations (DFT). The code also interfaced with the LAMMPS molecular dynamics simulator (Plimpton, 1995), within which we used the ReaxFF reactive force field for water (Van Duin et al., 2001). The code interfaced indirectly with these packages by writing relevant input files and sending jobs to a Linux cluster. This procedure could be improved by directly linking the code into these software packages.

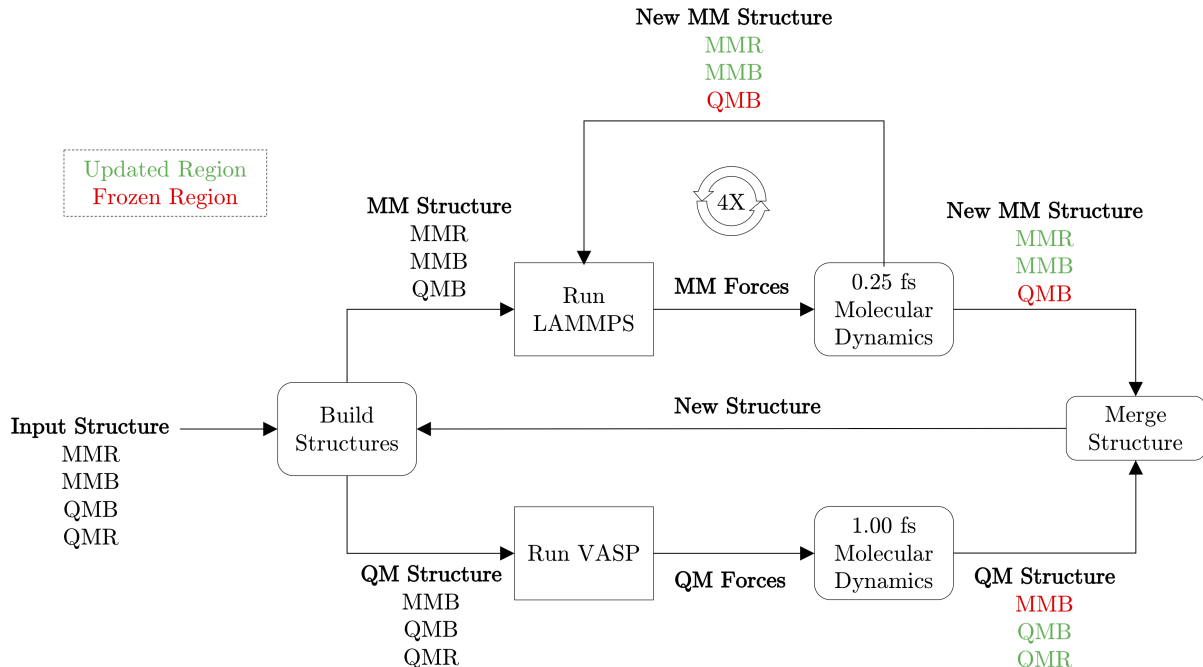


Figure 3: Implemented QM/MM MD scheme using a MM timestep 4x smaller than the QM timestep.

User Input

The user provides an input structure in VASP’s POSCAR format. As of now, the code assumes a 2D periodic slab geometry, in which a periodic slab of catalyst is covered in many layers of solvent. The user species BREQM’s regions and boundaries through z-coordinate cutoff parameters. The user also provides computational parameters for VASP and LAMMPS, so the user has total control over the QM and MM calculations.

Force Calculation

As diagrammed in Figure 3, the code first splits the input structure into QM and MM parts using the provided z-coordinate cutoffs. These structures are sent to VASP and LAMMPS, respectively, which compute various forces. The code then extracts the relevant forces from each software package’s output to capture the BREQM interactions. In the QM calculation, the MM region is included but remains unchanged. In the MM calculation, the QM region is included but remains unchanged.

