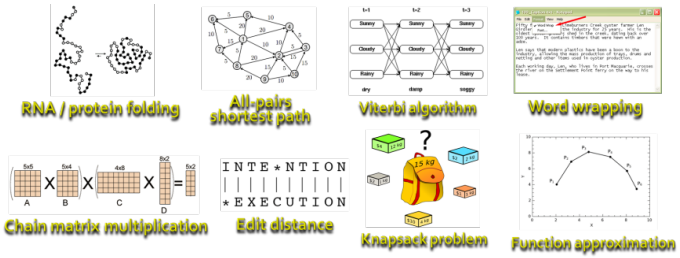
Automatic Discovery of Cache-oblivious Parallel Recursive Algorithms for Dynamic Programs

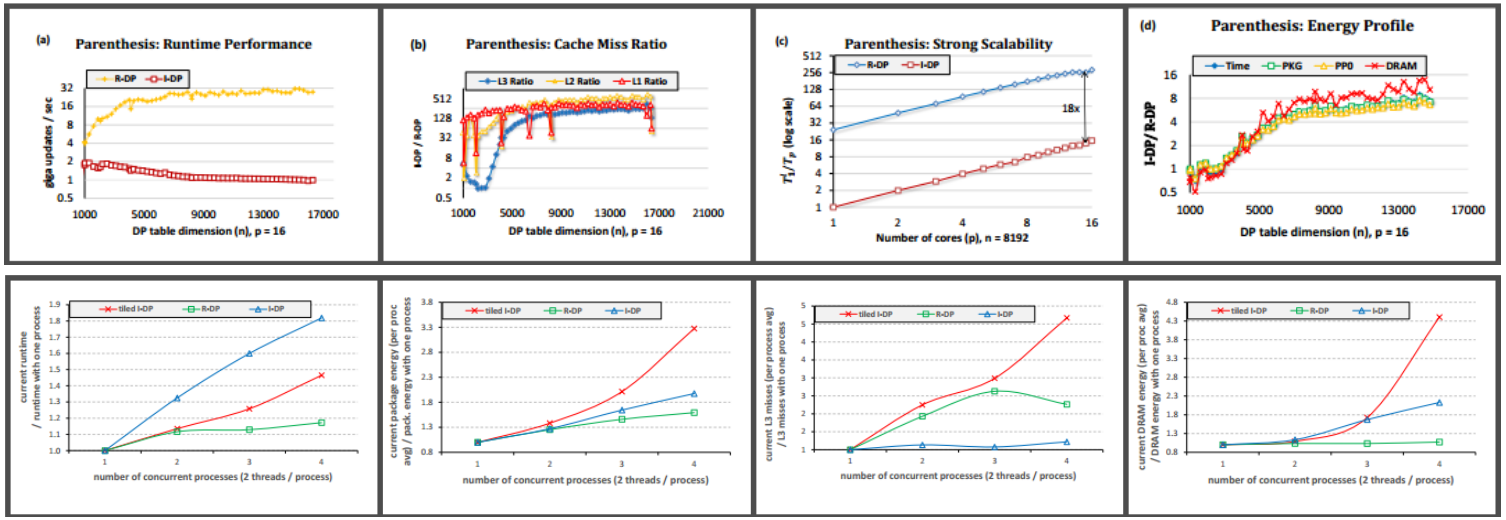
(with MIT and Fudan University)

Rezaul Chowdhury

- *Dynamic Programs (DP)*
  - arise in many application areas
  - traditionally implemented using inefficient nested loops

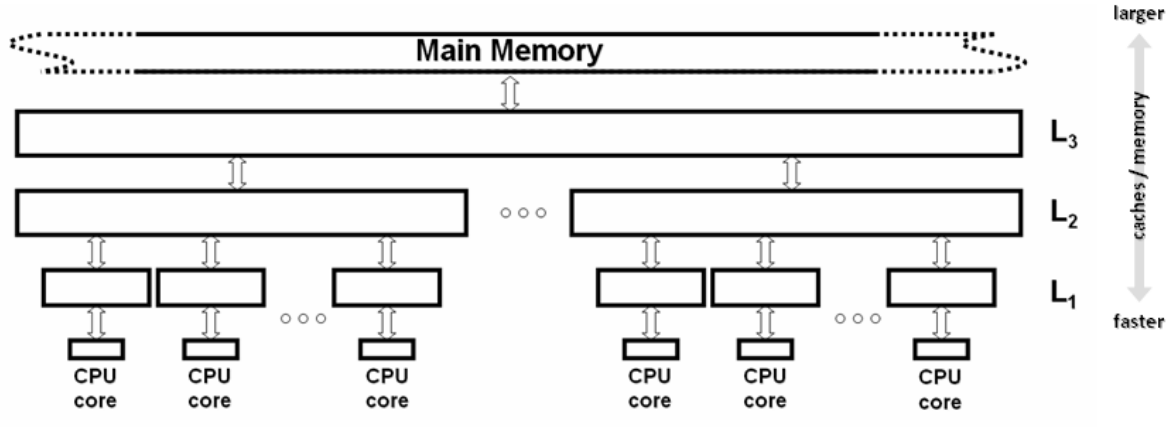


- Given an inefficient iterative DP implementation (I-DP), we automatically generate a *high-performing energy-efficient resource-oblivious parallel recursive algorithm* (R-DP) for solving the DP



Oblivious Algorithms for Multicores & Processor Networks

(with UT Austin and University of Padova)



- *Efficient* algorithms for multicores must address
  (1) caching issues, and (2) shared-memory parallelism

- Efficient algorithms *oblivious* of machine parameters (e.g., cores & caches)
  - can be written without worrying about machine configuration
  - can run efficiently on machines with a wide range of machine parameters
  - can simultaneously adapt to all levels of a multi-level cache hierarchy

