

Successes and challenges in computational chemistry for the energy sciences



Robert J. Harrison

harrisonrj@ornl.gov

robert.harrison@utk.edu



The DOE funding

- This work is funded by the U.S. Department of Energy, the division of Basic Energy Science, Office of Science, under contract DE-AC05-00OR22725 with Oak Ridge National Laboratory. This research was performed in part using
 - resources of the National Energy Scientific Computing Center which is supported by the Office of Energy Research of the U.S. Department of Energy under contract DE-AC03-76SF0098,
 - and the Center for Computational Sciences at Oak Ridge National Laboratory under contract DE-AC05-00OR22725 .



SciDAC

Scientific Discovery through Advanced Computing

An Integrated Approach to the Rational Design of Chemical Catalysts: NCCS Incite Project

Robert J. Harrison: PI

Oak Ridge National Laboratory,
University of Tennessee

Edoardo Aprà

Jerzy Bernholc

A.C. Buchanan III

Marco Buongiorno Nardelli

James M. Caruthers

W. Nicholas Delgass

David A. Dixon

Sharon Hammes-Schiffer

Duane D. Johnson

Manos Mavrikakis

Vincent Meunier

Mathew Neurock

Steven H. Overbury

William F. Schneider

William A. Shelton

David Sherrill

Bobby G. Sumpter

Kendall T. Thomson

Roberto Ansaloni

Carlo Cavazzoni

Oak Ridge National Laboratory

North Carolina State University

Oak Ridge National Laboratory

North Carolina State University

Purdue University

Purdue University

University of Alabama

Pennsylvania State University

University of Illinois at Urbana Champaign

University of Wisconsin at Madison

Oak Ridge National Laboratory

University of Virginia

Oak Ridge National Laboratory

University of Notre Dame

Oak Ridge National Laboratory

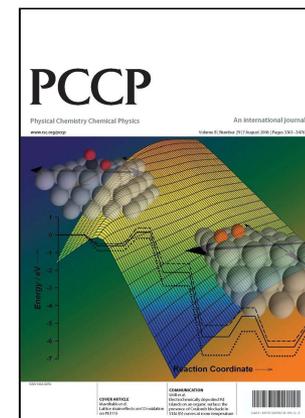
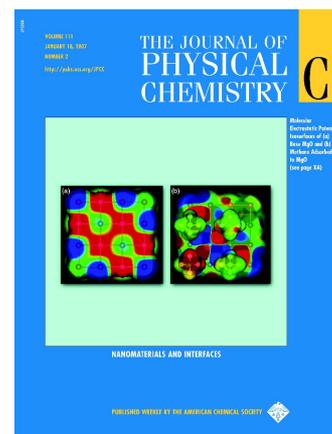
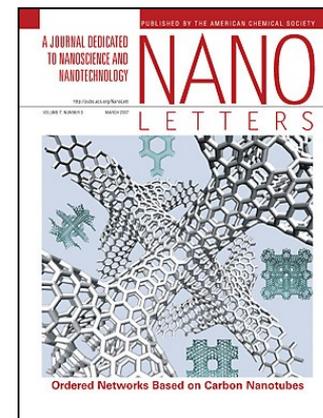
Georgia Institute of Technology

Oak Ridge National Laboratory

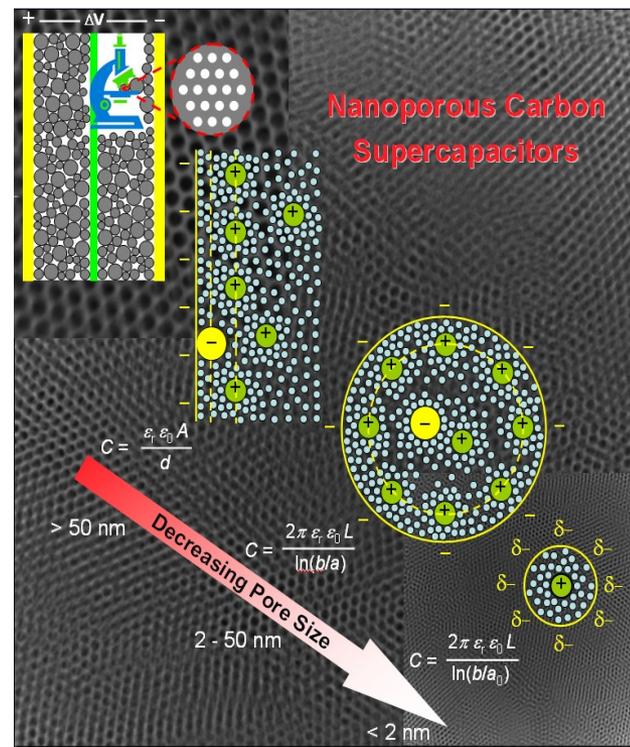
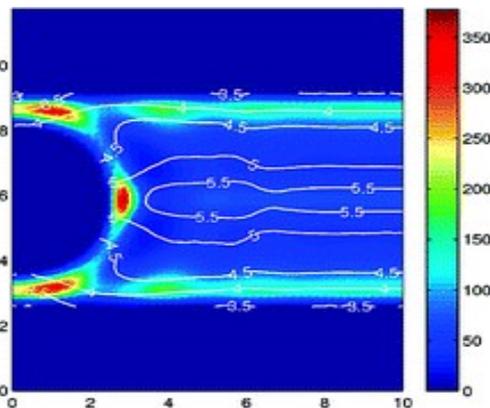
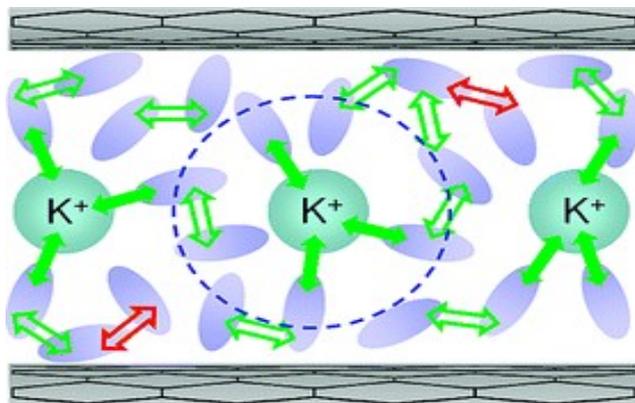
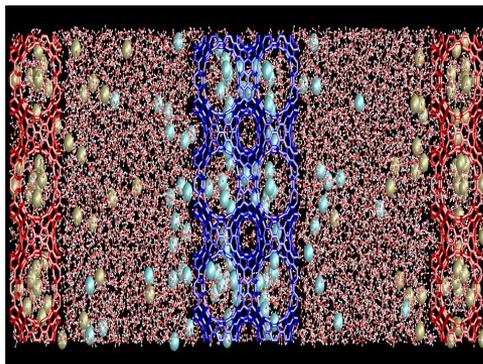
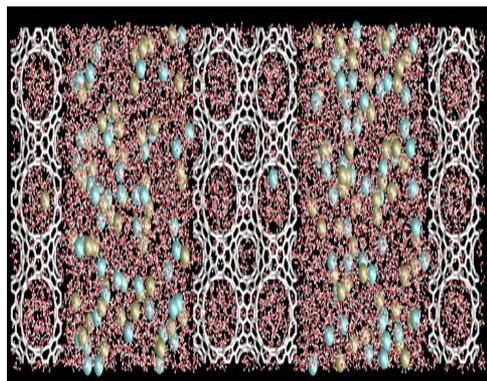
Purdue University

Cray

CINECA, Italy



Theory and modeling of Supercapacitors for Energy Storage (Sumpter and Meunier)



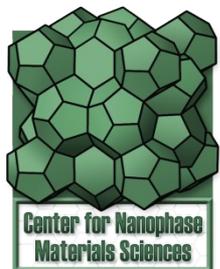
Supercapacitors are electrical energy storage devices that have large energy power and density. The density is tremendously improved by reducing pore sizes in the nanometer regimes. Computational studies show that partial desolvation is important, and leads to replacement of the conventional model by a “wire-in-a-cylinder or sandwich” model, which accounts for all the observed phenomena.

Challenges

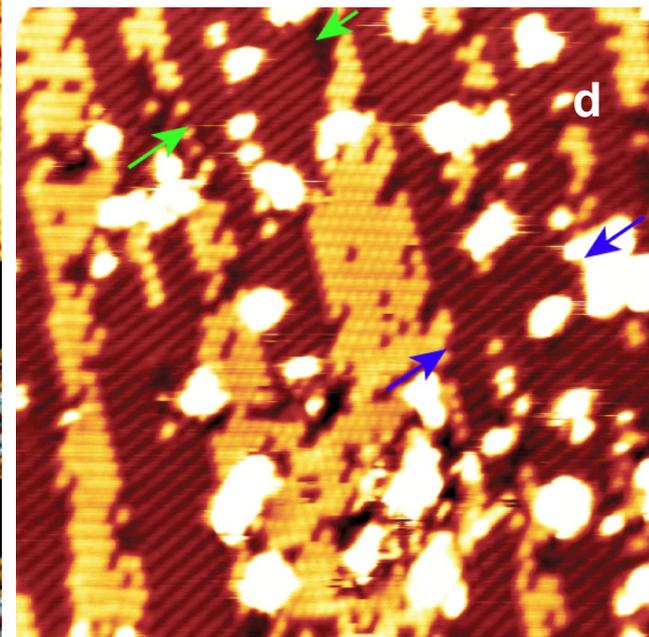
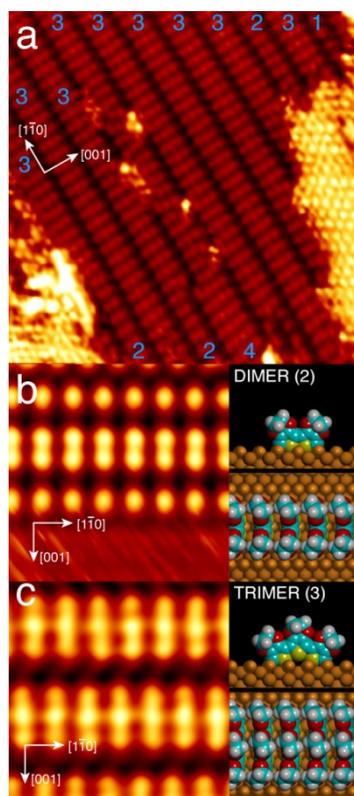
- Time and length scales
 - Need *ab initio* dynamics – strong scaling
- Limitations of current density functionals
 - Long-range interactions, exchange
- Speed and accuracy of many-body methods
 - Slow convergence with basis (solved by explicitly correlated wave function)
 - Poor scaling with system size $O(N^7)$
(solved by local-correlation & low-rank)

Step-by-step growth of epitaxially aligned polythiophene by surface-confined oligomerization

J. A. Lipton-Duffin, J. A. Miwa, M. Kondratenko, F. Cicoira, **B. G. Sumpter**, **V. Meunier**, D. F. Perepichka, F. Rosei



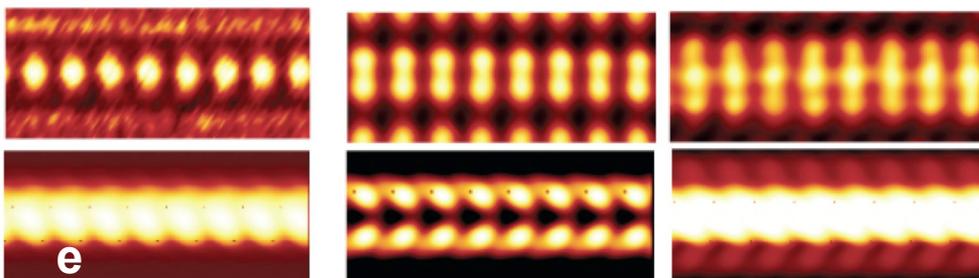
- Demonstrated the emergent surface-confined growth of ordered arrays of PEDOT chains using *in situ* STM imaging combined with first principles density functional theory calculations
- Developed controlled assembly of aromatic building blocks into ordered structures on a metallic substrate with mechanically and electrically robust contacts
- The epitaxially aligned conjugated systems show promise for use in future electronic and optoelectronic devices requiring high-quality active materials.



Monomer

Dimer

Trimer

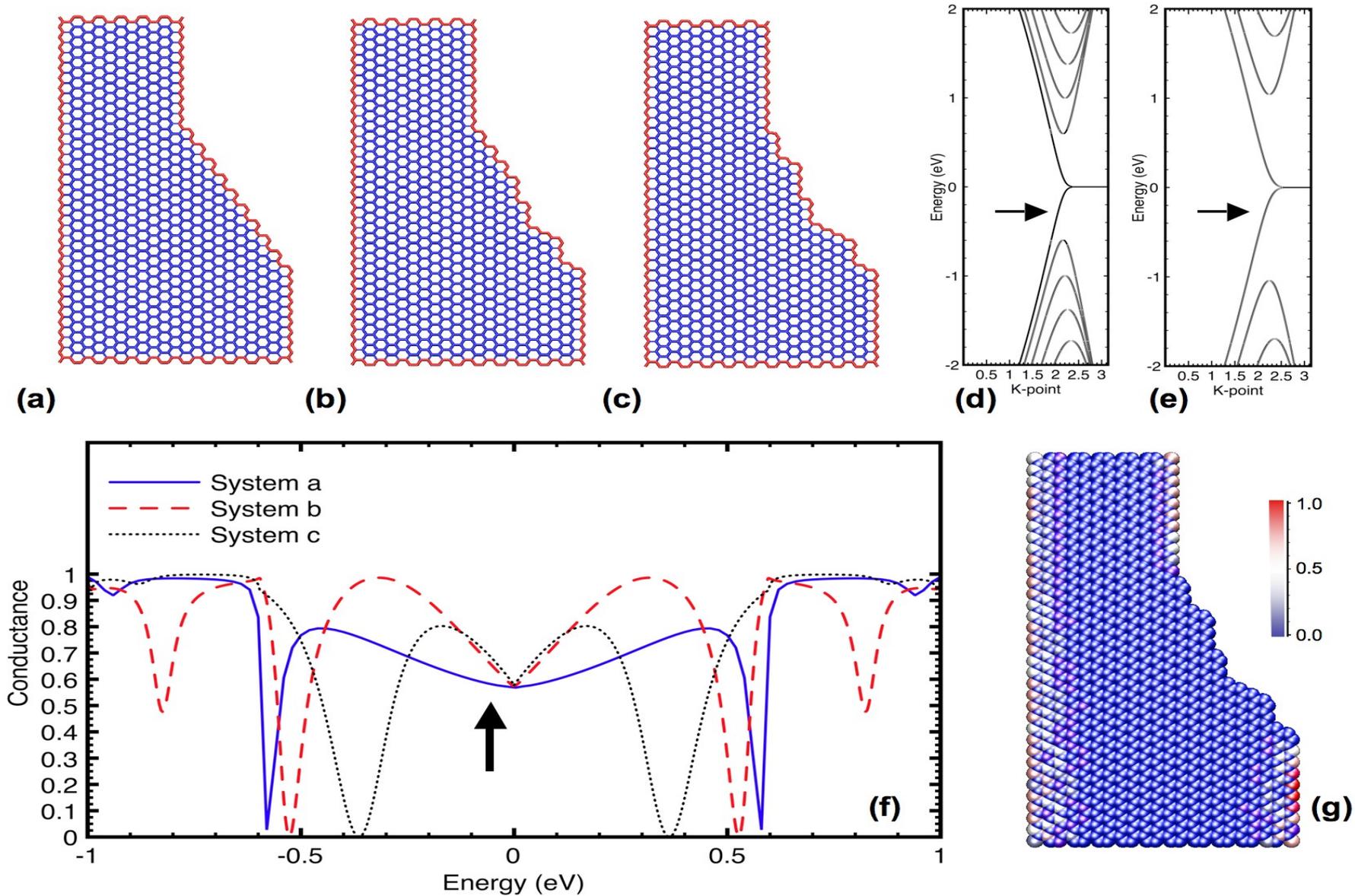


PEDOT oligomers, showing the formation of dimers, trimers and tetramers (a), dimers (b), trimers (c) and the characteristic length of polymers that are grown (d): arrows on (d) mark oligomers with 14 and 16 monomers. DFT computed geometries are shown on (b) and (c). (e) shows simulated (lower) STM images compared to the measured STM (upper).

