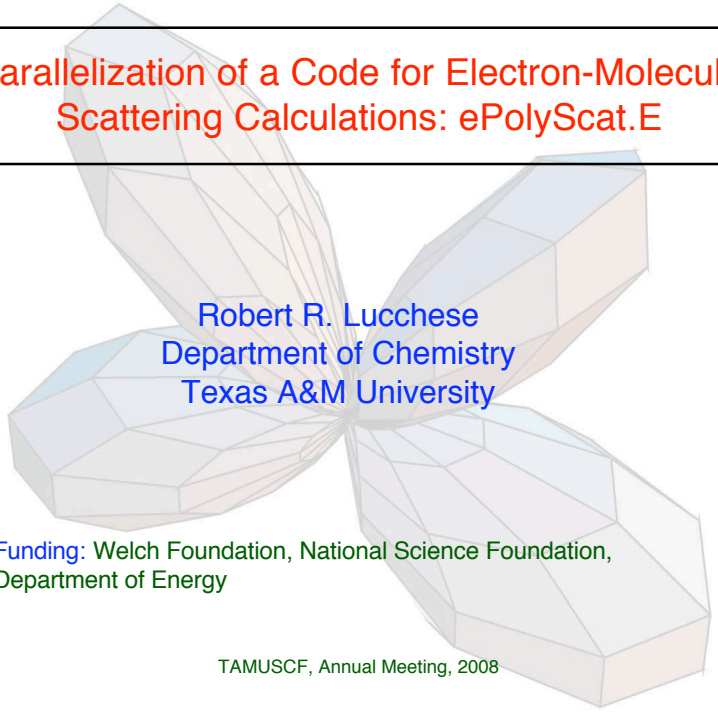
Parallelization of a Code for Electron-Molecule Scattering Calculations: ePolyScat.E



Robert R. Lucchese
Department of Chemistry
Texas A&M University

Funding: Welch Foundation, National Science Foundation, Department of Energy

Design Goals

- ePolyScat.D, a legacy serial code
  - Many separate executables
  - Coarse grained parallelization using shell scripts
  - Low memory usage
    - Heavy use of scratch disk files
    - Communication between executables through disk
  - Static array allocation
- ePolyScat.E, distributed-memory parallel ($N \leq 64$)
  - One executable
  - Limited disk I/O using only the master node
  - Dynamic array allocation
  - All intermediate data held in core
  - Fortran 90, MPI message passing

Basic Equations

The three dimensional Schrodinger equation for an electron interacting with a molecule

$$H\psi(r,\theta,\phi) = E\psi(r,\theta,\phi)$$

Use expansion in the angular coordinates

$$\psi(r,\theta,\phi) = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l} \frac{1}{r} f_{lm}(r) X_{lm}(\theta,\phi)$$

Differential equation for radial functions

$$\left\{ -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - E \right\} f_{lm}^{n}(r) + \sum_{l'm'} V_{lm,l'm'}^{\text{static}}(r) f_{l'm'}^{n}(r)$$
$$= \sum_{l'm'} \int_0^{\infty} V_{lm,l'm'}^{\text{exchange}}(r,r') f_{l'm'}^{n-1}(r') dr'$$

Distribute Grid on Processors

The problem is solved on a grid

$$f_{lm}(r) \Rightarrow f_{lmk} = f_{lm}(r_k) \quad k = 1,\text{K}\ ,M$$

Each of the *N* processor then has part of the grid

$$[1,\text{K}\ ,M_1]\ [M_1+1,\text{K}\ M_2]\ \text{L}\ [M_{N-1}+1,\text{K}\ M_N]$$

- The grid has different step sizes in different regions of the molecule
- Different values of $l_{\max}$ are used in different regions
- Different numbers of points are allocated to each processors

Exchange Integrals

The exchange integral contains a term from the Coulomb potential of the form

$$I_l(r) = \frac{1}{r^{l+1}} \int_0^r (r')^l \, g(r') dr' + r^l \int_r^\infty \frac{1}{(r')^{l+1}} g(r') dr'$$

Partial integrations are done independently on each processor and the sums are obtained in VDiff using MPI_SCAN

$$I^k = \sum_{j=1}^{k} P^j$$

Solution of the Inhomogeneous Diff Eq

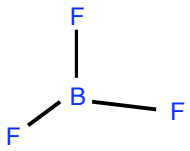
- The solution of the differential equation can also be written as a series of matrix multiplications

$$\mathbf{f}_N = \mathbf{T}_N \mathbf{T}_{N-1} \cdots \mathbf{T}_2 \mathbf{T}_1 \mathbf{f}_0$$