- We introduce the notion of *multicore-oblivious* algorithms
  - show how to achieve multicore-obliviousness
  - present optimal multicore-oblivious algorithms for fundamental problems

Pochoir: Efficient Stencil Computations Made Easy

(with MIT and Fudan University)

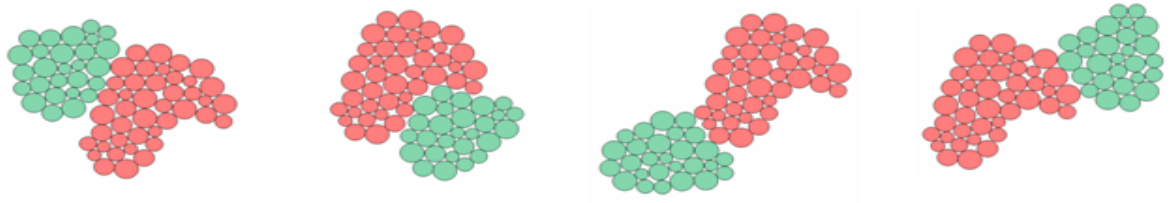
- Pochoir (pronounced "PO-shwar") is a compiler and run-time system for implementing stencil computations on multicores
- core- and cache-oblivious (optimizes private cache misses)
- Two major components
  - a stencil specification language embedded in C++
  - a source-to-source translator

    (specification → highly optimized parallel Cilk Plus code)

```
1.  Pochoir⟨ 2 ⟩ heat;
2.  Pochoir_Array⟨ double, 2, 2 ⟩ u( d1, d0 );
3.  heat.registerArray( u );
4.  Pochoir_Shape⟨ 2 ⟩ heat_shape[ ] = { { 1, 0, 0 }, { 0, 1, 0 }, { 0, -1, 0 }, { 0, -1, -1 }, { 0, 0, -1 }, { 0, 0, 1 } };
5.  heat.registerShape( heat_shape );
6.  Pochoir_Kernel_2D( heat_fn, t, x, y )
7.      u( t + 1, x, y ) = u( t, x, y ) + CX * ( u( t, x + 1, y ) – 2 * u( t, x, y ) + u( t, x – 1, y ) )
                                        + CY * ( u( t, x, y + 1 ) – 2 * u( t, x, y ) + u( t, x, y – 1 ) );
8.  Pochoir_Kernel_End
9.  heat.run( T, heat_fn );
```

F²Dock: Rigid-body Protein-Protein Docking

(with UT Austin and SCRIPPS Research)



- employs many novel ideas for ranking/filtering docked positions
  - outperforms other rigid-body docking software in accuracy
- first docking software to employ
  - non-uniform (error-bounded) & sparse uniform FFT
  - octree based tunable approximations (speed-accuracy tradeoff)
- parallelization
  - multithreaded
  - MPI-based distributed implementation
- front-end
  - graphical user interface
  - client-server mode (submits jobs to UT PRISM2 cluster by default)
- open source

Nuclear Astrophysics

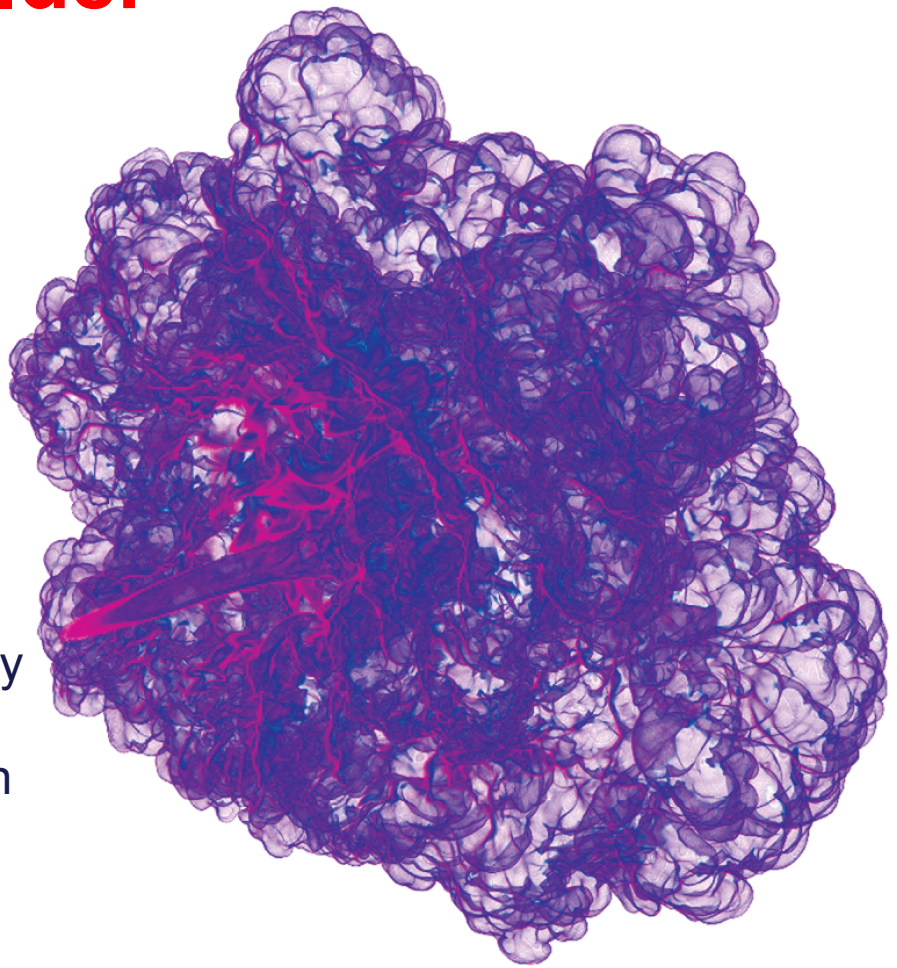
Alan Calder

Stellar Explosions:

- Thermonuclear (type Ia) and core collapse supernovae
- Classical novae
- Neutron stars and X-ray bursts.

