Combining Locally Optimized Interpolants with a Multilevel QM/MM Method for Complex Reactive Simulations

Michael R. Salazar
Department of Chemisty
Union University

Abstract

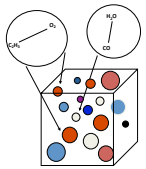
A method for performing QM/MM molecular dynamic simulations of complex reactive systems is presented. The **A**ccelerated **Mol**ecular **D**ynamics with **C**hemistry (AMolDC) program formulates the total potential of the system into time-dependent (adaptive) spatially-resolved groups over which multiple levels of QM theory may be employed in combination with an in-memory potential energy surface database (PESDBase) that is called by computationally facile local optimized interpolants that are capable of adapting to the density of the underlying grid and topology of the multidimensional potential and gradient surfaces of the variously formed groups. Accurate high-level QM calculations are utilized for reactive groups, while lower level QM or MM calculations are performed for the ubiquitous non-reactive collisions. AMolDC is linked to the NWChem, PSI3, and GAMESS electronic structure codes and the ReaxFF is used for molecular mechanical calculations. The combination of saving the QM potential energy surface data for the formulation of interpolants in subsequent calculations is employed for the expansion of system size and simulation time of complex chemical reactive processes where high levels of QM theory are required.

The energy conservation and computational properties of the method (without utilizing the interpolation component) have been investigated and reported (*J. Chem. Theory Comput.* **2010**, *6*, 18-25). These investigations show that, although energy is not a conserved property, smooth continuous (and time-reversible) simulations within the canonical ensemble (NVT) are feasible despite very large discontinuities in the total potential. The method also demonstrates linear computational scaling ($O(N)$) rather than the higher orders of QM theory ($O(N^3)$-$O(N^7)$).

In order to expand the system size and simulation time of the targeted complex reactive processes, AMolDC makes calls to extremely fast, numerical interpolants once the database is sufficiently populated for a given group. These interpolants are local (they use a subset of all the stored QM data for an individual group) and they are optimized (they depend on parameters that may be optimized to adapt to the density of the underlying grid and potential or gradient topology). The accuracy of the gradient interpolation has been examined and are reported for very large and diverse chemical systems (saturated and unsaturated hydrocarbon chains and rings) as a function of the underlying grid density. In addition, the computational savings associated with reducing the number of high-level QM calculations and employing local interpolants is reported.

Adaptive Multilevel QM/MM Theory

$$V(t) = \sum_{i}^{NGroups} \sum_{\alpha}^{N_i} V_{i\alpha} + \frac{1}{2} \sum_{i}^{NGroups} \sum_{\alpha}^{N_i} \sum_{j}^{NGroups} \sum_{\beta}^{N_j} V_{i\alpha,j\beta}$$



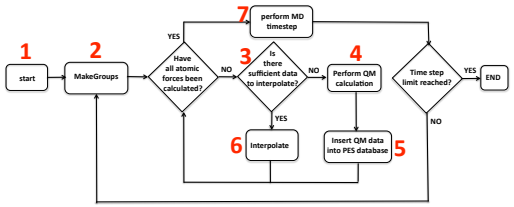
Potential/Gradients

Intragroup: Multilevel QM (QM_1,QM_2,QM_3,…)

Intergroup (MM)

Save all QM data (potential and Cartesian gradients) to database.
Linked to NWChem, GAMESS, PSI3 electronic structure codes, MM = ReaxFF
Numerical interpolation code accesses database to check grid density in order to perform interpolation

Accelerated Molecular Dynamics with Chemistry

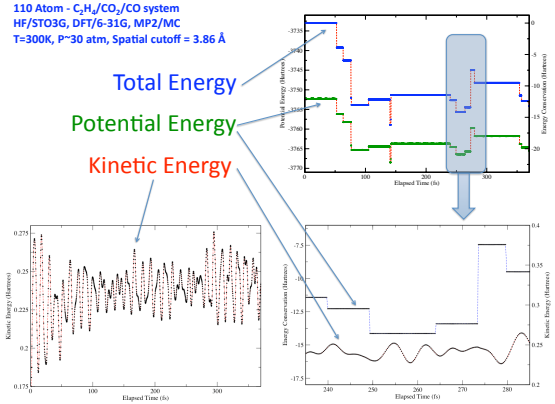


1: Initialize – coordinates, temperature, Nsteps, …
2: *MakeGroups*– divide simulation cell into link-listed subcells and form groups based on spatial cutoff
3: Search grid to get density of QM data
4: Perform NWChem, GAMESS, or PSI3 QM calculation
5: *PESDatabase*– insert QM data into database
6: *Interpolate* – perform numerical interpolation on the basis of underlying QM grid
7: MD step using Rapaport MD code