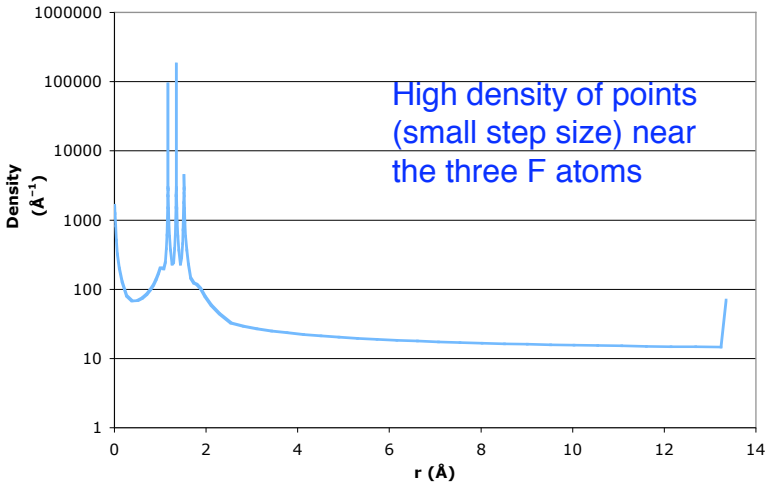
  where each $\mathbf{T}_k$ matrix resides on a different processor.
- Unfortunately, in order to obtain numerical stability, the computation of $\mathbf{T}_k$ depends in part on the value of the vector obtained on the *k*-1 processor.
- Current code in GHomo uses a pipeline for parallelization.

Test Runs on Hydra

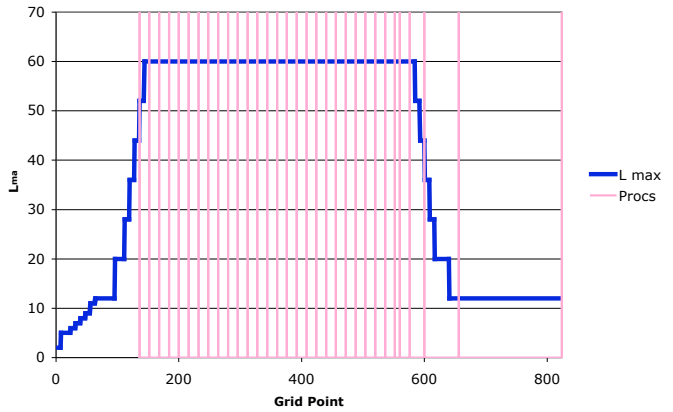
- Tests using 8, 16, 32, and 64 processors with 8 processors per node
- Elapsed time ranged from 3 hours for 8 processors to 1 hour with 64
- Production runs typically use 32 processors
- Test jobs compute the photoionization of BF_3 in an asymmetric planar geometry