Molecular Dynamics

The code implements a velocity verlet algorithm to simulate an NVE ensemble (fixed energy) and a Nosé-Hoover algorithm to simulate an NVT ensemble

(fixed temperature). In the case of velocity verlet, the user can periodically apply velocity rescaling according to the Maxwell-Boltzmann distribution. Thus, when using velocity verlet, the user can linearly heat and cool the system by occasionally rescaling the atom velocities. Alternatively, the user can maintain the temperature through rescaling; however, the Nosé-Hoover algorithm is recommended for simulating constant temperature.

Molecular Dynamics Timestep

Since MM-derived forces are less accurate than QM-derived forces, MM MD requires smaller timesteps than QM MD. For example, ReaxFF force field MD works best with a 0.25 fs timestep while DFT MD can easily tolerate a timestep as large as 1.00 fs.

To account for different timesteps, we developed a simple QM/MM MD scheme diagrammed in figure 3. Rather than perform each MD step on the entire system, we separately perform MD steps on the QM and MM parts of the system and then merge the structures together. The QM and MM parts of the system are technically out of phase by a single QM timestep. However, since simulations are performed over hundreds of timesteps, this is inconsequential.

Our implementation allows the timesteps to be chosen by the user. This allows for different dynamics durations for the MM and QM parts. For instance,

one might wish to run 100 fs of dynamics on the MM part for every timestep on the QM part. In doing this, one can ensure that the MM part is *always* in equilibrium with the QM part.

Bond Breaking

In QM/MM simulations, bonds can be broken when splitting the system into QM and MM structures. Within the BREQM methodology, bonds can break between the QMR and QMB and between the MMR and MMB (consider the structure building process in Figure 3). Bond breaking can have disastrous consequences, since it leads to incorrect QM-computed wavefunctions and MM-computed atomic charges, both of which are central when calculating per-atom forces.

We developed BREQM with heterogeneous catalysis in mind, so we were luckily able to side-step the problem of bond breaking. We did this by ensuring that the QMR/QMB and MMR/MMB interfaces are contained entirely within the solvent. Since the solvent is composed of small, individual molecules, we use automatically calculated bond information to ensure that solvent molecules are never broken. In essence, our code dynamically redraws the boundaries around solvent molecules.

References

- Asthana, A., and Wheeler, D.R. (2013). A polarizable reactive force field for water to enable molecular dynamics simulations of proton transport. *J Chem Phys* *138*, 174502.
- Bowler, D.R., and Miyazaki, T. (2012). O(N) methods in electronic structure calculations. *Rep Prog Phys* *75*, 036503.
- Cheng, T., Xiao, H., and Goddard, W.A., 3rd (2015). Free-Energy Barriers and Reaction Mechanisms for the Electrochemical Reduction of CO on the Cu(100) Surface, Including Multiple Layers of Explicit Solvent at pH 0. *J Phys Chem Lett* *6*, 4767-4773.
- Kresse, G., and Furthmuller, J. (1996). Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* *6*, 15-50.
- Kresse, G., and Furthmuller, J. (1996). Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B* *54*, 11169-11186.
- Kresse, G., and Hafner, J. (1993). Ab initio molecular dynamics for liquid metals. *Phys Rev B Condens Matter* *47*, 558-561.
- Kresse, G., and Hafner, J. (1994). Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Phys Rev B Condens Matter* *49*, 14251-14269.
- Lin, H., and Truhlar, D.G. (2006). QM/MM: what have we learned, where are we, and where do we go from here? *Theoretical Chemistry Accounts* *117*, 185-199.
- Mones, L., Jones, A., Gotz, A.W., Laino, T., Walker, R.C., Leimkuhler, B., Csanyi, G., and Bernstein, N. (2015). The adaptive buffered force QM/MM method in the CP2K and AMBER software packages. *J Comput Chem* *36*, 633-648.
- Plimpton, S. (1995). Fast Parallel Algorithms for Short-Range Molecular-Dynamics. *Journal of Computational Physics* *117*, 1-19.
- Van Duin, A.C., Dasgupta, S., Lorant, F., and Goddard, W.A. (2001). ReaxFF: a reactive force field for hydrocarbons. *J Phys Chem A* *105*, 9396-9409.
- Warshel, A., and Levitt, M. (1976). Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* *103*, 227-249.