Tractable ERI's in Hartree-Fock

Keywords. *computational chemistry*

this is just the gist that i think makes a good story. feel free to use your own language, etc.

first paragraph - short summary: i am pursuing more efficient methods for quantum simulation, by introducing algorithmic ideas from computer science as well as mathematical ideas from nonparametric estimation theory. (this latter part is optional depending on space - see below).

studying and designing molecules for drug design and materials design requires simulation of their behavior. for computational reasons, simulations at performed at various levels of fidelity, the quantum level being the most detailed. there we account for all the electrons, etc. at higher levels, gross approximations are made, including things like approximating whole side-chains at the protein folding level. note that the inability to simulate tractably while maintaining accuracy is the bottleneck for the protein folding needed for drug discovery.

the hartree-fock method is what's used at the quantum level [cite]. even the most accurate methods like David Sherrill's coupled-cluster methods [cite] build upon hartree-fock. it has two main computational parts, an outer svd and an inner N^4 loop. (show one equation.) for this reason only tens of atoms can be simulated at the quantum level of simulation. we focus our efforts on computing the inner loop tractably while maintaining very high accuracy.

the FMM [cite] is used here. however, the FMM was created for pairwise interactions. it makes the naive N^2 into O(N). it has been used in the hartree-fock context by making other hacky approximations like truncations. despite the promise of the introduction of the FMM to quantum chemistry, it hasn't opened the way to larger molecules. the rigorous error bounds of the FMM do not carry over to the entire computation since the FMM only deals with part of the overall N^4 computation. there is no FMM which can deal with 4-way interactions.

[don't go into all the details about the FMM that you did in the current writeup - first, the exact statements aren't completely accurate; also it is not needed]

in [nbody-nips-paper, riegel-tech-report 07], a kind of generalization of the FMM was shown, which extends to k-way interactions. the approach comes from a computer science viewpoint, and can be thought of as a kind of higher-order divide-and-conquer. this line of work was concerned with statistical problems, such as n-point correlation functions (which involve a k-way interaction) [gray comp-phys 94] and kernel density estimation (which can involve gaussian densities) [lee-paper 06]. the recursive approach can use adaptive space-partitioning trees which are more sensitive than oct-trees to non-uniformly distributed points, and can automatically guarantee rigorous error bounds on relative error (with respect to the true quantity) as opposed to absolute error.

describe how it works. figure can be made more compact by shortening the lines between the nodes. also use wrapfig latex package (ask ryan) to make it take up much less space. perhaps show one equation, like the hermite expansion for the guassian.

i am investigating the derivation of a fast algorithm for the inner loop of hartree-fock using this framework. this involves.... mention investigating recent data structures from theoretical computer science such as cover-trees [cite]... implementing it in software... testing on molecules of chemical interest such as... replacing the Gauss software to help the quantum chemistry community since good public codes don't exist... describe how sherrill and his group are providing chemistry intuition, data, etc. to us.

perhaps also get into the idea of using the theory of nonparametric estimation to select the gaussian bandwidths. mention the 1/r approximation, etc.

this work has the exciting potential of allowing bigger molecules... connect back to materials and drug design... etc.

Originality. This proposal represents my own work and ideas.