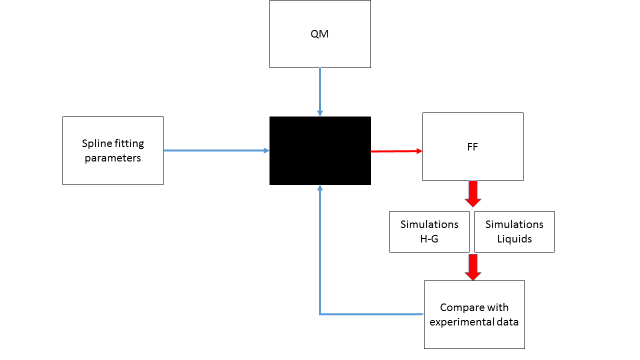
Force field mapping

# August 24, 2016

## Meeting notes

* Chris Bayly has a way to automatically find atom types based on SMILES and something similar to SMILES called SMIRKS. He can compare his automatic atom types with atom types in a FF he developed while at Merck (parm@Frosst <http://www.ccl.net/cca/data/parm_at_Frosst/parm_Frosst_note.pdf>) and check for convergence.
* GEM\*(Duke, Starovoytov et al. 2014) is a similar attempt to generate a FF from quantum data.
* (McDaniel and Schmidt 2016) have a next-generation FF review paper, which seems to focus on using SAPT.
* [Regularization](https://en.wikipedia.org/wiki/Regularization_(mathematics)) is a process to help avoid overfitting with parameters. Be cognizant.
* How are we different than Force Balance (Wang, Chen et al. 2013, Wang, Martinez et al. 2014)?
* Pure liquid properties we can match (start with the bold):
  + Diffusion
  + **Density**
  + **Enthalpy of vaporization**
  + Dielectric constant
  + Heat capacity
  + Enthalpy of mixing
* The [COBYLA](https://en.wikipedia.org/wiki/COBYLA) optimizer is in [scipy](http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.optimize.fmin_cobyla.html) and allows constraints on the parameters and does not require the gradient.
* The “[bag of bonds](http://pubs.acs.org/doi/pdf/10.1021/acs.jpclett.5b00831)” model (Hansen, Biegler et al. 2015) demonstrates a machine learning approach.



## To-do

1. Write a function to extract relevant properties from simulations.
2. Write a function to output proper parameter and topology files.
   1. Use parmed in python (python2 only).
   2. Use parmed system call.
   3. Write custom GAFF file and reprocess with tleap.
3. Write a function to compare computed and experimental properties.
4. Write a function that – for now – loops over values and reports the difference between the chosen parameters and the experimental values.
5. Start with ethanol.

# September 9, 2016

1. Link up with Niel’s QM data and try to optimize a mapping function for ethanol @oh and the pure density.
2. Figure out an effective way to grep for simulation errors. This is started in a new GH branch called (I think) `check\_simulation`.
3. Compare results with Sam’s density re-weighting.

References

Duke, R. E., et al. (2014). "GEM\*: A Molecular Electronic Density-Based Force Field for Molecular Dynamics Simulations." J Chem Theory Comput **10**(4): 1361-1365.

GEM\*, a force field that combines Coulomb and Exchange terms calculated with Hermite Gaussians with the polarization, bonded, and modified van der Waals terms from AMOEBA is presented. GEM\* is tested on an initial water model fitted at the same level as AMOEBA. The integrals required for the evaluation of the intermolecular Coulomb interactions are efficiently evaluated by means of reciprocal space methods. The GEM\* water model is tested by comparing energies and forces for a series of water oligomers and MD simulations. Timings for GEM\* compared to AMOEBA are presented and discussed.

Hansen, K., et al. (2015). "Machine Learning Predictions of Molecular Properties: Accurate Many-Body Potentials and Nonlocality in Chemical Space." J Phys Chem Lett **6**(12): 2326-2331.

Simultaneously accurate and efficient prediction of molecular properties throughout chemical compound space is a critical ingredient toward rational compound design in chemical and pharmaceutical industries. Aiming toward this goal, we develop and apply a systematic hierarchy of efficient empirical methods to estimate atomization and total energies of molecules. These methods range from a simple sum over atoms, to addition of bond energies, to pairwise interatomic force fields, reaching to the more sophisticated machine learning approaches that are capable of describing collective interactions between many atoms or bonds. In the case of equilibrium molecular geometries, even simple pairwise force fields demonstrate prediction accuracy comparable to benchmark energies calculated using density functional theory with hybrid exchange-correlation functionals; however, accounting for the collective many-body interactions proves to be essential for approaching the "holy grail" of chemical accuracy of 1 kcal/mol for both equilibrium and out-of-equilibrium geometries. This remarkable accuracy is achieved by a vectorized representation of molecules (so-called Bag of Bonds model) that exhibits strong nonlocality in chemical space. In addition, the same representation allows us to predict accurate electronic properties of molecules, such as their polarizability and molecular frontier orbital energies.

McDaniel, J. G. and J. R. Schmidt (2016). "Next-Generation Force Fields from Symmetry-Adapted Perturbation Theory." Annu Rev Phys Chem **67**: 467-488.

Symmetry-adapted perturbation theory (SAPT) provides a unique set of advantages for parameterizing next-generation force fields from first principles. SAPT provides a direct, basis-set superposition error free estimate of molecular interaction energies, a physically intuitive energy decomposition, and a seamless transition to an asymptotic picture of intermolecular interactions. These properties have been exploited throughout the literature to develop next-generation force fields for a variety of applications, including classical molecular dynamics simulations, crystal structure prediction, and quantum dynamics/spectroscopy. This review provides a brief overview of the formalism and theory of SAPT, along with a practical discussion of the various methodologies utilized to parameterize force fields from SAPT calculations. It also highlights a number of applications of SAPT-based force fields for chemical systems of particular interest. Finally, the review ends with a brief outlook on the future opportunities and challenges that remain for next-generation force fields based on SAPT.

Wang, L. P., et al. (2013). "Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data." J Chem Theory Comput **9**(1): 452-460.

We introduce ForceBalance, a method and free software package for systematic force field optimization with the ability to parametrize a wide variety of functional forms using flexible combinations of reference data. We outline several important challenges in force field development and how they are addressed in ForceBalance, and present an example calculation where these methods are applied to develop a highly accurate polarizable water model. ForceBalance is available for free download at <https://simtk.org/home/forcebalance>.

Wang, L. P., et al. (2014). "Building Force Fields: An Automatic, Systematic, and Reproducible Approach." J Phys Chem Lett **5**(11): 1885-1891.

The development of accurate molecular mechanics force fields is a significant challenge that must be addressed for the continued success of molecular simulation. We developed the ForceBalance method to automatically derive accurate force field parameters using flexible combinations of experimental and theoretical reference data. The method is demonstrated in the parametrization of two rigid water models, yielding new parameter sets (TIP3P-FB and TIP4P-FB) that accurately describe many physical properties of water.