Challenges

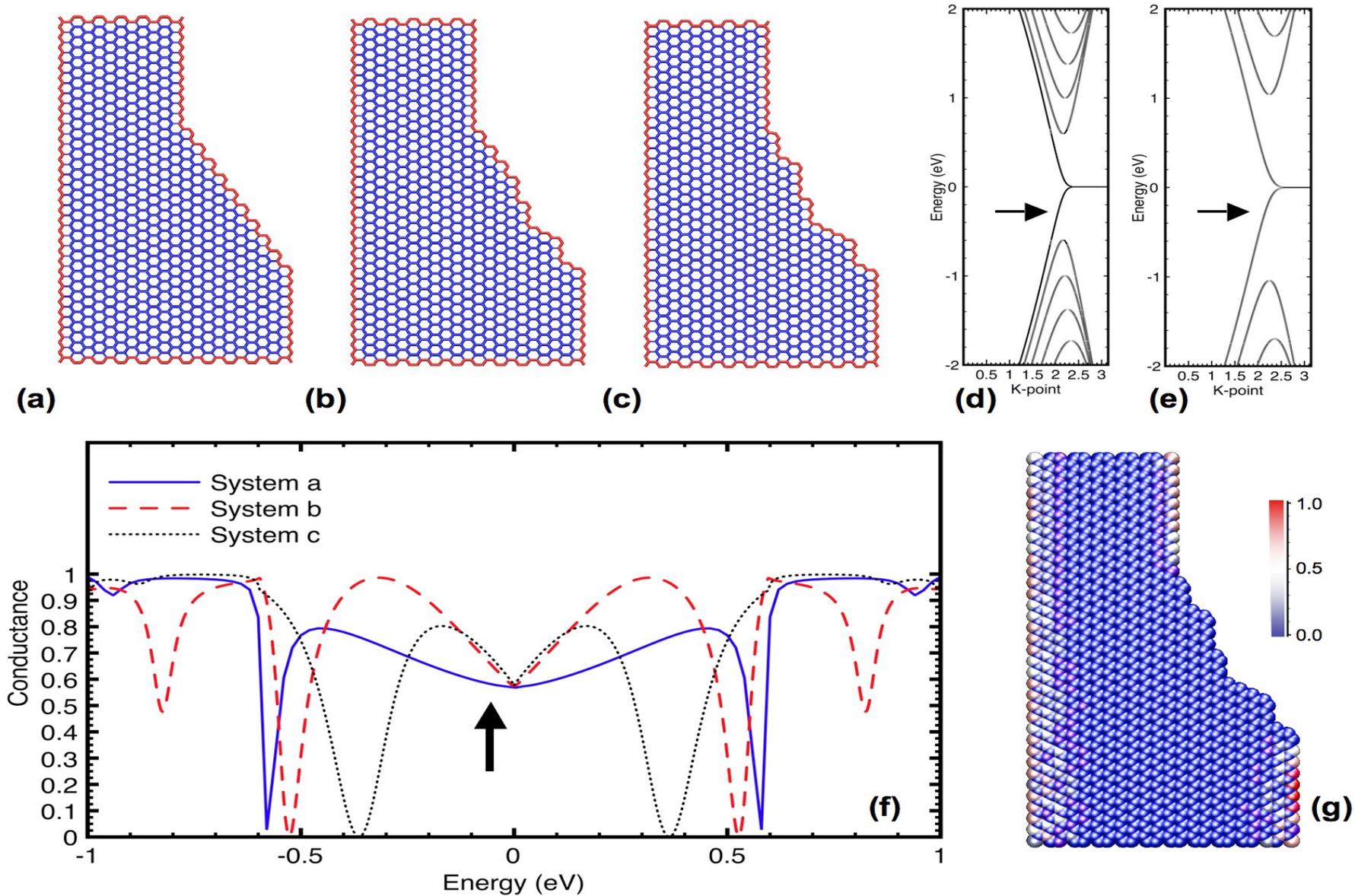
- Chemistry at interfaces
- Time and length scales
 - Need *ab initio* dynamics – strong scaling
- Limitations of current density functionals
 - Incompatible screening models (or fudge factors)
- Many-body methods for materials
 - Much less mature than for molecules (implementation complexity can be eliminated as done in chemistry)
 - Incompatible formulations (propagator methods less developed in chemistry)
 - Lots of calibration of numerous models still required
- Accurate treatment of 1e and 2e excited states

Formation of “sharp edges” in graphene nano ribbon: quantum transport and electronic structure calculations



Science 323, 1701 (2009) Jia, Dresselhus, Sumpter, Meunier, Terrones, ...

Formation of “sharp edges” in graphene nano ribbon: quantum transport and electronic structure calculations



Science 323, 1701 (2009) Jia, Dresselhus, Sumpter, Meunier, Terrones, ...

Challenges

- Time and length scales; rare events
- Coupling of electronic and nuclear motion
 - Electron-phonon interactions – inelastic transport (response calculation for 1000s of atoms)
- Open-shell systems – complex chemistry
- Current DFT functionals neglect long-range correlations
- Spectroscopy (Raman – 3rd derivatives!)

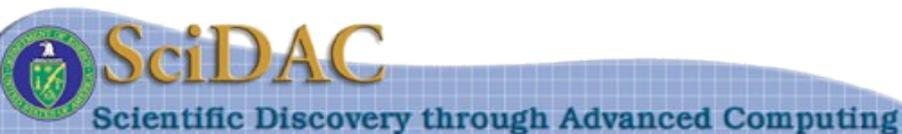
What's in the semantic gap?

Robert J. Harrison
Oak Ridge National Laboratory
and
The University of Tennessee, Knoxville

harrisonrj@ornl.gov



National Science Foundation
WHERE DISCOVERIES BEGIN



HPC futures we'd like to avoid



High-end simulation is the most credible vehicle for accelerating the application of knowledge from basic science to the design of new energy technologies.

Complexity constrains all of our ambitions (cost & feasibility).

We hear about the successes, but what about the failures?

What about the disciplines that are not computing?

O(1) programmers
O(10,000) nodes
O(100,000) processors
O(100,000,000) threads
and growing

- **Growing intrinsic complexity of problem**
- **Complexity kills ... sequential or parallel**
 - **Expressing concurrency at extreme scale**
 - **Managing the memory hierarchy**
- **Semantic gap (Colella)**
 - **Our equations are O(100) lines but**
 - **The program is O(100K) & growing**
 - **Why?**

Wish list

- Eliminate gulf between theoretical innovation in small groups and realization on high-end computers
- Eliminate the semantic gap so that efficient parallel code is no harder than doing the math
- Enable performance-portable “code” that can be automatically migrated to future architectures
- Reduce cost at all points in the life cycle

- Much of this is pipe dream – but what can we aspire to?

Guy L. Steele, Jr.

From Wikipedia, the free encyclopedia

Guy Lewis Steele Jr., (pronounced */ˈsti:l/* as in *steel*), also known as "The Great Quux"^[1] and GLS (*/ˈɡlɪs/*), is an American [computer scientist](#) who has played an important role in designing and documenting several computer [programming languages](#).

- Lisp, Scheme, Java, Fortress

fibonacci($n:\mathbb{Z}$): \mathbb{Z} requires $\{ n \geq 0 \} =$
if $n < 2$ then n else *fibonacci*($n - 1$) + *fibonacci*($n - 2$) end

Guy Steele's example

- Natural language problem description

```
I will give you some text - please remove  
white space and give me the list of words
```

- The wetware of an undistracted first grader is capable of running this two-line “program” with no more guidance

Missing ingredients

- Natural language problem description
 - Implicit & fuzzy definitions of verbs, nouns, ...
 - Incomplete without more context
 - E.g., what if an American first grader was given
- وقال خليل إبراهيم في برنامج ضيف المنتصف على قناة الجزيرة اليوم إن "ما يحدث في الدوحة هو تزوير لأن الأطراف الرئيسية غير موجودة والحرب دائرة على الأرض وتمتد للمدن".
- No ordering specified – parallelism
- Does not include an algorithm

Conventional solution

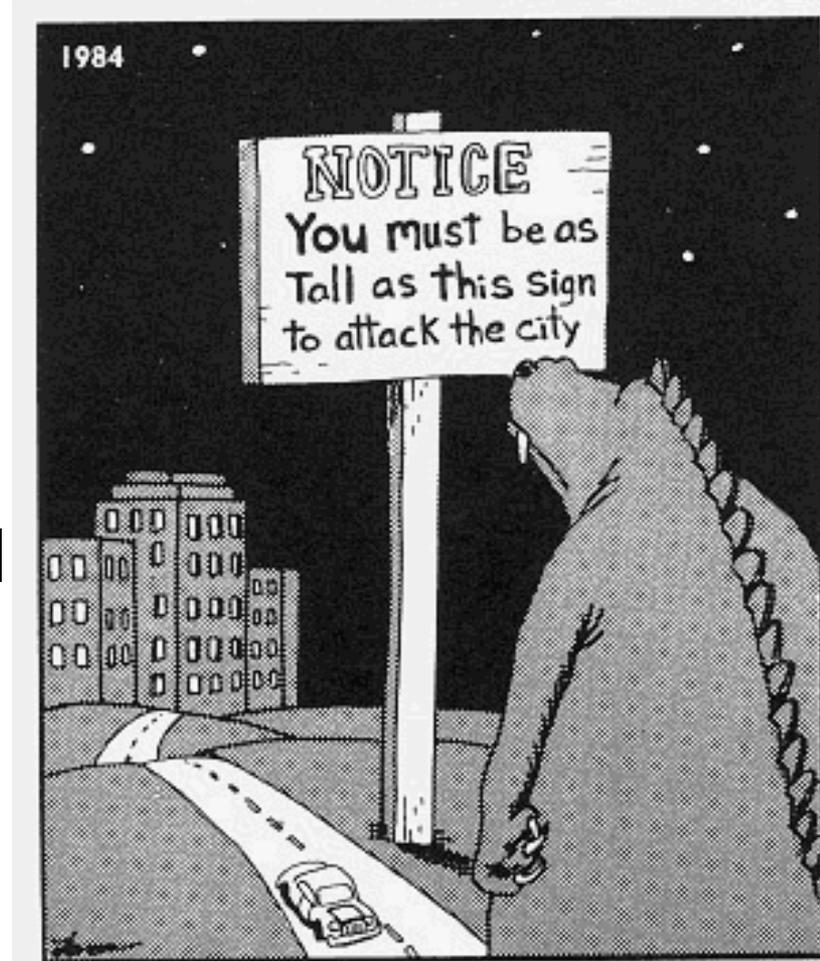
- Problem statement + brain
→ algorithm
- Algorithm + language + brain
→ program
- Computer + program + input
→ result
- The brain is
 - Expensive
 - Finite
 - Not growing exponentially



This is clearly Guy's brain
- Flashes of inspiration
- Blue aura of authority

Impact of sustained exponential growth

- We are only beginning to realize the transforming power of computing as an enabler of innovation and discovery.
- A characteristic of exponential growth is that we will make as much progress in the next doubling cycle as we've made since the birth of the field:
 - 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, ...

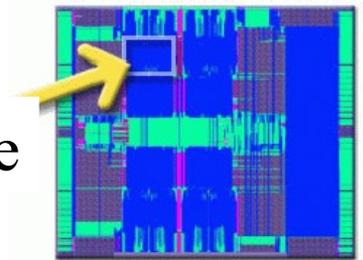
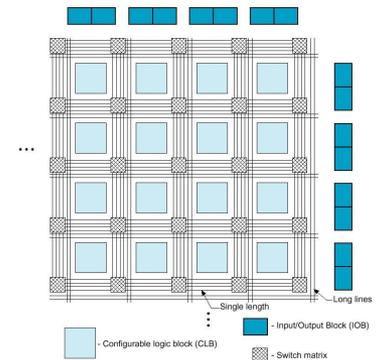


Scientific vs. WWW software

- Why are we not experiencing the same nearly exponential growth in functionality?
 - Level of investment or number of developers?
 - Lack of software interoperability and standards?
 - Competition not cooperation between groups?
 - Shifting scientific objectives?
 - Our problems are intrinsically harder?
 - Failure to embrace/develop higher levels for composing applications?
 - Differing impact of hardware complexity?

What is a processor in 2015?

- Probably not an x86
- E.g., Nvidia C2060 480 cores
 - ~1TF peak (single), ~0.5TF (double)
- LANL RoadRunner – hybrid
 - 1 PF of IBM Cell + Opteron
- Cyclops
 - 160 simple thread units per chip
 - no cache - S/W managed on-chip memory
- FLOPs are cheap; bandwidth is expensive



There is no escape

- Trickle-down computing
 - 2010's national supercomputer is
 - 2013's campus resource
 - 2016's group cluster
 - 2019's personal server
- Why petascale computing now?
 - Because you have large problems now
 - Because you want to prepare for the future

Back to the Gap – what did Guy's brain do?