Radial Grid Density

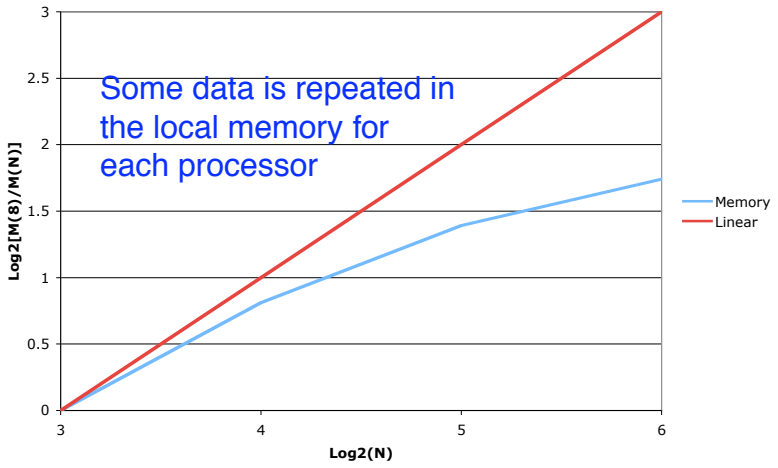


Partial Wave Expansion

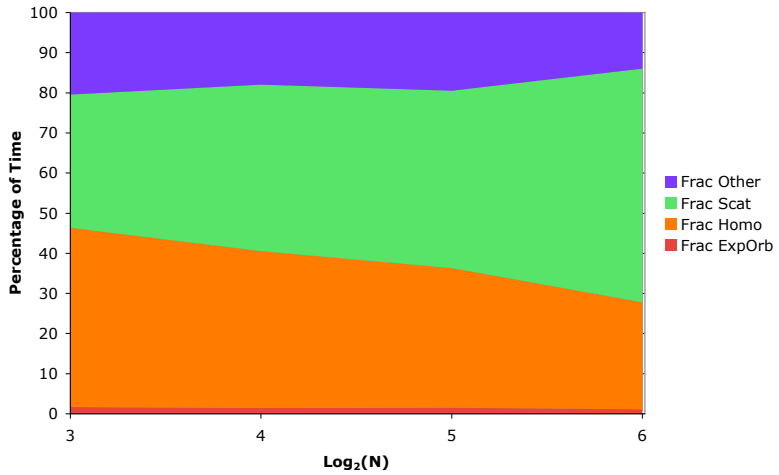


Distribution of grid on 32 processors, work ~ $L^{1.5}$

Memory Usage

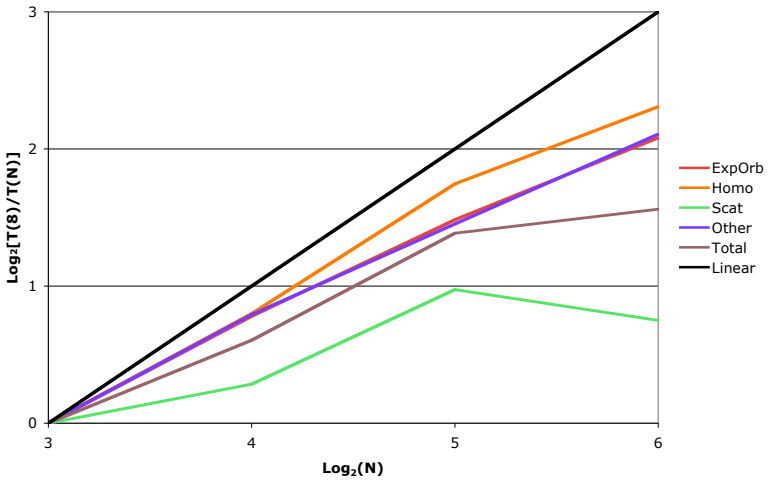


Distribution of Time in ePolyScat



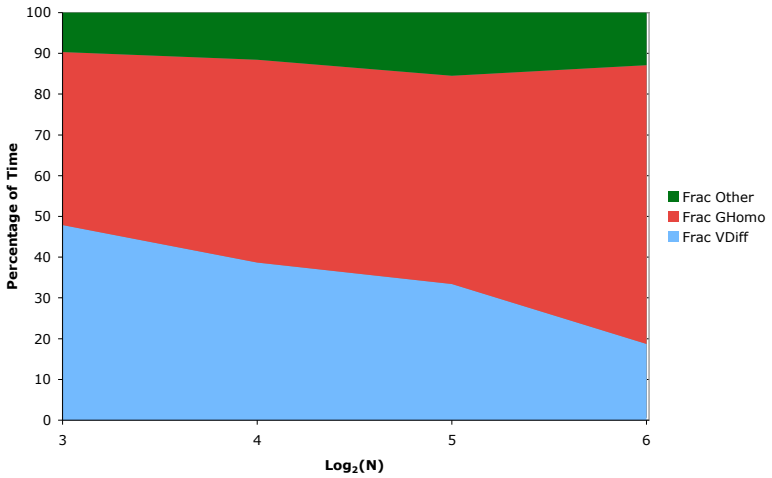
TAMUSCF, Annual Meeting, 2008

Speedup in ePolyScat



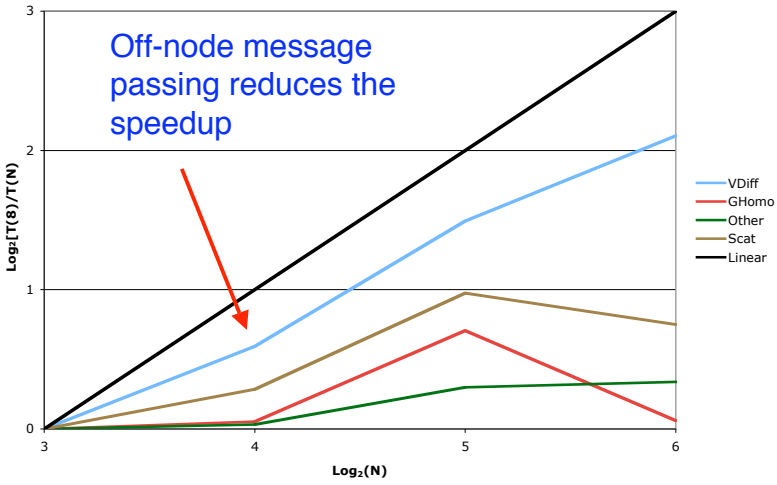
TAMUSCF, Annual Meeting, 2008

Distribution of Time in Scat



TAMUSCF, Annual Meeting, 2008

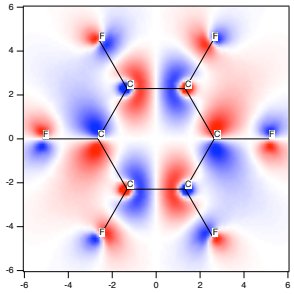
Speedup in Scat



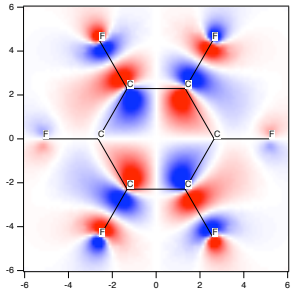
TAMUSCF, Annual Meeting, 2008

Photoionization of C_6F_6 Leading to the $C\,^2B_{2u}$ State of $C_6F_6^+$

$b_{2u} \rightarrow ke_{2g}$



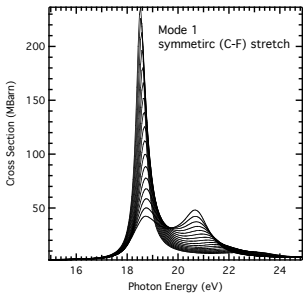
σ^*(C-F) and σ^*(C-C)
E=18.5 eV, Γ = 0.31 eV

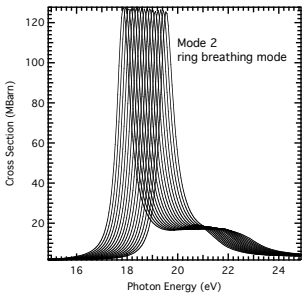


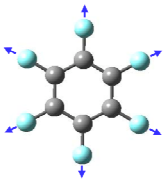
σ^*(C-F) and π^*(C-C)