Computational Science:

- Hydrodynamics and radiation hydrodynamics
- Verification, validation, and uncertainty quantification
- Basic Physics of turbulent combustion
- Computational Science Education



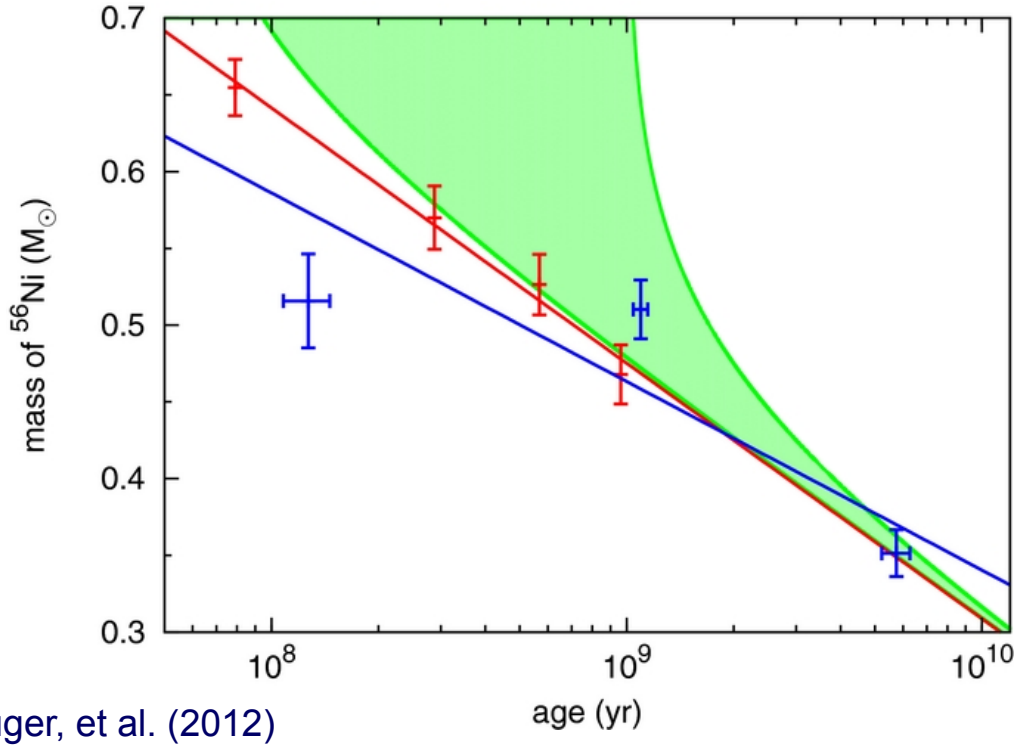
Type Ia Supernovae

Thermonuclear (Type Ia) Supernovae are bright explosions that serve as distance indicators for cosmological studies.

Research focuses on understand the mechanism of the explosion and determining systematic effects on the brightness and the intrinsic scatter.

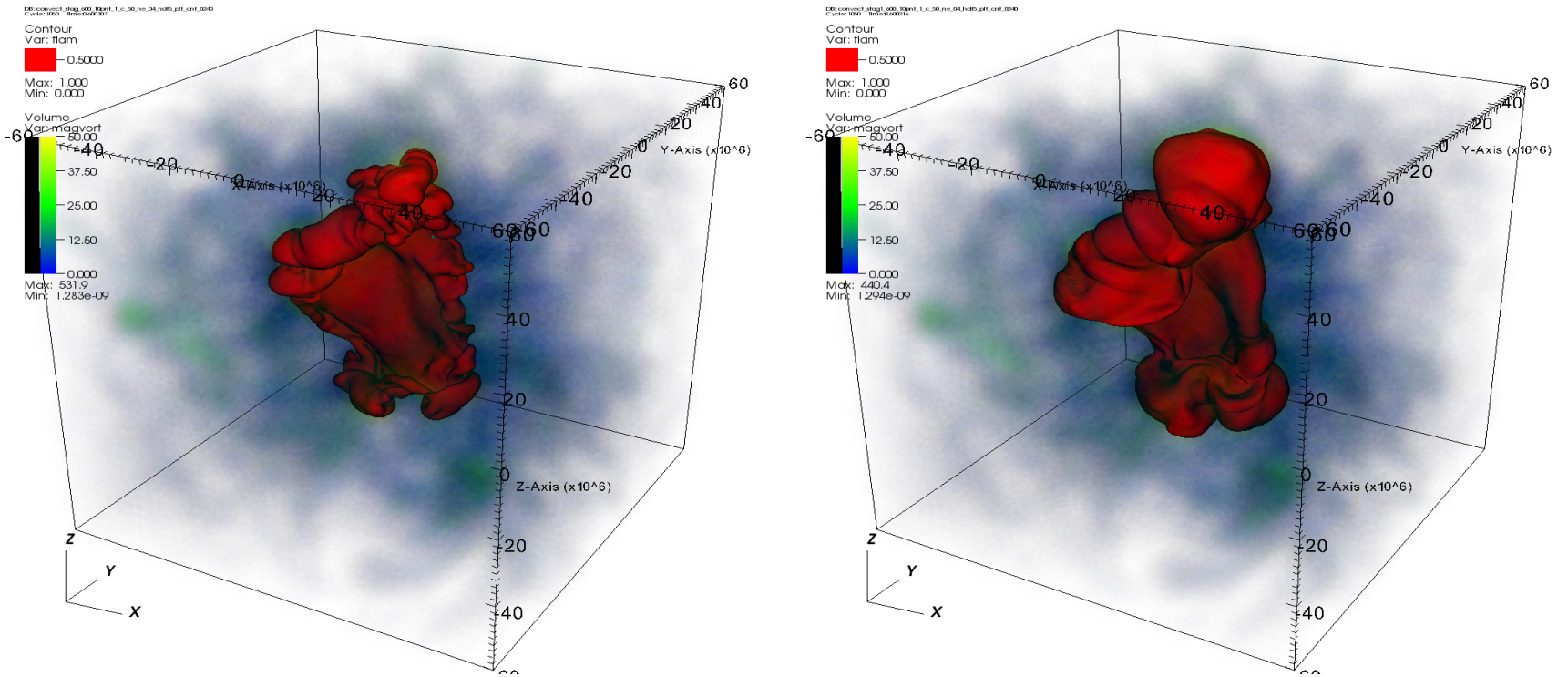
Figure: mass of radioactive nickel, the source of brightness, vs. age of progenitor (red points).