- Complete the specification with mathematical rigor
 - Including aspects of representation
- Provide an algorithm
- Whence came the algorithm?
 - Derived from the specification?
 - Instantly and unconsciously pattern matched against decades of prior experience?
- Express the algorithm as Fortress code



Fortunately for scientific HPC

- Mathematical rigor is the norm
 - Which partially explains why some disciplines are late to the table
- Unfortunately, also the norm are neglect or ignorance of
 - difference between algebraic and floating-point numbers,
 - accuracy, stability, and other properties of common algorithms
 - parallel algorithms and programming

Frameworks

- NWChem
- MADNESS
- ChemES – chemistry end-station
- Frameworks
 - Increase productivity; hide complexity
 - Interface disciplines
 - Capture knowledge
 - Open HPC to wider community
 - Expensive, communal projects

Molecular Science Software Project

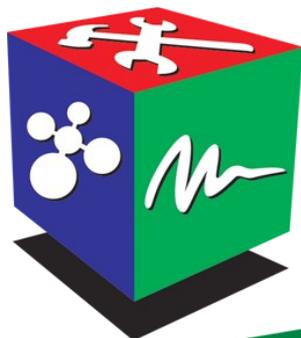


PNNL

**Yuri Alexeev,
Eric Bylaska,
Bert deJong,
Mahin Hackler,
Karol Kowalski,
Lisa Pollack,
Tjerk Straatsma,
Marat Valiev,
Theresa Windus**

ORNL

**Edo Apra,
Vincent Meunier
Robert Harrison**



MS³

MOLECULAR SCIENCE
SOFTWARE SUITE



ECCE

EXTENSIBLE COMPUTATIONAL
CHEMISTRY ENVIRONMENT



NWChem

HIGH-PERFORMANCE COMPUTATIONAL
CHEMISTRY SOFTWARE



GA TOOLS

PARALLEL COMPUTING LIBRARIES
AND SOFTWARE TOOLS

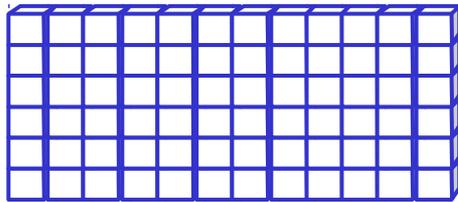
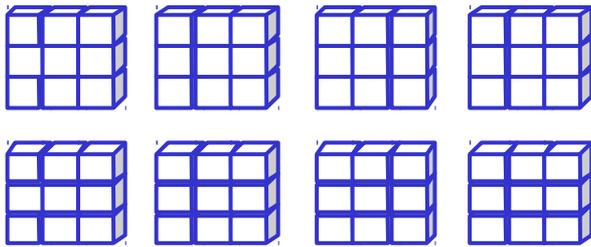
**Manoj Krishnan, Jarek Nieplocha,
Bruce Palmer, Vinod Tipparaju**



**Gary Black,
Brett Didier,
Todd Elsenthagen,
Sue Havre,
Carina Lansing,
Bruce Palmer,
Karen Schuchardt,
Lisong Sun
Erich Vorpapel**

Global Arrays (technologies)

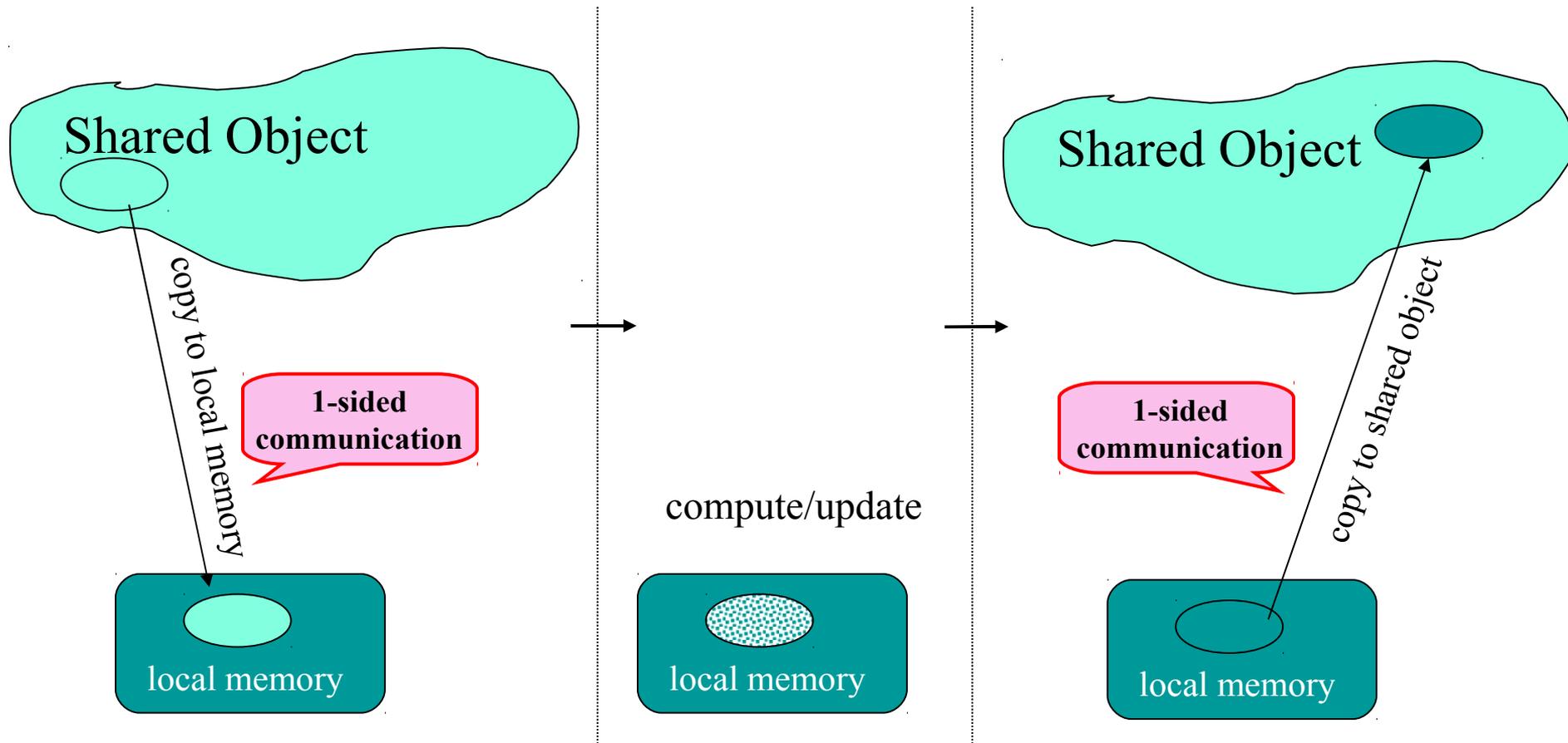
Physically distributed data



Single, shared data structure

- Shared-memory-like model
 - Fast local access
 - NUMA aware and easy to use
 - MIMD and data-parallel modes
 - Inter-operates with MPI, ...
- BLAS and linear algebra interface
- Ported to major parallel machines
 - IBM, Cray, SGI, clusters,...
- Originated in an HPCC project
- Used by most major chemistry codes, financial futures forecasting, astrophysics, computer graphics
- Supported by DOE
- A legacy of Jarek Nieplocha, PNNL

Non-uniform memory access model of computation



Dynamic load balancing

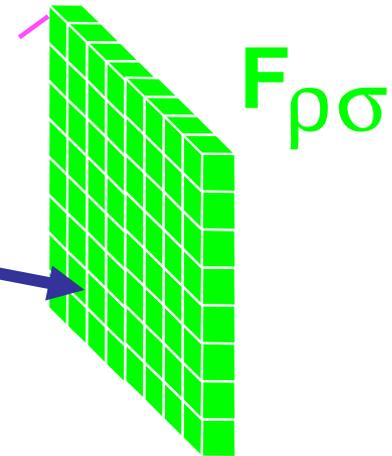
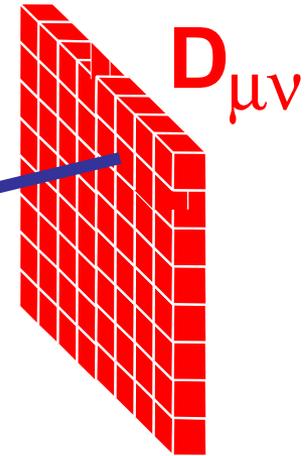
While ((task = SharedCounter()) < max)

call ga_get()

(do work)

call ga_acc()

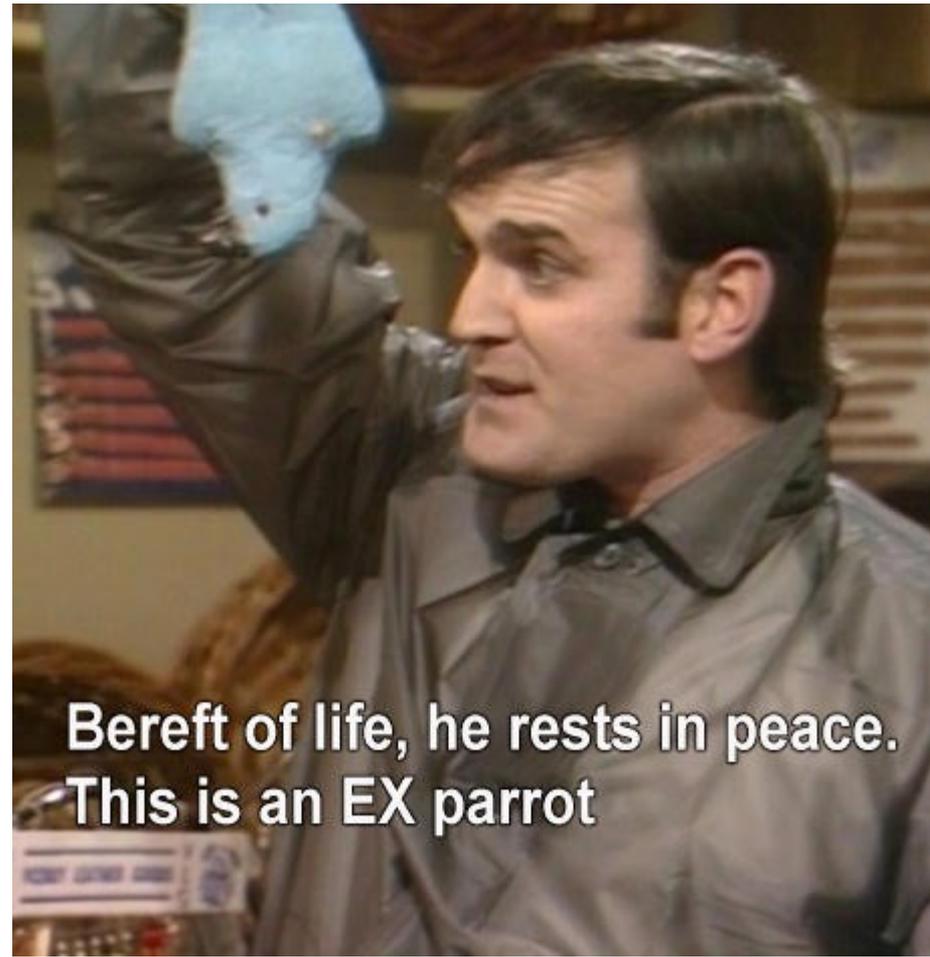
End while
Barrier()



Dead code

7 December 1969

- Requires human labor
 - to migrate to future architectures, or
 - to exploit additional concurrency, or
 - ...
- By these criteria most extant code is dead
- Sanity check
 - How much effort is required to port to hybrid cpu+GPGPU?



Next generation ORNL NLCF

- ORNL has proposed a system to meet DOE's requirement for 20-40 PF of compute capability split between the Oak Ridge and Argonne LCF centers
- ORNL's proposed system will be based on accelerator technology
 - includes software development environment
- We plan to deploy the system in late 2011 with users getting access in 2012
- Watch this space for more details soon

<http://www.nccs.gov/>



The language of many-body physics

$$\Phi_{GW} = \frac{1}{2} \text{Hartree} - \frac{1}{2} \text{Fock} - \frac{1}{4} \text{bubble}_1 - \frac{1}{6} \text{bubble}_2 - \frac{1}{8} \text{bubble}_3 - \dots$$