E=21.0 eV, Γ = 0.49 eV

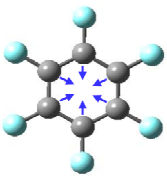
TAMUSCF, Annual Meeting, 2008

Response to the Two Symmetric Modes



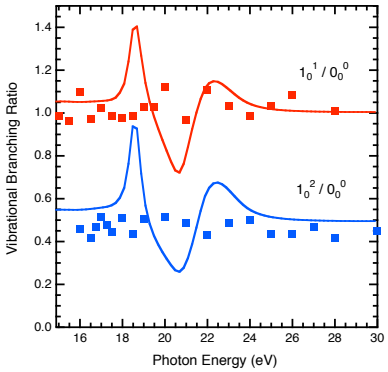




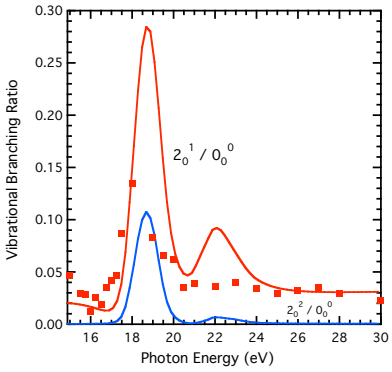


TAMUSCF, Annual Meeting, 2008

C_6F_6 Vibrational Branching Ratios



Mode 1



Mode 2

Experiments of E. Poliakoff and coworkers, Louisiana State University

Conclusions

- Parallelization has been achieved using up to 32 processors
- The combination of a parallel code and a significant computational resource, i.e. hydra, has allowed us to consider much larger systems
- To go beyond 32 processors
  - Additional work on the bottleneck, GHomo
  - Consider larger systems
    - systems with no symmetry
    - electron correlation, i. e. multichannel
  - Additional coarse grained parallelization, e. g. more that one scattering energy at one time
- http://www.chem.tamu.edu/rgroup/lucchese/ePolyScat.E2.manual/manual.html