The study provided the first theoretical explanation for the observed trend of dimmer supernovae in older galaxies (blue points).



Kruger, et al. (2012)

Fundamental Physics: the interaction between a flame and turbulence



Burning enhanced by a sub-grid-scale turbulence-flame interaction model (right). An important ingredient of type Ia supernova models.

Computer Simulations of clouds and climate

Marat Khairoutdinov

Paleoclimate simulations using advanced climate model

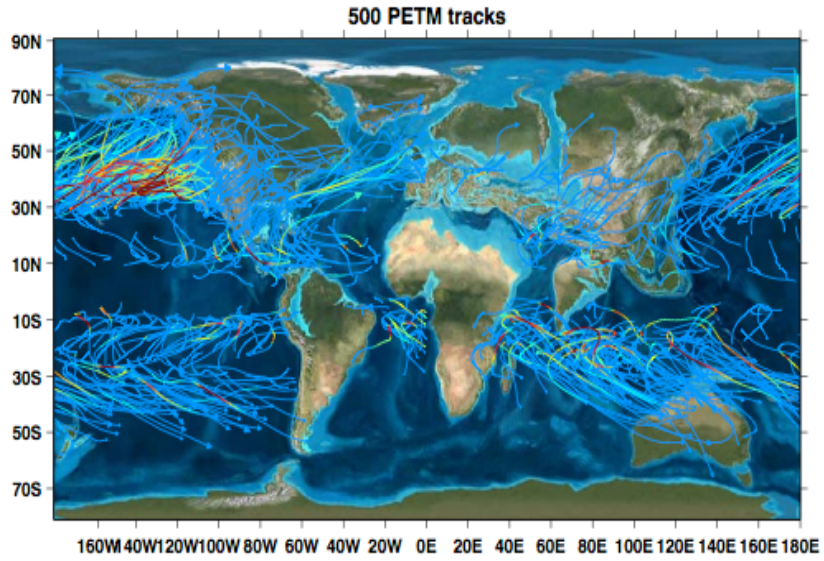
According to paleo reconstructions, it has been tens of millions of years since the Earth had the levels of CO_2 and corresponding radiative forcing that we may experience in just 100 years from now.

Perhaps we can use the past to tell us what is awaiting us in the future ...

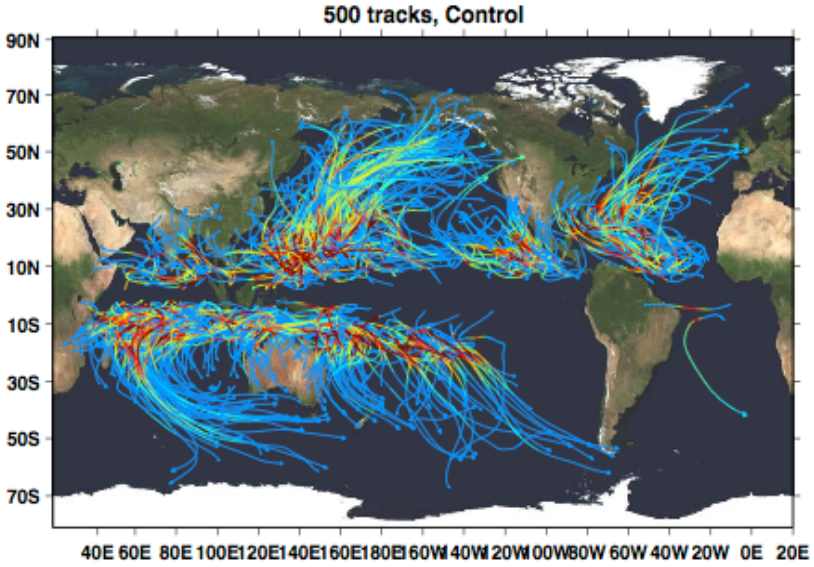
Unlike simulations of the future, there are observational constraints on simulations of the past...

Earth 55 million years ago during Paleocene-Eocene Thermal Maximum (PETM)
Simulated possible tracks of hurricanes