Hartree

Fock

Infinite chain of **dressed**
electron-hole bubbles

CCSD Doubles Equation

$$\begin{aligned} \bar{h}[a,b,i,j] = & \text{sum}[f[b,c]^*t[i,j,a,c],\{c\}] - \text{sum}[f[k,c]^*t[k,b]^*t[i,j,a,c],\{k,c\}] + \text{sum}[f[a,c]^*t[i,j,c,b],\{c\}] - \text{sum}[f[k,c]^*t[k,a]^*t[i,j,c,b],\{k,c\}] \\ & - \text{sum}[f[k,j]^*t[i,k,a,b],\{k\}] - \text{sum}[f[k,c]^*t[j,c]^*t[i,k,a,b],\{k,c\}] - \text{sum}[f[k,i]^*t[j,k,b,a],\{k\}] - \text{sum}[f[k,c]^*t[i,c]^*t[j,k,b,a],\{k,c\}] \\ & + \text{sum}[t[i,c]^*t[j,d]^*v[a,b,c,d],\{c,d\}] + \text{sum}[t[i,j,c,d]^*v[a,b,c,d],\{c,d\}] + \text{sum}[t[j,c]^*v[a,b,i,c],\{c\}] - \text{sum}[t[k,b]^*v[a,k,i,j],\{k\}] \\ & + \text{sum}[t[i,c]^*v[b,a,j,c],\{c\}] - \text{sum}[t[k,a]^*v[b,k,j,i],\{k\}] - \text{sum}[t[k,d]^*t[i,j,c,b]^*v[k,a,c,d],\{k,c,d\}] - \text{sum}[t[i,c]^*t[j,k,b,d]^*v[k,a,c,d], \\ & \{k,c,d\}] - \text{sum}[t[j,c]^*t[k,b]^*v[k,a,c,i],\{k,c\}] + 2*\text{sum}[t[j,k,b,c]^*v[k,a,c,i],\{k,c\}] - \text{sum}[t[j,k,c,b]^*v[k,a,c,i],\{k,c\}] \\ & - \text{sum}[t[i,c]^*t[j,d]^*t[k,b]^*v[k,a,d,c],\{k,c,d\}] + 2*\text{sum}[t[k,d]^*t[i,j,c,b]^*v[k,a,d,c],\{k,c,d\}] - \text{sum}[t[k,b]^*t[i,j,c,d]^*v[k,a,d,c],\{k,c,d\}] \\ & - \text{sum}[t[j,d]^*t[i,k,c,b]^*v[k,a,d,c],\{k,c,d\}] + 2*\text{sum}[t[i,c]^*t[j,k,b,d]^*v[k,a,d,c],\{k,c,d\}] - \text{sum}[t[i,c]^*t[j,k,d,b]^*v[k,a,d,c],\{k,c,d\}] \\ & - \text{sum}[t[j,k,b,c]^*v[k,a,i,c],\{k,c\}] - \text{sum}[t[i,c]^*t[k,b]^*v[k,a,j,c],\{k,c\}] - \text{sum}[t[i,k,c,b]^*v[k,a,j,c],\{k,c\}] \\ & - \text{sum}[t[i,c]^*t[j,d]^*t[k,a]^*v[k,b,c,d],\{k,c,d\}] - \text{sum}[t[k,d]^*t[i,j,a,c]^*v[k,b,c,d],\{k,c,d\}] - \text{sum}[t[k,a]^*t[i,j,c,d]^*v[k,b,c,d],\{k,c,d\}] \\ & + 2*\text{sum}[t[j,d]^*t[i,k,a,c]^*v[k,b,c,d],\{k,c,d\}] - \text{sum}[t[j,d]^*t[i,k,c,a]^*v[k,b,c,d],\{k,c,d\}] - \text{sum}[t[i,c]^*t[j,k,d,a]^*v[k,b,c,d],\{k,c,d\}] \\ & - \text{sum}[t[i,c]^*t[k,a]^*v[k,b,c,j],\{k,c\}] + 2*\text{sum}[t[i,k,a,c]^*v[k,b,c,j],\{k,c\}] - \text{sum}[t[i,k,c,a]^*v[k,b,c,j],\{k,c\}] \\ & + 2*\text{sum}[t[k,d]^*t[i,j,a,c]^*v[k,b,d,c],\{k,c,d\}] - \text{sum}[t[j,d]^*t[i,k,a,c]^*v[k,b,d,c],\{k,c,d\}] - \text{sum}[t[j,c]^*t[k,a]^*v[k,b,i,c],\{k,c\}] \\ & - \text{sum}[t[j,k,c,a]^*v[k,b,i,c],\{k,c\}] - \text{sum}[t[i,k,a,c]^*v[k,b,j,c],\{k,c\}] + \text{sum}[t[i,c]^*t[j,d]^*t[k,a]^*t[l,b]^*v[k,l,c,d],\{k,l,c,d\}] \\ & - 2*\text{sum}[t[k,b]^*t[l,d]^*t[i,j,a,c]^*v[k,l,c,d],\{k,l,c,d\}] - 2*\text{sum}[t[k,a]^*t[l,d]^*t[i,j,c,b]^*v[k,l,c,d],\{k,l,c,d\}] \\ & + \text{sum}[t[k,a]^*t[l,b]^*t[i,j,c,d]^*v[k,l,c,d],\{k,l,c,d\}] - 2*\text{sum}[t[j,c]^*t[l,d]^*t[i,k,a,b]^*v[k,l,c,d],\{k,l,c,d\}] \\ & - 2*\text{sum}[t[j,d]^*t[l,b]^*t[i,k,a,c]^*v[k,l,c,d],\{k,l,c,d\}] + \text{sum}[t[j,d]^*t[l,b]^*t[i,k,c,a]^*v[k,l,c,d],\{k,l,c,d\}] \\ & - 2*\text{sum}[t[i,c]^*t[l,d]^*t[j,k,b,a]^*v[k,l,c,d],\{k,l,c,d\}] + \text{sum}[t[i,c]^*t[l,a]^*t[j,k,b,d]^*v[k,l,c,d],\{k,l,c,d\}] \\ & + \text{sum}[t[i,c]^*t[l,b]^*t[j,k,d,a]^*v[k,l,c,d],\{k,l,c,d\}] + \text{sum}[t[i,k,c,d]^*t[j,l,b,a]^*v[k,l,c,d],\{k,l,c,d\}] \\ & + 4*\text{sum}[t[i,k,a,c]^*t[j,l,b,d]^*v[k,l,c,d],\{k,l,c,d\}] - 2*\text{sum}[t[i,k,c,a]^*t[j,l,b,d]^*v[k,l,c,d],\{k,l,c,d\}] \\ & - 2*\text{sum}[t[i,k,a,b]^*t[j,l,c,d]^*v[k,l,c,d],\{k,l,c,d\}] - 2*\text{sum}[t[i,k,a,c]^*t[j,l,d,b]^*v[k,l,c,d],\{k,l,c,d\}] + \text{sum}[t[i,k,c,a]^*t[j,l,d,b]^*v[k,l,c,d], \\ & \{k,l,c,d\}] + \text{sum}[t[i,c]^*t[j,d]^*t[k,l,a,b]^*v[k,l,c,d],\{k,l,c,d\}] + \text{sum}[t[i,j,c,d]^*t[k,l,a,b]^*v[k,l,c,d],\{k,l,c,d\}] \\ & - 2*\text{sum}[t[i,j,c,b]^*t[k,l,a,d]^*v[k,l,c,d],\{k,l,c,d\}] - 2*\text{sum}[t[i,j,a,c]^*t[k,l,b,d]^*v[k,l,c,d],\{k,l,c,d\}] + \text{sum}[t[j,c]^*t[k,b]^*t[l,a]^*v[k,l,c,i], \\ & \{k,l,c\}] + \text{sum}[t[l,c]^*t[j,k,b,a]^*v[k,l,c,i],\{k,l,c\}] - 2*\text{sum}[t[l,a]^*t[j,k,b,c]^*v[k,l,c,i],\{k,l,c\}] + \text{sum}[t[l,a]^*t[j,k,c,b]^*v[k,l,c,i],\{k,l,c\}] \\ & - 2*\text{sum}[t[k,c]^*t[j,l,b,a]^*v[k,l,c,i],\{k,l,c\}] + \text{sum}[t[k,a]^*t[j,l,b,c]^*v[k,l,c,i],\{k,l,c\}] + \text{sum}[t[k,b]^*t[j,l,c,a]^*v[k,l,c,i],\{k,l,c\}] \\ & + \text{sum}[t[j,c]^*t[l,k,a,b]^*v[k,l,c,i],\{k,l,c\}] + \text{sum}[t[i,c]^*t[k,a]^*t[l,b]^*v[k,l,c,j],\{k,l,c\}] + \text{sum}[t[l,c]^*t[i,k,a,b]^*v[k,l,c,j],\{k,l,c\}] \\ & - 2*\text{sum}[t[l,b]^*t[i,k,a,c]^*v[k,l,c,j],\{k,l,c\}] + \text{sum}[t[l,b]^*t[i,k,c,a]^*v[k,l,c,j],\{k,l,c\}] + \text{sum}[t[i,c]^*t[k,l,a,b]^*v[k,l,c,j],\{k,l,c\}] \\ & + \text{sum}[t[j,c]^*t[l,d]^*t[i,k,a,b]^*v[k,l,d,c],\{k,l,c,d\}] + \text{sum}[t[j,d]^*t[l,b]^*t[i,k,a,c]^*v[k,l,d,c],\{k,l,c,d\}] \\ & + \text{sum}[t[j,d]^*t[l,a]^*t[i,k,c,b]^*v[k,l,d,c],\{k,l,c,d\}] - 2*\text{sum}[t[i,k,c,d]^*t[j,l,b,a]^*v[k,l,d,c],\{k,l,c,d\}] \\ & - 2*\text{sum}[t[i,k,a,c]^*t[j,l,b,d]^*v[k,l,d,c],\{k,l,c,d\}] + \text{sum}[t[i,k,c,a]^*t[j,l,b,d]^*v[k,l,d,c],\{k,l,c,d\}] + \text{sum}[t[i,k,a,b]^*t[j,l,c,d]^*v[k,l,d,c], \\ & \{k,l,c,d\}] + \text{sum}[t[i,k,c,b]^*t[j,l,d,a]^*v[k,l,d,c],\{k,l,c,d\}] + \text{sum}[t[i,k,a,c]^*t[j,l,d,b]^*v[k,l,d,c],\{k,l,c,d\}] + \text{sum}[t[k,a]^*t[l,b]^*v[k,l,i,j], \\ & \{k,l\}] + \text{sum}[t[k,l,a,b]^*v[k,l,i,j],\{k,l\}] + \text{sum}[t[k,b]^*t[l,d]^*t[i,j,a,c]^*v[l,k,c,d],\{k,l,c,d\}] + \text{sum}[t[k,a]^*t[l,d]^*t[i,j,c,b]^*v[l,k,c,d], \\ & \{k,l,c,d\}] + \text{sum}[t[i,c]^*t[l,d]^*t[j,k,b,a]^*v[l,k,c,d],\{k,l,c,d\}] - 2*\text{sum}[t[i,c]^*t[l,a]^*t[j,k,b,d]^*v[l,k,c,d],\{k,l,c,d\}] \\ & + \text{sum}[t[i,c]^*t[l,a]^*t[j,k,d,b]^*v[l,k,c,d],\{k,l,c,d\}] + \text{sum}[t[i,j,c,b]^*t[k,l,a,d]^*v[l,k,c,d],\{k,l,c,d\}] + \text{sum}[t[i,j,a,c]^*t[k,l,b,d]^*v[l,k,c,d], \\ & \{k,l,c,d\}] - 2*\text{sum}[t[l,c]^*t[i,k,a,b]^*v[l,k,c,j],\{k,l,c\}] + \text{sum}[t[l,b]^*t[i,k,a,c]^*v[l,k,c,j],\{k,l,c\}] + \text{sum}[t[l,a]^*t[i,k,c,b]^*v[l,k,c,j],\{k,l,c\}] \\ & + v[a,b,i,j] \end{aligned}$$

$$\bar{h}_{ij}^{ab} = \left\langle \begin{array}{c} a \ b \\ i \ j \end{array} \middle| e^{-\hat{T}_1 - \hat{T}_2} \hat{H} e^{\hat{T}_1 + \hat{T}_2} \middle| \mathbf{0} \right\rangle$$

The Tensor Contraction Engine: A Tool for Quantum Chemistry

**Oak Ridge National
Laboratory**

David E. Bernholdt,
Venkatesh Choppella, *Robert
Harrison*

**Pacific Northwest National
Laboratory**
So Hirata

Louisiana State University
J Ramanujam,

Ohio State University

Gerald Baumgartner, Alina
Bibireata, Daniel Cociorva,
Xiaoyang Gao, Sriram
Krishnamoorthy, Sandhya
Krishnan, Chi-Chung Lam,
Quingda Lu, *Russell M.
Pitzer, P Sadayappan,*
Alexander Sibiryakov

University of Waterloo

Marcel Nooijen, Alexander
Auer

<http://www.cis.ohio-state.edu/~gb/TCE/>

Pesky details of an incomplete spec.

- Some tensors have symmetries w.r.t. index permutations

$$\langle p q | r s \rangle = \langle p s | r q \rangle = \dots$$

- Others have predictable block sparsity

$$b_{2u} \times b_{3u} = b_{1g}$$

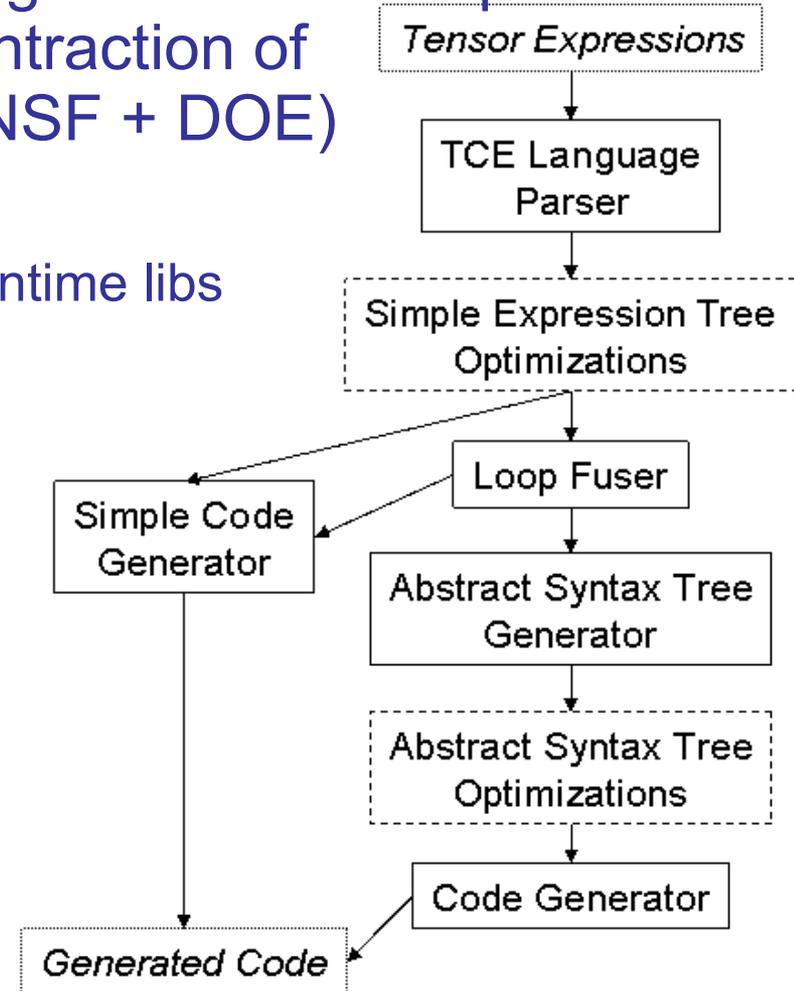
- Huge impact on memory use and algorithm cost
- ... one year later ...

Tensor Contraction Engine (TCE)

- Sadayappan et al. Proc. IEEE, 93, 2005
- High-level domain-specific language for a class of problems in chemistry/physics based on contraction of large multi-dimensional tensors (NSF + DOE)
- Specialized optimizing compiler
 - Produces F77+GA code, linked to runtime libs

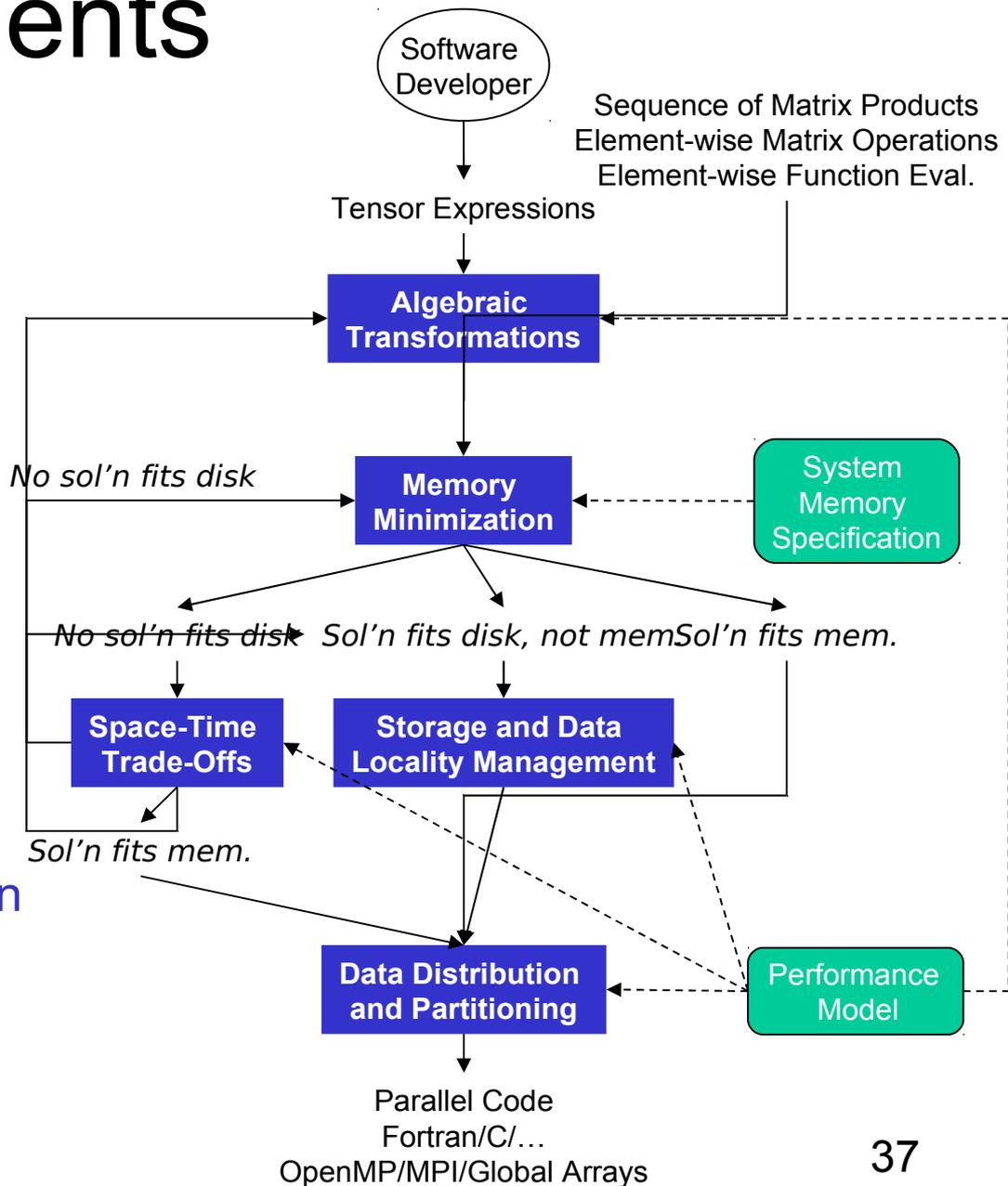
```
range V = 3000;
range O = 100;
index a,b,c,d,e,f : V;
index i,j,k,l : O;
mlimit = 100GB;

procedure P(in A[V,V,O,O], in B[V,V,V,O],
            in C[V,V,O,O], in D[V,V,V,O],
            out S[V,V,O,O])=
begin
  S[a,b,i,j] == sum[ A[a,c,i,k] * B[b,e,f,l]
                    * C[d,f,j,k] * D[c,d,e,l],
                    {c,e,f,k,l}];
end
```

$$S_{abij} = \sum_{cefk} A_{acik} B_{befl} C_{dfjk} D_{cdel}$$


TCE Components

- Algebraic Transformations
 - Minimize operation count
- Memory Minimization
 - Reduce intermediate storage via loop fusion (LCPC'03)
- Space-Time Transformation
 - Trade-offs between storage and recomputation (PLDI'02)
- Data Locality Optimization
 - Optimize use of storage hierarchy via tiling (ICS'01, HiPC'03, IPDPS'04)
- Data Dist./Comm. Optimization
 - Optimize parallel data layout (IPDPS'03)
- Integrated System
 - (SuperComputing'02, Proc. IEEE 05)



