

NSF Research Proposal

Bill March

Computer simulations are widely used in biology and chemistry. *Add a better opening sentence. Discuss some compelling applications here.*

Determining Parameters. All of these simulations require a tradeoff between detail and scale. Detailed quantum mechanics methods can require quadratic space and $O(N^4)$ time. On the other hand, simplified models can easily lose the details necessary to capture interesting behavior. Therefore, one of the primary challenges of developing effective computer simulations is to find the correct simplifications. In quantum mechanical methods, sets of functions, called basis sets, are combined to approximate the wavefunction. Basis sets are carefully optimized for both speed and accuracy and often are specialized for particular systems. In classical simulations on proteins, a set of parameters are fixed, called a force field. These included atomic charges which are developed to accurately reproduce experiments. Determining a force field is extremely time consuming, and force fields rarely generalize to other systems. New algorithms are necessary to allow larger and more accurate simulations and reduce the need for scientists to spend time developing elaborate parameter sets.

Many-Body Interactions. One of the key (and largely unexamined) simplifying assumptions in nearly all methods is that all interactions between particles can be computed through sums of interactions between pairs. Since computing the interaction between all k -tuples scales as $O(N^k)$, higher-order interactions are generally considered intractable. While this assumption is reasonable for many systems, it fails for some. Three-body terms in liquid argon account for about 10% of the total energy (cite). Such terms are also believed to be important in simulations of water [5]. These terms have received little attention, not because they are physically insignificant, but because of the difficulties with computing them.

Efficient Many-Body Potentials. We propose a general framework for making many-body potentials computationally feasible for large systems. (This paragraph needs to say more, but I'm not sure what). In particular, we will test our framework on the well-known Axilrod-Teller potential for dispersion among triples of atoms [1]. This will allow us to demonstrate the applicability of our method on a well-known instance and will hopefully promote more investigation of many-body potentials.

Fast Two-Body Computations. In general, two-body potentials scale as $O(N^2)$. For many potentials, it is possible to truncate such computations for distant particles. However, for long-range effects, such as the Coulomb electrostatic potential, this simple method is too inaccurate, so we must find a better approximation for distant interactions.

FMM. The Fast Multipole Method [3] makes large-scale electrostatic computations feasible. The FMM groups charged particles in the nodes of a space-partitioning tree, and computes interactions between pairs of particles by considering pairs of nodes. Interactions between nearby groups of particles can be computed exhaustively, as in the truncation method. When two nodes

are distant, we approximate their interactions with a single computation. We approximate the potential for each node with a Taylor series. This *multipole expansion* is what makes the FMM efficient.

N-body Problems in Statistics. The FMM is effective for the Coulomb potential between pairs of particles. We will apply ideas from statistics in order to extend the FMM to include more general potentials involving many particles. Gray and Moore [2] have suggested a framework for accelerating more general computations between pairs of points. This framework has been applied to computing the Fast Gauss Transform [4], which is a computation similar to many potentials. *Improve this!*

In order to compute all pairs interactions between two sets of points, each set is organized in a binary space partitioning tree, like in the FMM. We can consider the computation between two nodes, one from each tree. If the nodes are distant enough to make a valid approximation, we can *prune* further computation. Otherwise, we recurse on both nodes simultaneously, and consider the four pairs of children. This dual-tree recursion can easily be extended to computations involving more than pairs. Instead of employing two trees, we can use k , and recurse in each simultaneously. Pruning can occur as before.

Many-Body Potentials. We will apply this framework to computing the Axilrod-Teller three-body potential. We will develop multipole expansions for this potential, then build a multi-tree framework to handle the recursive approximations. We also intend to test this algorithm on systems of chemical interest in order to demonstrate its effectiveness.

Progress. *Do we have results for monopole version yet?*

This is a good place to talk about collaborators, why GT is a good place for this, etc.

These techniques (meaning series expansions and trees) can be extended to treat many other multibody potentials. We are also considering extensions of this framework to other problems in computational chemistry. As mentioned above, quantum mechanics simulations utilize basis functions. The rate-limiting step in the fundamental Hartree-Fock method is computing electron repulsion integrals (ERI's) among sets of four basis functions. *Give the formula here.* This computation is very similar to a four-body potential, and we believe that our methods can be extended to accelerate this and other methods.

References

- [1] B. M. Axilrod and E. Teller. Interaction of the van der waals type between three atoms. *J. Chem. Phys.*, 11(299), 1943.
- [2] A. Gray and A. W. Moore. N-body problems in statistical learning. In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems 13 (December 2000)*. MIT Press, 2001.
- [3] L. Greengard and V. Rokhlin. A Fast Algorithm for Particle Simulations. *Journal of Computational Physics*, 73:325–248, 1987.
- [4] D. Lee, A. Gray, and A. Moore. Dual-tree fast gauss transforms. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 747–754. MIT Press, Cambridge, MA, 2006.
- [5] T. P. Lybrand and P. A. Kollman. Water-water and water-ion potential functions including terms for many body effects. *Journal of Chemical Physics*, 83:2923–2933, 1985.