55 million years ago



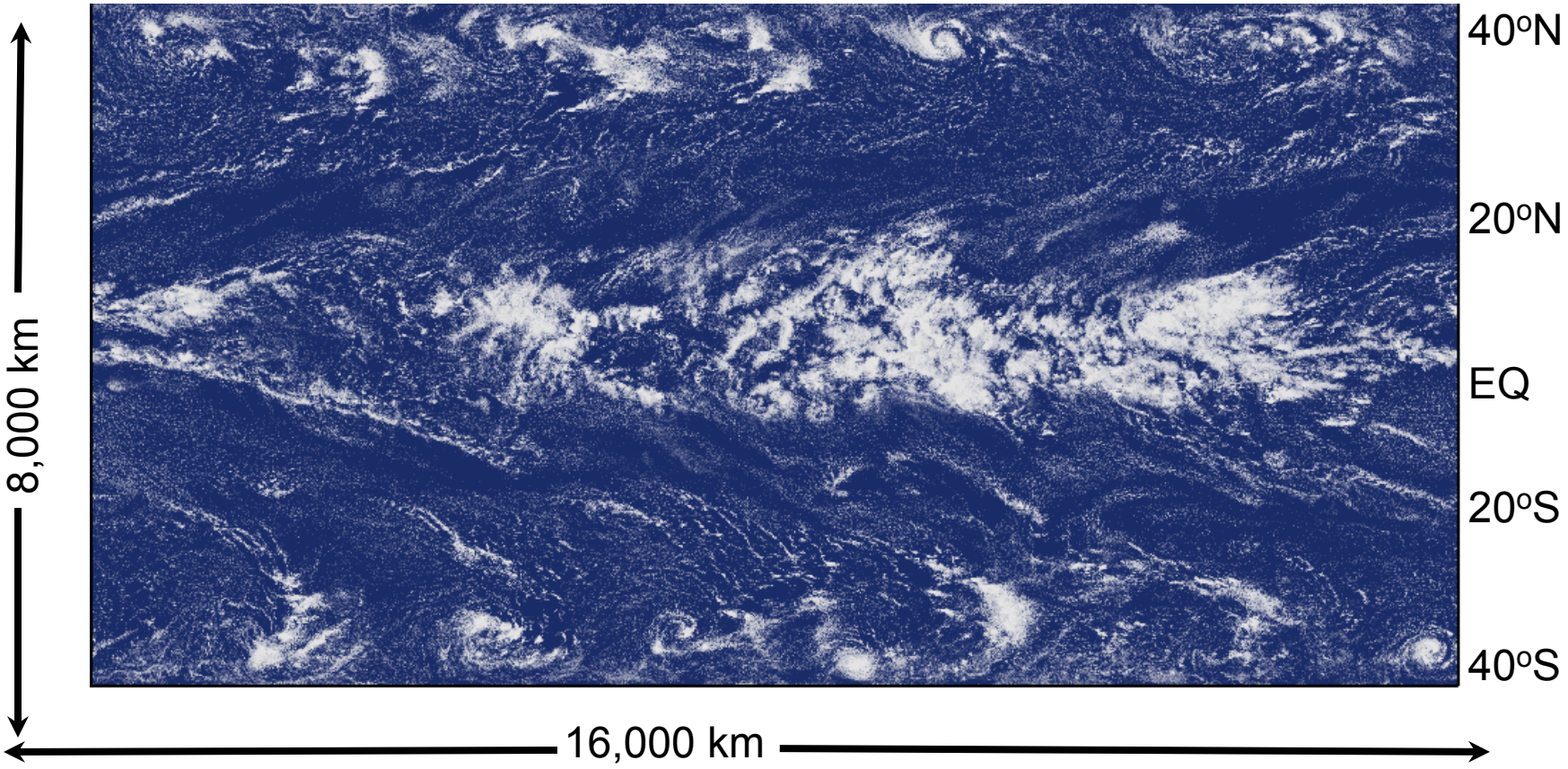
Today



Cloud-resolving simulation of tropical weather systems

Simulations like that help us to understand how climate regulates itself, how cloud systems organize on large scales, and what can happen in the warmer world.

Simulated view of a cloud field in Tropics as would be seen from a satellite



Each pixel of this image represents 4x4 grid cells of the numerical grid. A 100 day-long simulation takes about one month of nonstop computations using 2,048 processors of the IBM BlueGene/L supercomputer. It would take more than 150 years for a home desktop PC to produce such a simulation.

Large-eddy simulation of storms

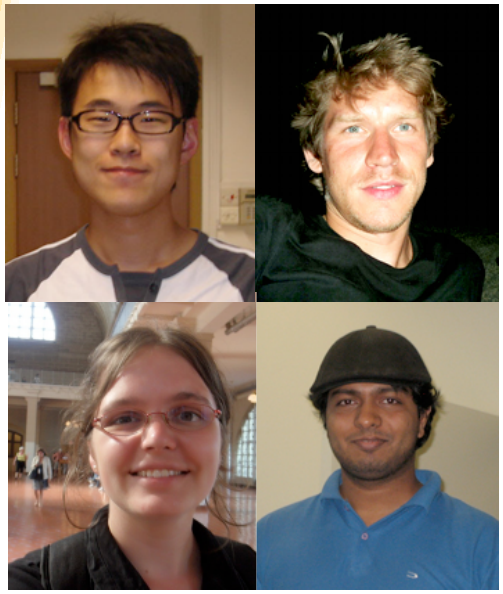
High-resolution simulations of convective storms (thunderstorms) help to improve representation of clouds in weather-prediction and climate models.

Visualization by Ian Glenn, U Utah

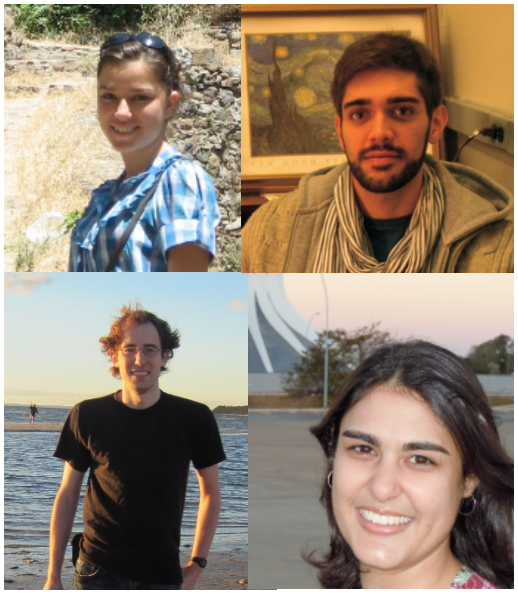
Visualization of a precipitating convective cloud system in the Tropics simulated by the System of Atmospheric Modeling (SAM). The model solves 3D Navier-Stokes equations coupled with the physical models of cloud microphysics and radiative transfer. More than one billion grid cells were used. Note, only 1/10th of the actual domain width is shown. The image is produced using 3D radiative transfer algorithm.