Productivity of TCE

- The tensor contraction expressions for the higher members of the Coupled Cluster family of models can be generated relatively easily, but the effort to manually generate Fortran code is quite significant
- The code development time for other models of comparable complexity can be reduced from years to days/weeks
- More than 25 methods implemented → » 5 years to hand code

Theory	#Terms	#F77Lines	Year
CCD	11	3209	1978
CCSD	48	13213	1982
CCSDT	102	33932	1988
CCSDTQ	183	79901	1992

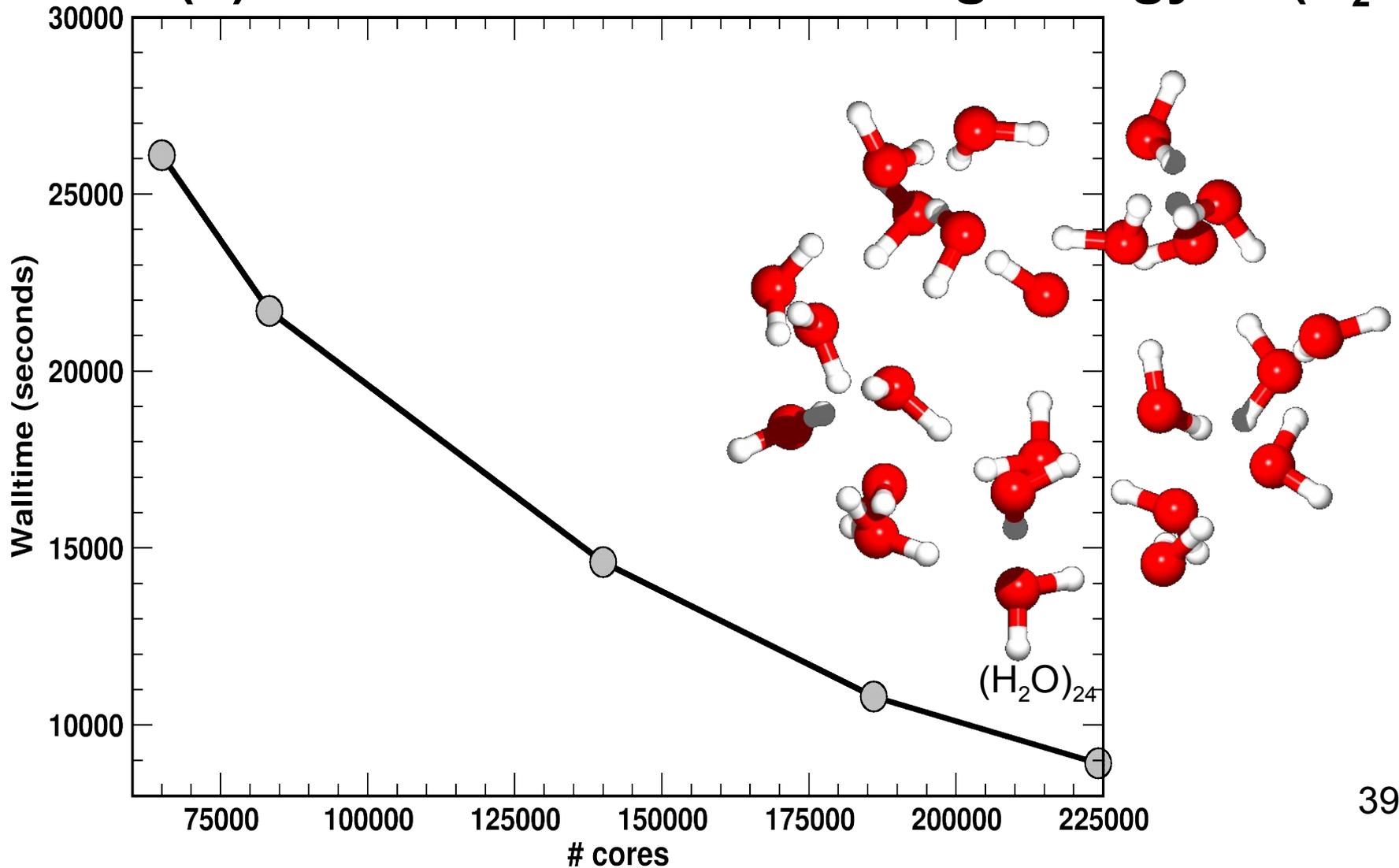


Result of first hand implementation of
CCSDTQ 38

NWChem CCSD(T) – 1.31 PFLOP/s

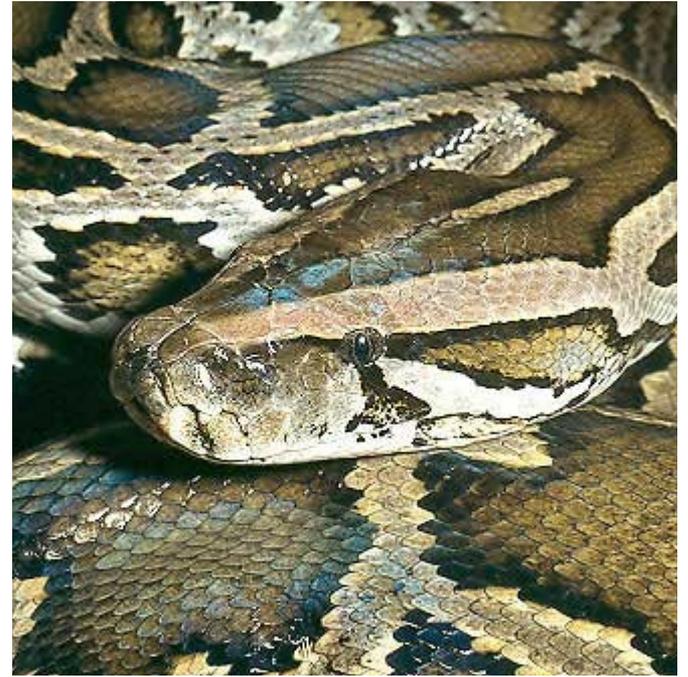
E. Aprà, R.J. Harrison, W.A. deJong, A.P. Rendell, V. Tipparaju and R.M. Olson

CCSD(T) benchmark of the binding energy of $(\text{H}_2\text{O})_{24}$



Python vs. Java

- The initial Python prototype written by chemists works but has lots of “issues” with memory, speed, ...
- The OSU TCE generates better code, respects bounds on memory use, but is written in Java by C/S graduate students
- And none of the chemists have a clue how it works and none of them know Java
- Guess which is in use

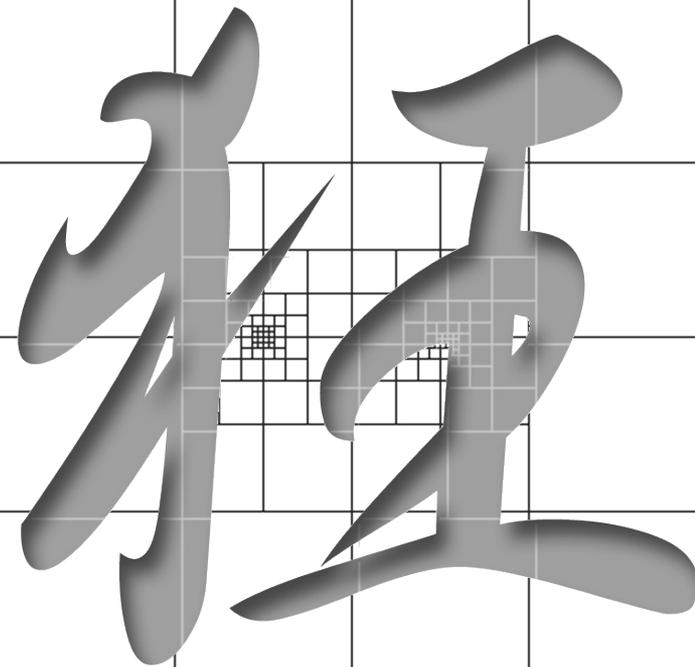


M

A

D

N



T



Multiresolution
Adaptive
Numerical
Scientific
Simulation

S



S

Multiresolution Adaptive Numerical Scientific Simulation

*Ariana Beste¹, George I. Fann¹, Robert J. Harrison^{1,2},
Rebecca Hartman-Baker¹, Judy Hill¹, Jun Jia¹,*

¹Oak Ridge National Laboratory

²University of Tennessee, Knoxville

in collaboration with



National Science Foundation
WHERE DISCOVERIES BEGIN

*Gregory Beylkin⁴, Lucas Monzon⁴,
Martin Mohlenkamp⁵, and Hideo Sekino⁶*

⁴University of Colorado

⁵Ohio University

⁶Toyohashi Technical University, Japan

harrisonrj@ornl.gov



Funding

- MADNESS started as a DOE SciDAC project and the majority of its support still comes from the DOE
- DOE SciDAC, divisions of Advanced Scientific Computing Research and Basic Energy Science, under contract DE-AC05-00OR22725 with Oak Ridge National Laboratory, in part using the National Center for Computational Sciences.
- DARPA HPCS2: HPCS programming language evaluation
- NSF CHE 0625598: Cyber-infrastructure and Research Facilities: Chemical Computations on Future High-end Computers
- NSF CNS-0509410: CAS-AES: An integrated framework for compile-time/run-time support for multi-scale applications on high-end systems
- NSF OCI-0904972: Computational chemistry and physics beyond the petascale

What is MADNESS?

- A general purpose numerical environment for reliable and fast scientific simulation
 - Applications already in nuclear physics, chemistry, atomic physics, material science, with investigations beginning in climate and fusion.
- A general purpose parallel programming environment designed for the petascale
 - Standard C++ with concepts from Cilk, Charm++, HPCS languages, with a multi-threaded runtime that dynamically manages task dependences, scheduling and provides global data view.
 - Compatible by design with existing applications



George Fann



Judy Hill



Gregory Beylkin



Rebecca
Hartman-Baker

Jun Jia
Tetsuya Kato
Justus Calvin
J. Pei



Ariana Beste



Eduard Valeyev



Alvaro Vasquez



Hideo Sekino



Robert Harrison



Nicholas Vence



Takahiro Ii



Scott Thornton



Matt Reuter



Paul Sutter

Why MADNESS

- MADNESS

- Reduces S/W complexity since programmer not responsible for managing dependencies, scheduling, or placement
- Reduces S/W complexity through MATLAB-like level of composition of scientific problems with guaranteed speed and precision
- Reduces numerical complexity by enabling solution of integral instead of differential equations
- Framework makes latest techniques in applied math and physics available to wide audience

The math behind the MADNESS

- Discontinuous spectral element basis
 - High-order convergence ideally suited for modern computer technology
- Multi-resolution analysis for fast algorithms
 - Sparse representation of many integral operators
 - Precision guaranteed through adaptive refinement
- Separated representations of operators and functions
 - Enable efficient computation in many dimensions

Essential techniques for fast computation

- Multiresolution

$$V_0 \subset V_1 \subset \dots \subset V_n$$
$$V_n = V_0 + (V_1 - V_0) + \dots + (V_n - V_{n-1})$$

- Low-separation rank

$$f(x_1, \dots, x_n) = \sum_{l=1}^M \sigma_l \prod_{i=1}^d f_i^{(l)}(x_i) + \mathcal{O}(\varepsilon)$$
$$\|f_i^{(l)}\|_2 = 1 \quad \sigma_l > 0$$

- Low-operator rank

$$A = \sum_{\mu=1}^r u_\mu \sigma_\mu v_\mu^T + \mathcal{O}(\varepsilon)$$
$$\sigma_\mu > 0 \quad v_\mu^T v_\lambda = u_\mu^T u_\lambda = \delta_{\mu\nu}$$

Integral Formulation

- Solving the integral equation
 - Eliminates the derivative operator and related “issues”
 - Converges as fixed point iteration *with no preconditioner*

$$\left(-\frac{1}{2}\nabla^2 + V\right)\Psi = E\Psi$$

$$\Psi = -2\left(-\nabla^2 - 2E\right)^{-1}V\Psi$$

$$= -2G^*(V\Psi)$$

$$(G^*f)(r) = \int ds \frac{e^{-k|r-s|}}{4\pi|r-s|} f(s) \quad \text{in 3D ; } k^2 = -2E$$

Such Green's Functions (bound state Helmholtz, Poisson) can be rapidly and accurately applied with a single, sparse matrix vector product. ⁴⁹

High-level composition

- Close to the physics

$$E = \langle \psi | -\frac{1}{2} \nabla^2 + V | \psi \rangle + \langle \psi \psi | \psi \psi \rangle$$

```
operatorT G = CoulombOperator(k, rlo, thresh);
functionT rho = psi*psi;
double twoe = inner(G(rho),rho);
double pe = 2.0*inner(Vnuc*psi,psi);
double ke = 0.0;
for (int axis=0; axis<3; axis++) {
    functionT dpsi = diff(psi,axis);
    ke += inner(dpsi,dpsi);
}
double energy = ke + pe + twoe;
```

