Description of the QM/MM method

Basic overview: QM/MM is a method that allows systems with locally chemically active regions to be treated at an *ab initio* level while the remainder of the system (referred to as the bath) is treated using a simplified model or an empirical "force field". An example would be an enzyme in water in which chemical processes at the active site are treated at the *ab initio* level while the majority of the water and the protein is modeled using CHARMM.

The basic components of a QM/MM calculation are:

1. A description of the electronic structure or *ab initio* calculation (a basis set and some level of theory).

2. An empirical force field to treat the bath.

3. A model for how the *ab initio* and bath interact with each other.

4. Two "boxes" are defined. The former refers to the full simulation cell encompassing the entire system and is called the Large-Box. The latter refers to the box which encloses the *ab initio* or chemically active region is called the CP-Box.

5. Specification of boundary conditions, e.g. periodic boundary conditions, on the Large-Box.

In order to perform QM/MM calculations, the following data is needed:

1. A set of electronic orbitals which, ultimately, are described in terms of a set of coefficients for expansion in a basis set. In PINY_MD, this would be a plane wave basis set. For Car-Parrinello calculations, we also need a set of fictitious velocities and forces for the coefficients (defined in the CP-box).

2. A set of positions, velocities, and forces for atoms treated at the *ab initio* level.

3. A set of positions, velocities, and forces for atoms treated using the force field.

The following is a description of the basic steps needed in a QM/MM calculation assuming that a plane wave basis set is used for the electronic orbitals and that the electronic structure is treated using density functional theory (This is the scheme that is currently implemented in PINY_MD).

Unlike plane-wave based *ab initio* MD, which uses two reciprocal space grids, one for the orbitals and one for the density, and one real-space grid, QM/MM calculations use three reciprocal space grids and two real space grids. The two real-space grids, referred to as the Large-Sparse and Small-Dense real space grids, respectively, correspond to the discretization of the Large-Box and CP-box, respectively. The three reciprocal space grids are as follows: There are the two reciprocal space grids used in standard CP calculations for the CP box, i.e. the reciprocal space grid for the coefficients and the density, respectively. The third reciprocal space grid called the Large-lowcutoff-reciprocal space is defined in the Large-Box. The Large-Sparse Grid and Large-density-Lowcutoff reciprocal space do not contain many elements (See Flow Chart).

In addition, the short range electron-atom interactions are calculated on the Small-highcutoff reciprocal space (inside the CP-Box) and the long range interactions on the Large-Lowcutoff-reciprocal space (inside the Large-Box).

1. Given N orbitals Small-orbital-highcutoff reciprocal space (within the CP-Box) a set of N FFTs is performed in order to generate the orbitals in real space and use them to calculate the density on the Small-Dense real-space grid (within the CP-Box).

2. The density on on the Small-Dense real-space grid is then interpolated in real space to the Large-Sparse real-space grid using Cardinal B-splines.

3. Perform an FFT of the density on both the Small-Dense real-space grid and the Large-Sparse real-space grid to obtain reciprocal-space densities on the Small-density-highcutoff reciprocal space and the large-density-lowcutoff reciprocal.

4. The calculation now branches as in ordinary CP. The Small-dense real-space density is used to calculate the exchange and correlation potential, while the two reciprocal-space densities are used to calculate the long and short range parts of the Hartree and local pseudopotentials. The long range part includes the interaction of the electrons with the atoms in the force-field region of the system. The Kohn-Sham potential is collected.

5. The calculations are then brought back together as in standard CP, and the real-space and reciprocal-space energies are combined. The Small-highcutoff reciprocal space potential is Fourier transformed back to small-dense real-space grid where it is combined with the exchange and correlation contribution. The Large-lowcutoff reciprocal space potential is transformed back to the Large-sparse real-space and B-Spline interpolated back to the Small-dense real-space grid and added in. The total is then multiplied on the orbitals in real-space and the result Fourier transformed back to reciprocal space to generate the final electronic forces.

6. The calculation of the intra- and intermolecular forces on the atoms described by the force field is an independent calculation that can occur before, after, or (preferably) during the calculation of electronic forces. The load balancing will depend sensitively on the size of this part of the system.