Marivi Fernandez-Serra

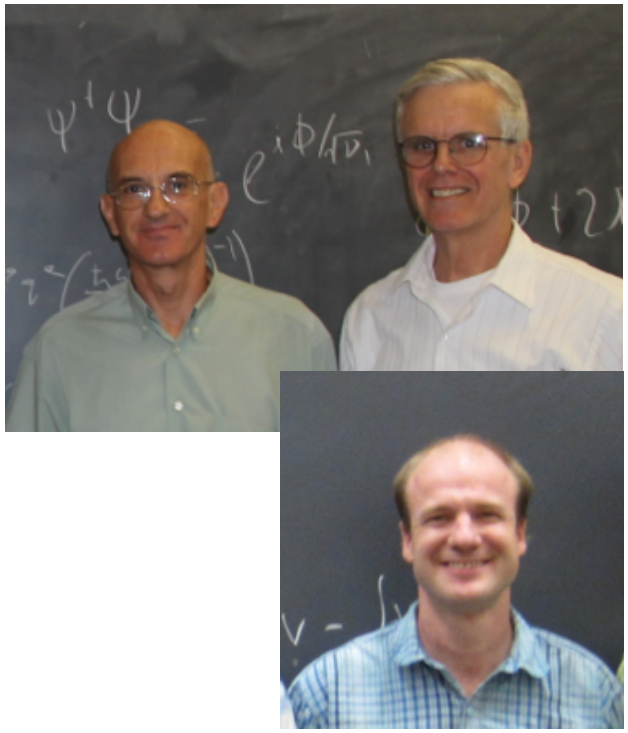
Past Members



Current Members



Close Collaborators



Madrid: Rafa Ramirez
San Sebastian: Emilio Artacho, Fabiano Corsetti
SIESTA team

DOE Early Career DE-SC0003871
DOE: DE-FG02-09ER16052

siesta

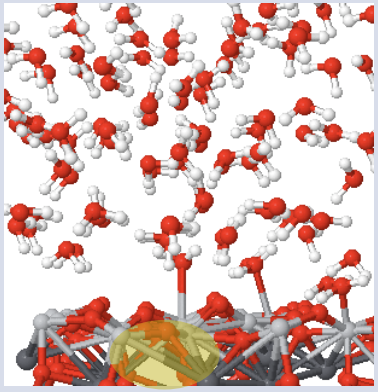
iACS INSTITUTE FOR ADVANCED COMPUTATIONAL SCIENCE

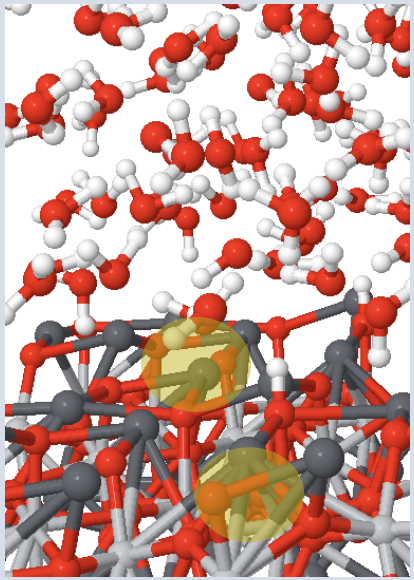
Water/interfaces by first principles

Ferroelectric Surfaces:
Investigating the role of polarity on water/substrate interactions

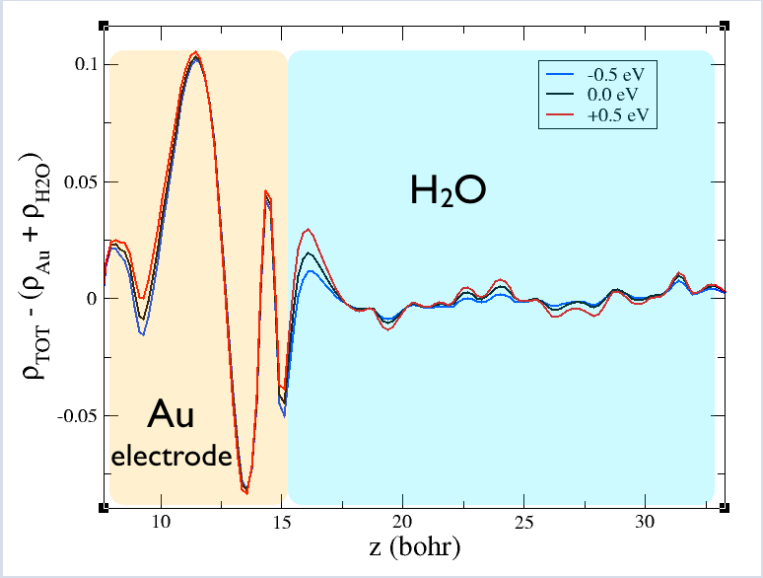
PbO Surfaces
dissociation+OH^- in solution