Let

$$\Omega = [-20, 20]^3$$

$$r = x \rightarrow \sqrt{x_0^2 + x_1^2 + x_2^2}$$

$$g = x \rightarrow \exp(-r(x))$$

$$v = x \rightarrow -r(x)^{-1}$$

In

$$\psi = \mathcal{F} g$$

$$\nu = \mathcal{F} v$$

$$S = \langle \psi | \psi \rangle$$

$$V = \langle \psi | \nu * \psi \rangle$$

$$T = \frac{1}{2} * \sum_{i=0}^2 (\langle \nabla_i \psi | \nabla_i \psi \rangle)$$

$$\text{print } S, V, T, \frac{T + V}{S}$$

End

H atom Energy

H atom actual source

```
Let
  Omega = [-20, 20]^3
  r = x -> sqrt(x_0^2 + x_1^2 + x_2^2)
  g = x -> exp(-r(x))
  v = x -> -r(x)^-1
In
  psi = F g
  nu = F v
  S = < psi | psi >
  V = < psi | nu * psi >
  T = 1/2 * sum_i=0^2 < del_i psi | del_i psi >
  print S, V, T, (T + V)/S
End
```

Let

$$\Omega = [-20, 20]^6$$

$$r1 = x \rightarrow \sqrt{x_0^2 + x_1^2 + x_2^2}$$

$$r2 = x \rightarrow \sqrt{x_3^2 + x_4^2 + x_5^2}$$

$$r12 = x \rightarrow \sqrt{(x_0 - x_3)^2 + (x_1 - x_4)^2 + (x_2 - x_5)^2}$$

$$g = x \rightarrow \left(1 + \frac{1}{2} * r12(x)\right) * \exp(-2 * (r1(x) + r2(x)))$$

$$v = x \rightarrow -\frac{2}{r1(x)} - \frac{2}{r2(x)} + \frac{1}{r12(x)}$$

In

$$\psi = \mathcal{F} g$$

$$\nu = \mathcal{F} v$$

$$S = \langle \psi | \psi \rangle$$

$$V = \langle \psi | \nu * \psi \rangle$$

$$T = \frac{1}{2} * \sum_{i=0}^5 (\langle \nabla_i \psi | \nabla_i \psi \rangle)$$

$$\text{print } S, V, T, \frac{T + V}{S}$$

End

He atom
Hylleraas
2-term
6D

Let

$$\Omega = [-20, 20]^3$$

$$r = x \rightarrow \sqrt{x_0^2 + x_1^2 + x_2^2}$$

$$g = x \rightarrow \exp(-2 * r(x))$$

$$v = x \rightarrow -\frac{2}{r(x)}$$

In

$$\nu = \mathcal{F} v$$

$$\phi = \mathcal{F} g$$

$$\lambda = -1.0$$

for $i \in [0, 10]$

$$\phi = \phi * \|\phi\|^{-1}$$

$$V = \nu - \nabla^{-2} (4 * \pi * \phi^2)$$

$$\psi = -2 * (-2 * \lambda - \nabla^2)^{-1} (V * \phi)$$

$$\lambda = \lambda + \frac{\langle V * \phi | \psi - \phi \rangle}{\langle \psi | \psi \rangle}$$

$$\phi = \psi$$

print "iter", i, "norm", $\|\phi\|$, "eval", λ

end

End

He atom Hartree- Fock

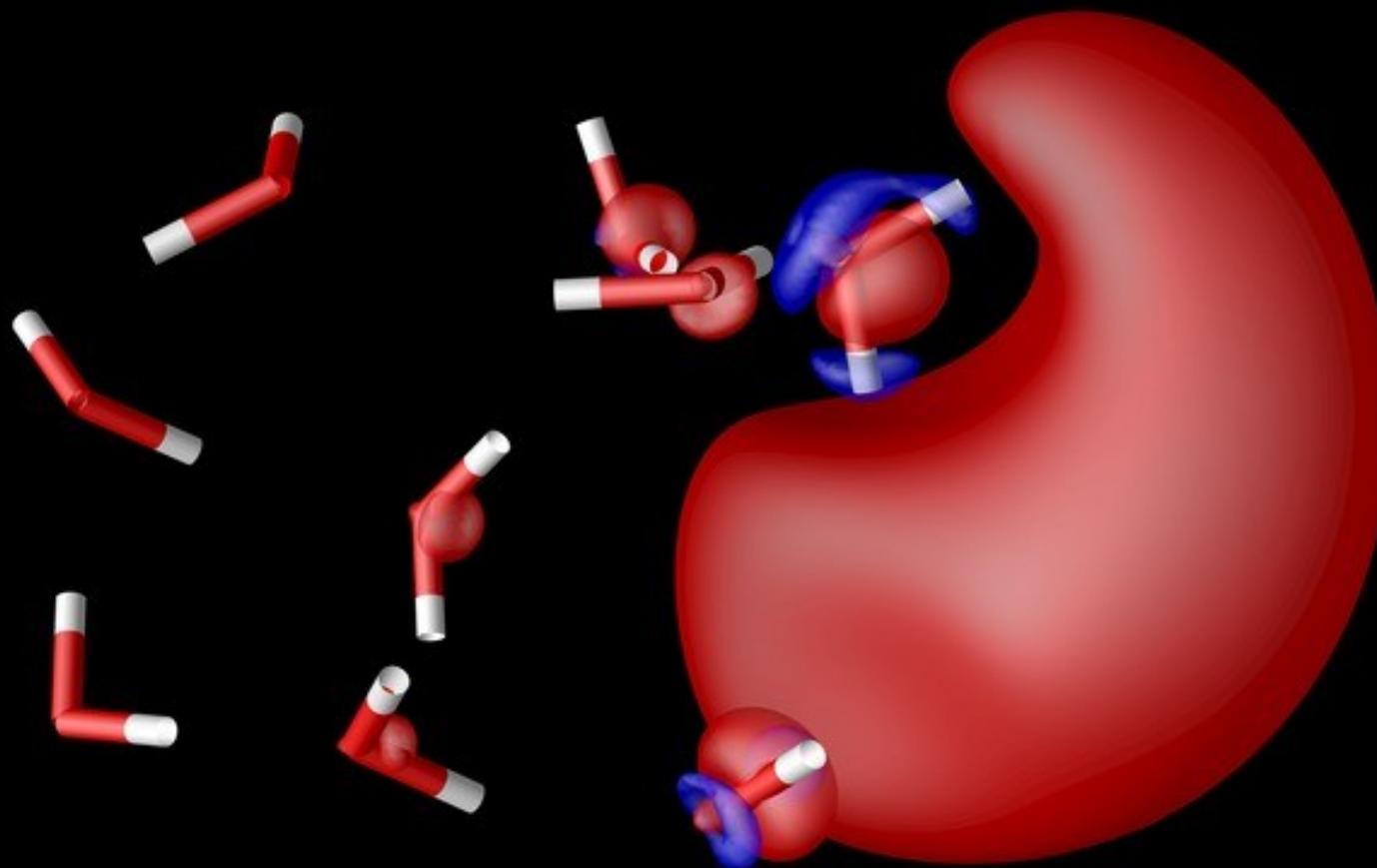
Hartree-Fock

- What I really wanted to type was

$$\min_{\phi} E[\phi] \quad \text{s.t.} \quad \|\phi\|_2 = 1$$

- But had to
 - Provide E (or rather $dE/d\phi$)
 - Describe inexact-Newton algorithm with stopping criterion
 - Transform to integral representation for efficiency and accuracy
- Can automate some steps, c.f. Maple, Mathematica
 - But properties of computation in the underlying basis are crucial for accuracy and efficiency
 - So let's go back and ask why is this working ...

Molecular HF and DFT



Energy and
gradients

ECPs coming
(Sekino)

Response
properties
(Vasquez and
Sekino)

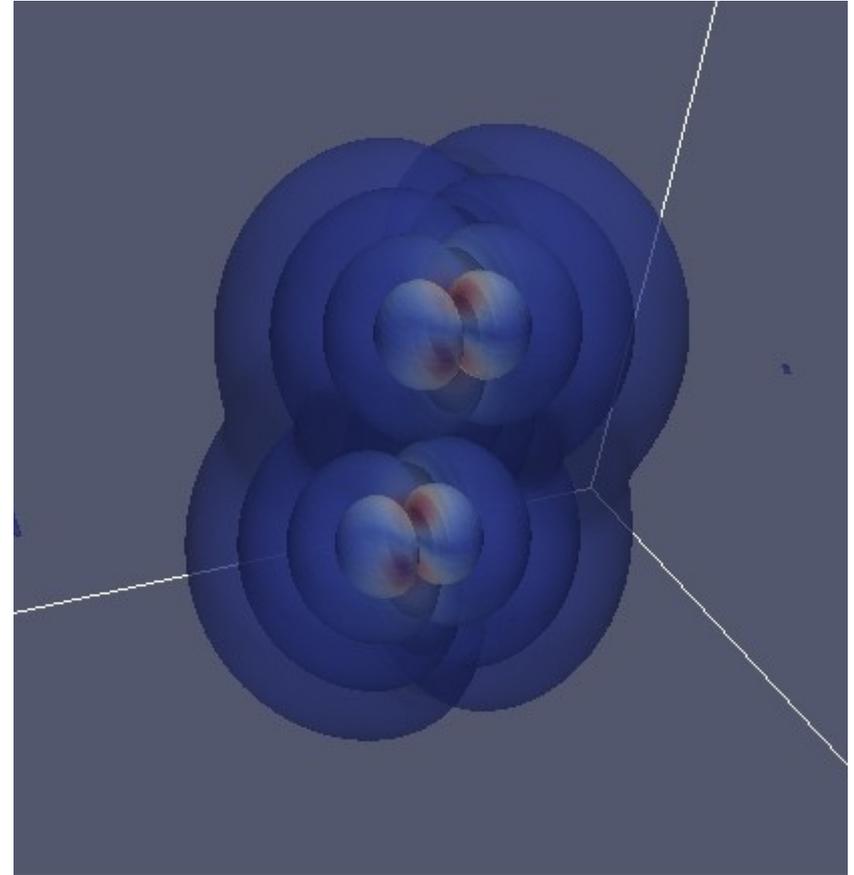
Still not as
functional as
previous
python version

*Spin density
of solvated
electron*⁵⁶

Nuclear physics

Pei, Fann, Ou, Nazarewicz

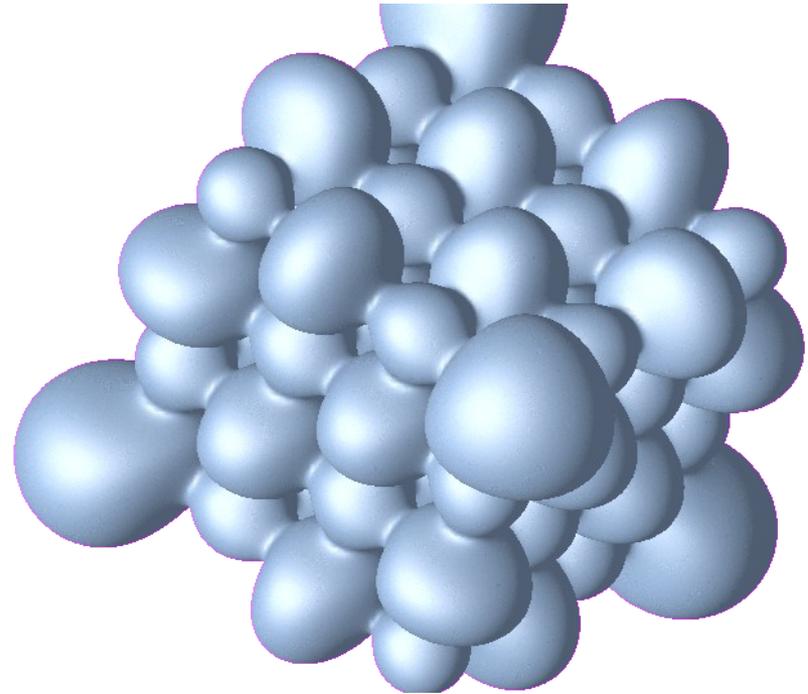
- DOE UNDEF
- Nuclei & neutron matter
- ASLDA
- Hartree-Fock Bogliobulov
- Spinors
- Gamov states



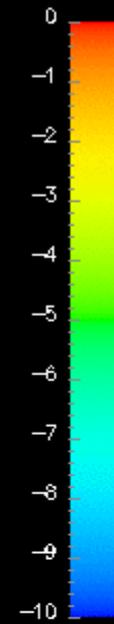
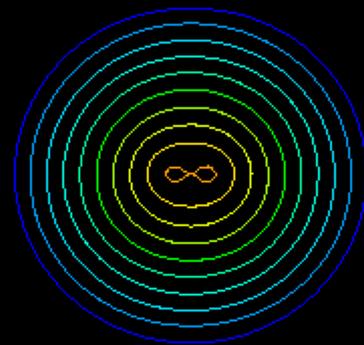
Imaginary part of the seventh eigen function
two-well Wood-Saxon potential

Solid-state physics

- Thornton, Eguiluz and Harrison (UT)
 - NSF OCI-0904972:
Computational chemistry and physics beyond the petascale
- Full band structure with LDA and HF for periodic systems
- In development: hybrid functionals, response theory, post-DFT methods such as GW and model many-body Hamiltonians via Wannier functions



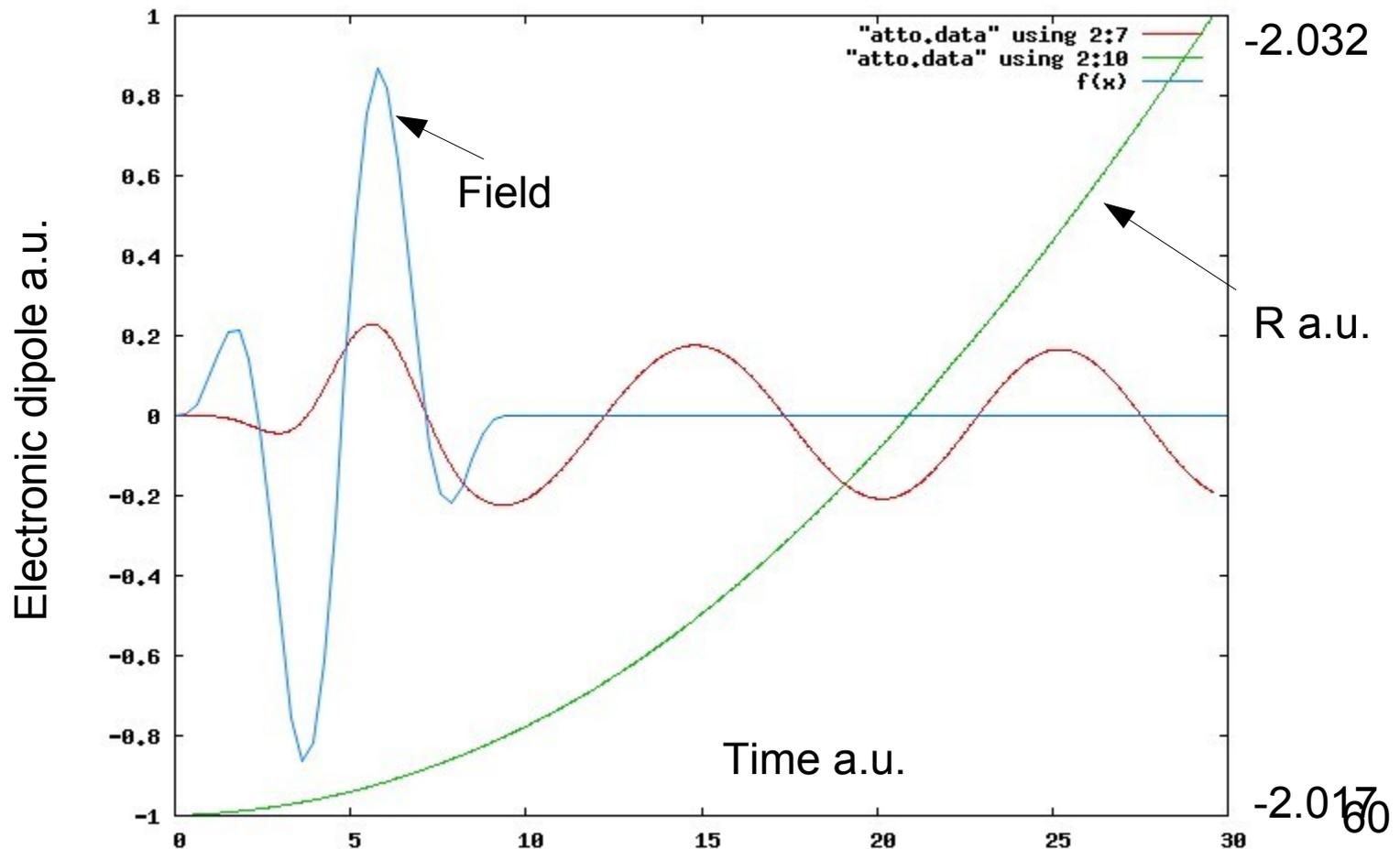
Coulomb potential isosurface in LiF



H₂⁺
molecule in
laser field
(fixed
nuclei)