Continuity Properties

110 Atom - $C_2H_4/CO_2/CO$ system
HF/STO3G, DFT/6-31G, MP2/MC
T=300K, P~30 atm, Spatial cutoff = 3.86 Å

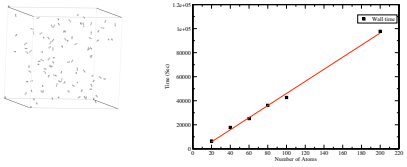
Total Energy
Potential Energy
Kinetic Energy



Computational Properties

Serial

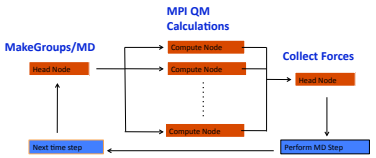
H_2 simulation
400 atm., 1000K
HF/6-31G**/ReaxFF
1000 time steps

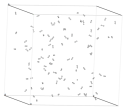


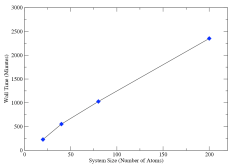
Parallel

H_2 (20 – 200 atoms)
HF/cc-pVTZ/ReaxFF
400 atm, 1000 K
1000 time steps
(Step = 0.25 fs)

7 Node Cluster –
1 Head Node
6 Compute Nodes



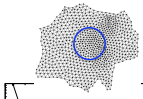




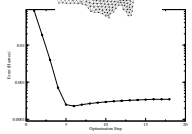
Interpolation Methodology

$$V(P) = \sum_{i}^{NClosest} c_i \sqrt{(r_{1,P} - r_{1,i})^2 + (r_{2,P} - r_{2,i})^2 + (r_{3,P} - r_{3,i})^2} + \ldots + \Delta$$

Local – use NClosest grid points



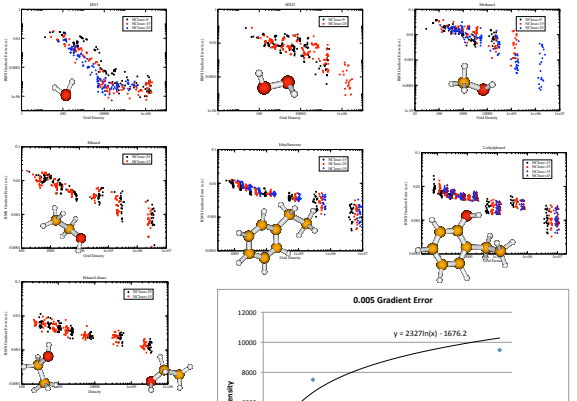
Optimized – choose Δ to minimize first difference in interpolated gradients at point P



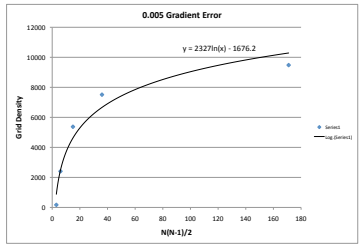
Interpolation Properties

Run MD simulation using ReaxFF, save PES data, calculate grid density, interpolate, and evaluate error.

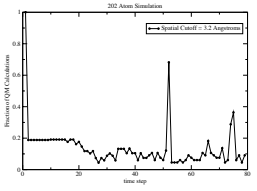
$$\rho = \frac{NClosest}{r_{rms}^3}, r_{rms} = \sqrt{\frac{1}{NClosest * NDim} \sum_{i}^{NClosest} \sum_{j>1}^{NAtoms} \sum^{NAtoms} (r_{P,ij} - r_{1,ij})^2}$$



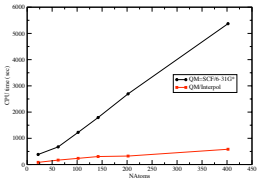
Grid density required to maintain 0.005 au error in gradients



Database population rate



Computational scaling employing interpolation



Acknowledgements

Acknowledgment is made to the donors of the American Chemical Society Petroleum Research Fund for support of this research. I thank Adrivan Duin for assistance with the ReaxFFcode and Wesley Allen and TrygveHelgaker for helpful conversations with issues related to the gradient conversions.