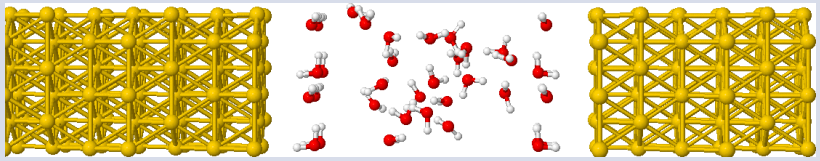
TiO_2 Surfaces
No dissociation





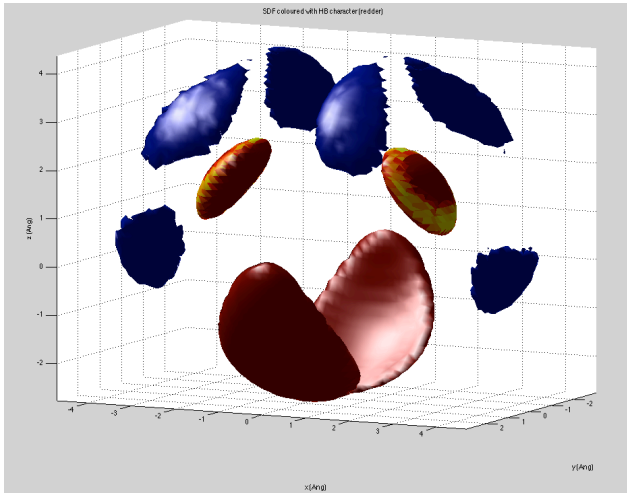
Electrochemical interface:
metal/water under applied bias coupling non equilibrium transport methods with molecular dynamics



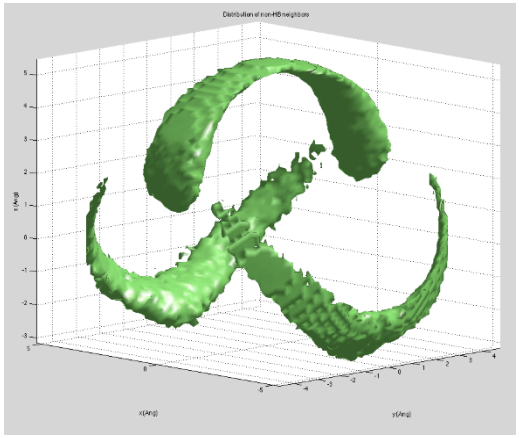


Presence of Interstitial, antitetrahedral molecules necessary to increase diffusivity

1st and second Hbond neighbors



vdW neighbors



Standard DFT water largely underestimates the density.

van der Waals interactions increase the density (too much), and improve compressibility but most importantly, they point to the source of the problem: vdW bonds are necessary to accurately reproduce the structure and dynamics of liquid water.

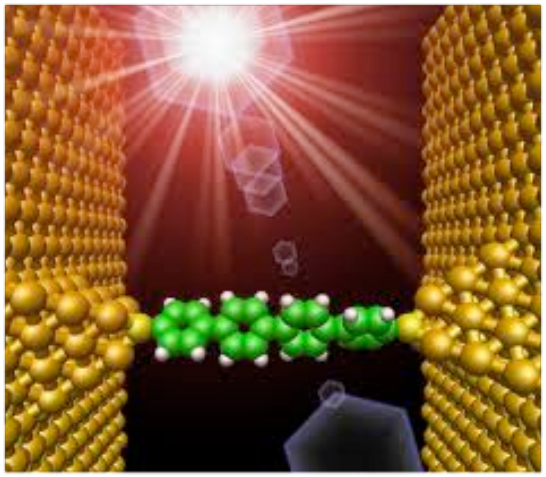
While neither of current vdW density functionals is perfect, at least we can already simulate liquid water at ambient temperature.

Newly awarded grant: DMREF: High-Throughput Mapping of Functional Dielectric/Metallic Heterostructures.

Matthew Reuter

Electron transport through molecules:

- How does electric current traverse a **quantum** system?
- What is the conductance of a single molecule?
- What physics determines this behavior?
- What effects lack classical analogs?



Applications include:

- Photovoltaics
- Scanning probe microscopies
- Molecular electronics

Our goals:

- Provide better interpretations of experimental data
- Develop & implement more accurate computational frameworks

Ratner, Nature Nanotech. **8**, 378 (2013).

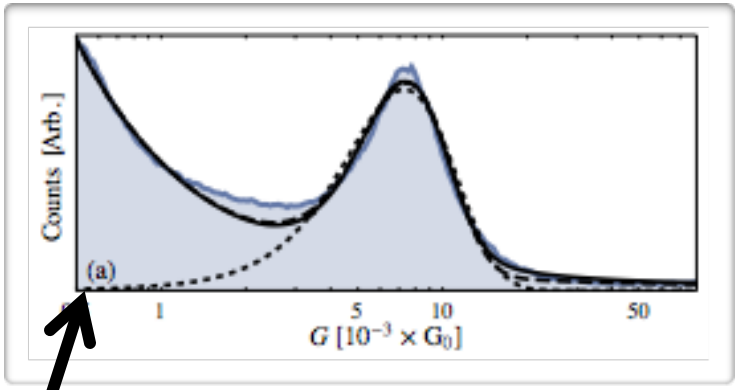
Interpreting Experiments

Common experiments are statistical

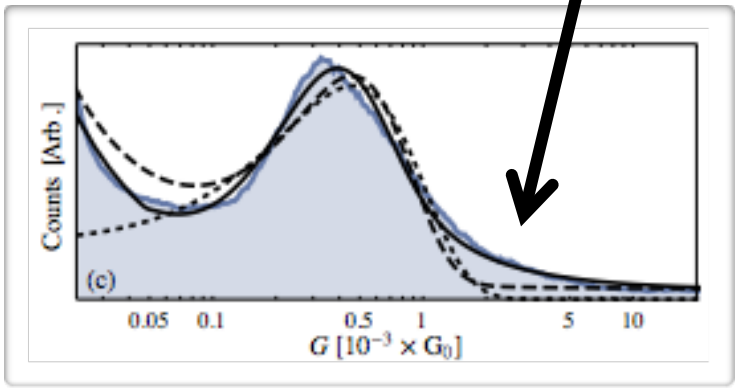
- High variability from one measurement to the next
- Lots of measured conductances reported as a histogram

Our work:

- Examine the statistics. What does the line shape tell us about the underlying molecular behavior?
- Develop computational tools to qualitatively simulate histograms; quantitatively extract information from experimental histograms



Solid line: our best model (so far)



Experimental data courtesy of the Venkataraman group.
Quan, Pitler, Ratner, & Reuter. *In preparation* (2015).