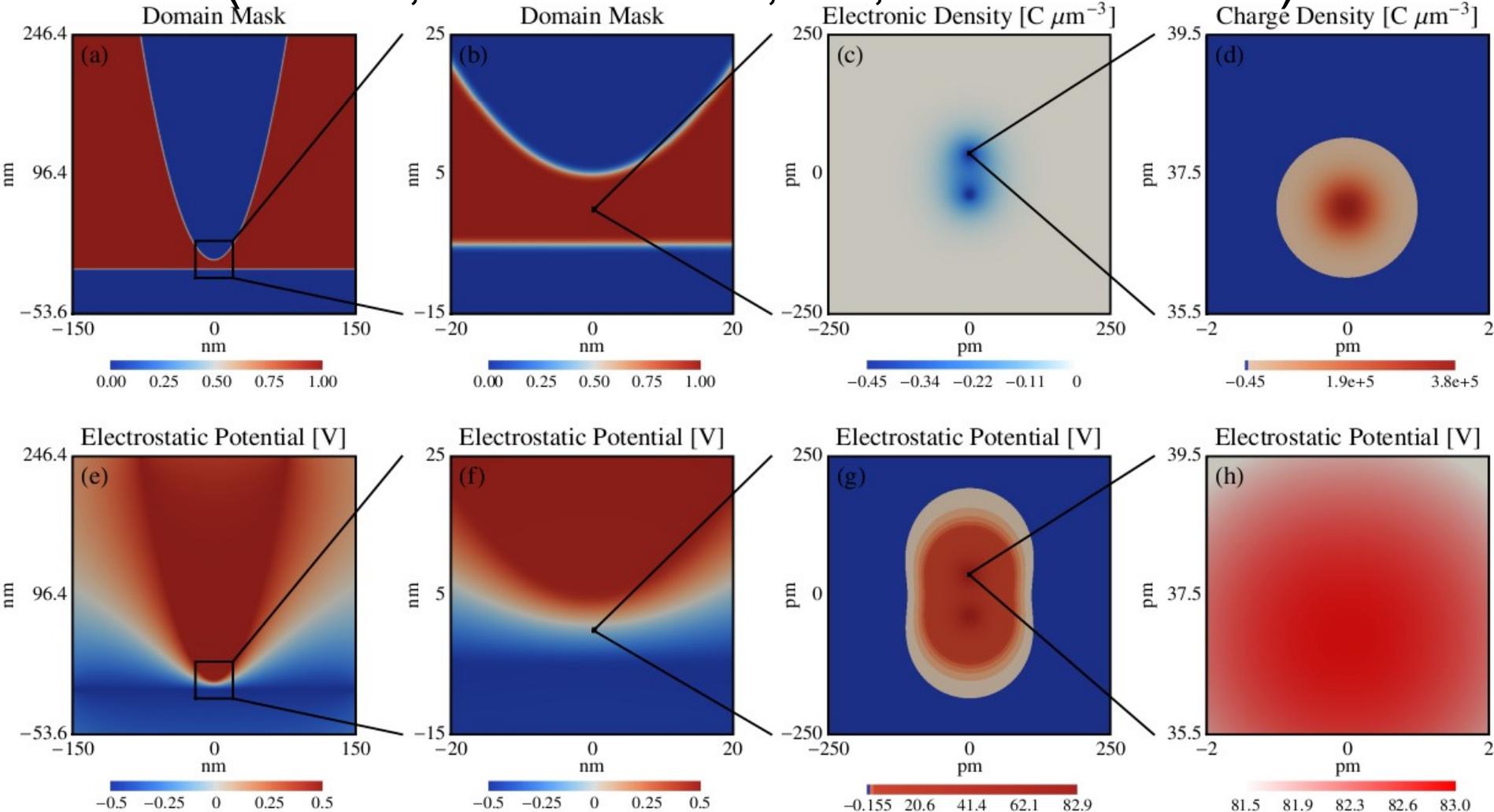
Dynamics of H_2^+ in laser

- 4D – 3 electronic + internuclear coordinate
 - First simulation with quantum nuclei and non-collinear field (field below is transverse)



Nanoscale photonics

(Reuter, Northwestern; Hill, Harrison ORNL)



Diffuse domain approximation for interior boundary value problem ; long-wavelength Maxwell equations; Poisson equation; Micron-scale Au tip 2 nm above Si surface with H₂ molecule in gap – 10^7 difference between shortest and longest length scales.

Runtime Objectives

- Scalability to 1+M processors ASAP
- Runtime responsible for
 - scheduling and placement, managing data dependencies, hiding latency, and medium to coarse grain concurrency
- Compatible with existing models
 - MPI, Global Arrays
- Borrow successful concepts from Cilk, Charm++, Python
- Anticipating next gen. languages

Key elements

- Futures for hiding latency and automating dependency management
- Global names and name spaces
- Non-process centric computing
 - One-sided messaging between objects
 - Retain place=process for MPI/GA legacy
- Dynamic load balancing
 - Data redistribution, work stealing, randomization

Futures

- Result of an asynchronous computation
 - Cilk, Java, HPCLs
- Hide latency due to communication or computation
- Management of dependencies
 - Via callbacks

```
int f(int arg);
ProcessId me, p;

Future<int> r0=task(p, f, 0);
Future<int> r1=task(me, f, r0);

// Work until need result

cout << r0 << r1 << endl;
```

Process “me” spawns a new task in process “p” to execute $f(0)$ with the result eventually returned as the value of future $r0$. This is used as the argument of a second task whose execution is deferred until its argument is assigned. Tasks and futures can register multiple local or remote callbacks to express complex and dynamic dependencies.

Global Names

- Objects with global names with different state in each process
 - C.f. shared[threads] in UPC; co-Array
- Non-collective constructor; deferred destructor
 - Eliminates synchronization

```
class A : public WorldObject<A>{  
    int f(int);  
};  
ProcessID p;  
A a;  
Future<int> b = a.task(p, &A::f, 0);
```

A task is sent to the instance of a in process p. If this has not yet been constructed the message is stored in a pending queue. Destruction of a global object is deferred until the next user synchronization point.

Global Namespaces

- Specialize global names to containers

- Hash table done
- Arrays, etc., planned

- Replace global pointer (process+local pointer) with more powerful concept

- User definable map from keys to “owner” process

```
class Index; // Hashable
class Value {
    double f(int);
};
```

```
WorldContainer<Index, Value> c;
Index i, j; Value v;
c.insert(i, v);
Future<double> r =
    c.task(j, &Value::f, 666);
```

A container is created mapping indices to values.

A value is inserted into the container.

A task is spawned in the process owning key j to invoke $c[j].f(666)$.

Summary

- Huge computational resources are rushing towards us
 - Tremendous scientific potential
 - Tremendous challenges in
 - Research,
 - Education, and
 - Community
- We need radical changes in how we compose scientific S/W
- UT and ORNL
 - Think of us if you have some good students looking for challenging graduate or postdoctoral study



Time evolution

- Multiwavelet basis not optimal
 - Not strongly band limited
 - Explicit methods very unstable
(DG introduces flux limiters, we use filters)
- Semi-group approach
 - Split into linear and non-linear parts

$$\dot{u}(x, t) = \hat{L}u + N(u, t)$$

$$u(x, t) = e^{\hat{L}t} u(x, 0) + \int_0^t e^{\hat{L}(t-\tau)} N(u, \tau) d\tau$$

- Trotter-Suzuki methods

$$e^{A+B} = e^{A/2} e^B e^{A/2} + O(\|[[[A, B], A] \dots]\|)$$

- Time-ordered exponentials
- Chin-Chen gradient correction (JCP 114, 7338, 2001)

Exponential propagator

- Imaginary time Schrodinger equation
 - Propagator is just the heat kernel

$$\left(-\frac{1}{2} \nabla^2 + V(x) \right) \psi(x, t) = \dot{\psi}(x, t)$$

$$\psi(x, t) \simeq e^{\nabla^2 t/4} e^{-V t} e^{\nabla^2 t/4} \psi(x, 0)$$

$$e^{\nabla^2 t/2} f(x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}} f(y) dy$$

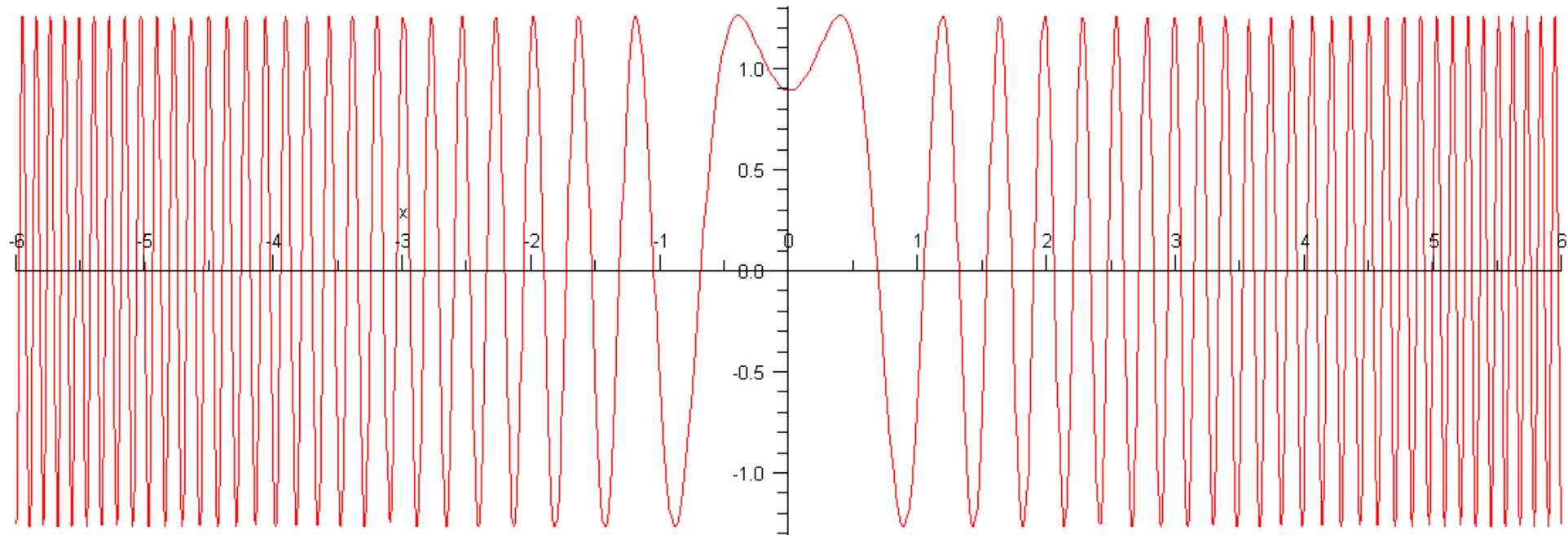
$$\lim_{t \rightarrow \infty} \psi(x, t) = \psi_0(x)$$

- Wrap in solver to accelerate convergence

Exponential propagator

- Free-particle propagator in real time

$$\psi(x, t) = e^{i\nabla^2 t/2} \psi(x, 0) = \frac{1}{\sqrt{2\pi i t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2it}} \psi(y, 0) dy$$

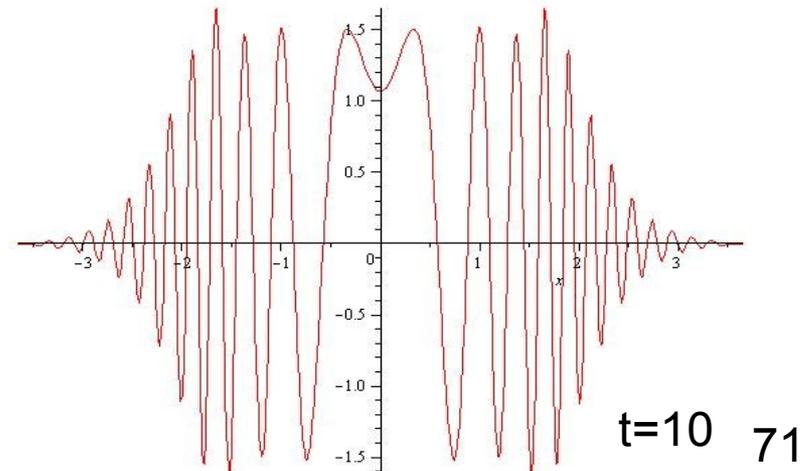
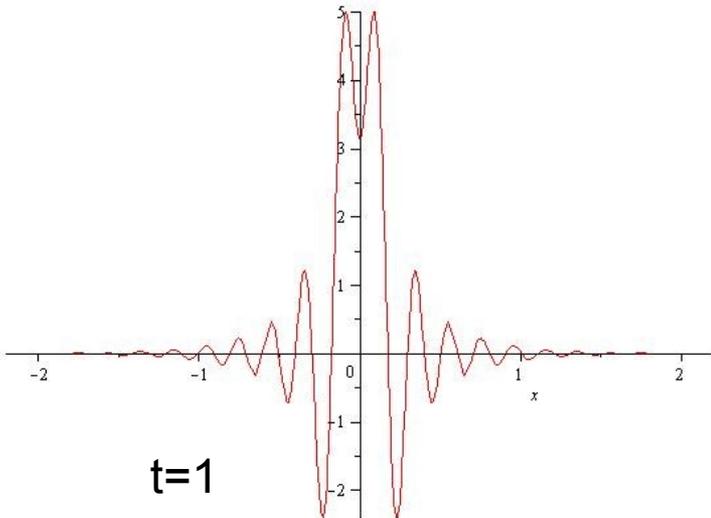


Exponential propagator

- Combine with projector onto band limit

$$\hat{G}_0(k, t, c) = e^{-i \frac{k^2 t}{2}} \left(1 + (k/c)^{30} \right)^{-1}$$

$$h = \frac{\pi}{c} \quad t_{crit} = \frac{2h^2}{\pi}$$



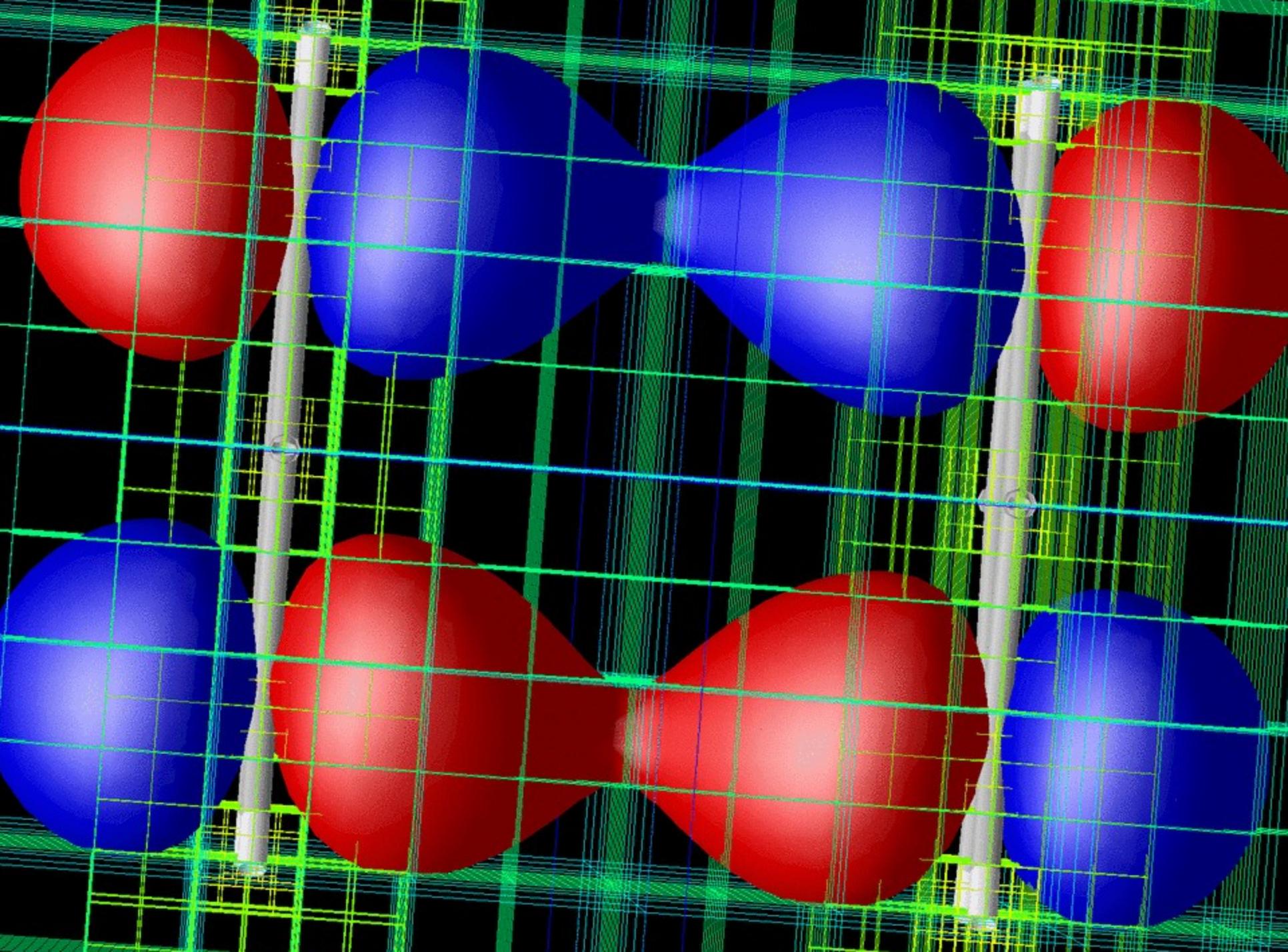
Path to linear scaling HF & DFT

- Need speed and precision
 - Absolute error cost $O(N \ln N / \epsilon)$
 - Relative error cost $O(N \ln 1 / \epsilon)$
- Coulomb potential
- HF exchange potential
- Orbital update
- Orthogonalization, localization, diagonalization
- Linear response properties

Summary

- MADNESS is a general purpose framework for scientific simulation
 - Conceived for the next (not the last) decade
 - Aims to makes scientific HPC more productive by reducing various sources of complexity
 - Deploys advanced numerical and C/S methods
- Multiple science applications at various levels of completeness

<http://code.google.com/p/m-a-d-n-e-s-s>



Separated form for integral operators

$$T * f = \int ds K(r-s) f(s)$$

- Approach

- Represent the kernel over a finite range as a sum of products of 1-D operators (often, not always, Gaussian)

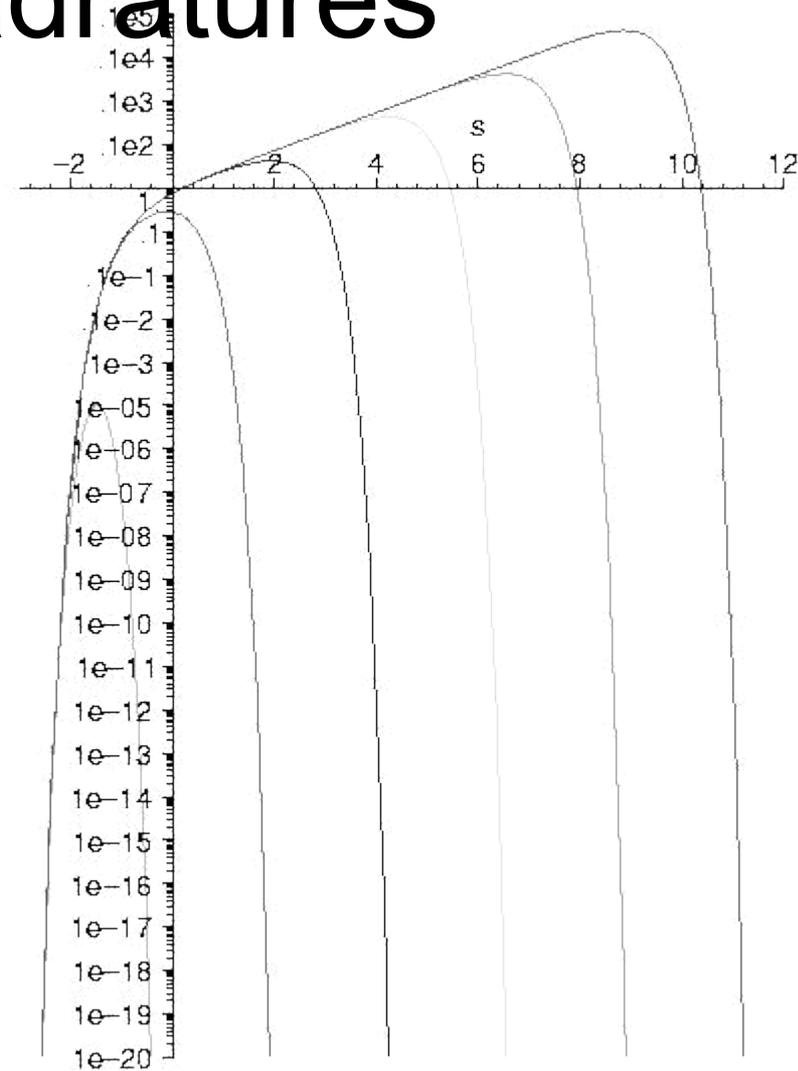
$$r_{ii',jj',kk'}^{n,l-l'} = \sum_{\mu=0}^M X_{ii'}^{n,l_x-l'} Y_{jj'}^{n,l_y-l'} Z_{kk'}^{n,l_z-l'} + O(\varepsilon)$$

- Only need compute 1D transition matrices (X,Y,Z)
- SVD the 1-D operators (low rank away from singularity)
- Apply most efficient choice of low/full rank 1-D operator
- Even better algorithms not yet implemented

Accurate Quadratures

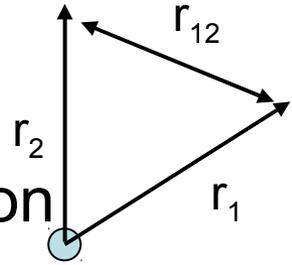
$$\begin{aligned} \frac{e^{-\mu r}}{r} &= \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-x^2 t^2 - \mu^2/4t^2} dt \\ &= \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2 e^{2s} - \mu^2 e^{-2s}/4 + s} ds \end{aligned}$$

- Trapezoidal quadrature
 - Geometric precision for periodic functions with sufficient smoothness
- Beylkin & Monzon
 - Further reductions, but not yet automated



The kernel for $x=1e-4, 1e-3, 1e-2, 1e-1, 1e0$.

Electron correlation



- All defects in mean-field model are ascribed to electron correlation

- Singularities in Hamiltonian imply for a two-electron atom

$$\Psi(r_1, r_2, r_{12}) = 1 + \frac{1}{2} r_{12} + \dots \quad \text{as } r_{12} \rightarrow 0$$

- Include the inter-electron distance in the wavefunction

- E.g., Hylleraas 1938 wavefunction for He

$$\Psi(r_1, r_2, r_{12}) = \exp(-\xi(r_1 + r_2)) (1 + a r_{12} + \dots)$$

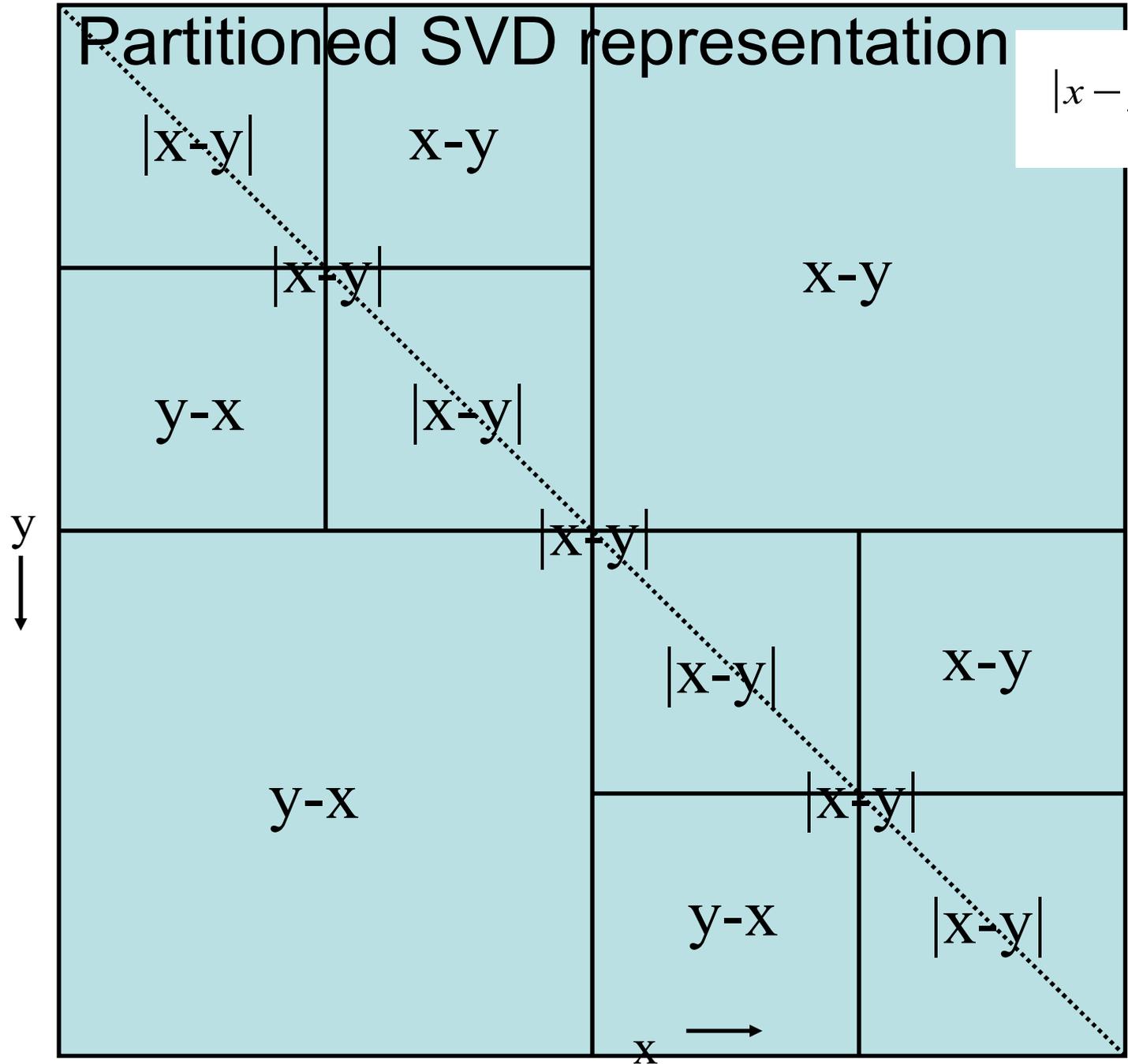
- Potentially very accurate, but not systematically improvable, and (until recently) not computationally feasible for many-electron systems

- Configuration interaction expansion – slowly convergent

$$\Psi(r_1, r_2, \dots) = \sum_i c_i \left| \phi_1^{(i)}(r_1) \phi_2^{(i)}(r_2) \dots \right|$$

Partitioned SVD representation

$$|x - y| = \sum_{\mu=1}^r f_{\mu}(x) g_{\mu}(y)$$



r = separation rank

In 3D, ideally must be one box removed from the diagonal

Diagonal box has full rank

Boxes touching diagonal (face, edge, or corner) have increasingly low rank

Away from diagonal $r = O(-\log \epsilon)$

Preliminary results for He atom (Yanai, 2005)

Computational details:

- 5-th order multiwavelets
- Wavelet threshold: 2×10^{-5}
- SVD threshold: 2×10^{-6}
- Exponential correlation factor

Perturbative wavefunction:

- Maximum refinement: $n=4$
- Memory: 132M in full partitioned SVD form
~10GB without SVD

	Variational E	ΔE	residual
HF	-2.861 61		
Iter. 0	-2.871 08		0.414 73
1	-2.894 92	-0.023 84	0.017 28
2	-2.900 43	-0.005 51	0.007 94
3	-2.902 18	-0.001 75	0.003 84
4	-2.902 88	-0.000 70	0.002 02
5	-2.903 20	-0.000 32	0.001 25
6	-2.903 39	-0.000 20	0.000 91
...
12	-2.903 73	-0.000 04	0.000 36
13	-2.903 73	+0.000 004	0.000 32
14	-2.903 77	-0.000 04	0.000 28
exact	-2.903 74 (E(HF)=-2.861 68)		Energy is variational (small non-variational is just truncation err)
Hylleraas (6 terms)	-2.903 24		
Löwdin and Redei	-2.895 4		
cc-pV6Z	-2.903 48 (FCI) (E(HF)= -2.861 67)		