Improving Computation

Ab initio simulations are common

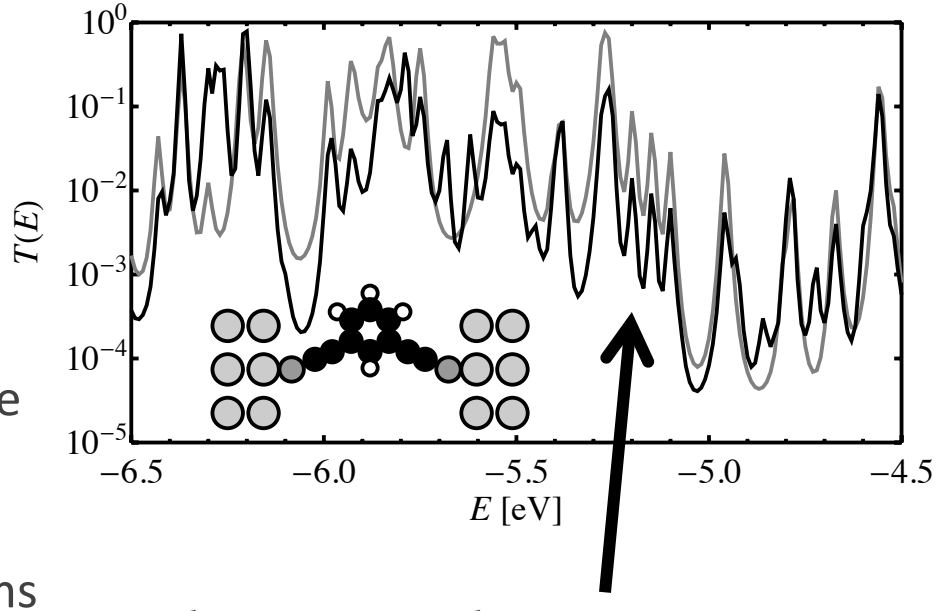
- Several known "white elephant" problems:
- Numerical artifacts
- Unphysical behavior (e.g., ghost transmission)

→ Transmissions are usually too large

Our work:

- Diagnose causes for these problems (e.g., poor system partitioning)
- Implement computational tools that are not plagued by these problems

→ More accurate simulations!



Preliminary results:
Gray line — Common formalism
Black line — Our tools

Although these issues sound pedantic, they have big effects!

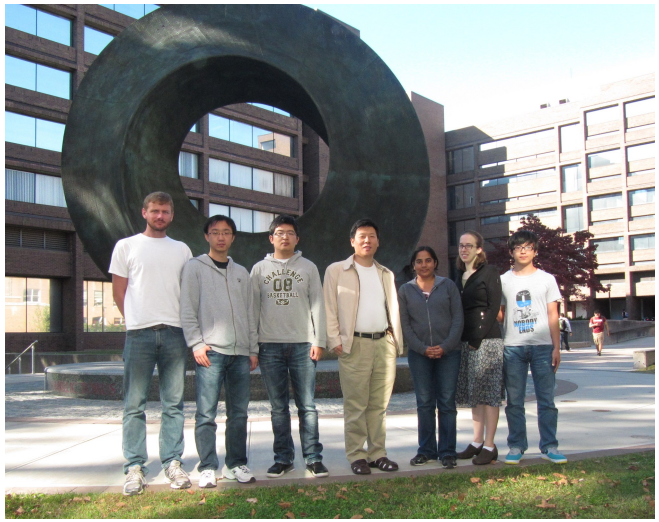
Reuter & Harrison. J. Chem. Phys. **139**, 114104 (2013); **140**, 177104 (2014).

Jim Jiao

NumGeom Group in Computational & Applied Mathematics

NumGeom group focuses on high-performance numerical and geometric computations

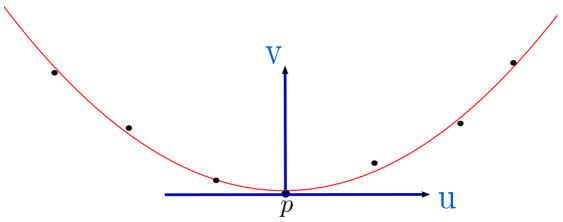
- **Numerical methods**: accurate and stable methods for general approximations or solving PDEs
- **Geometric algorithms**: methods for dynamic surfaces; data structures and algorithms for meshing; interfaces in multiphysics coupling
- **HPC**: efficient and scalable multigrid solvers; high-productivity programming environment



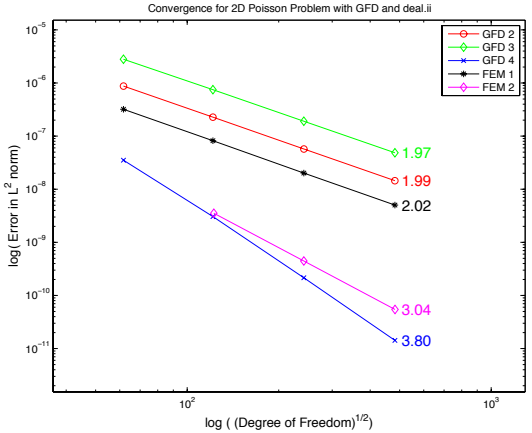
From left to right: Tristan Delaney, Hongxu Liu, Cao Lu, Prof. Xiangmin Jiao, Navamita Ray, Rebecca Conley, and Xinglin Zhao

Highlights 1: Unified Theoretical Framework of Numerical Methods

- **WLS**: Weighted least squares provides more flexible framework than interpolation for accurate and stable methods over point clouds or unstructured meshes
- **GFD**: Based on WLS, GFD generalizes finite-difference methods to unstructured meshes, delivering higher-order accuracy and stability (student participants: Hongxu Liu, Rebecca Conley, et al.)



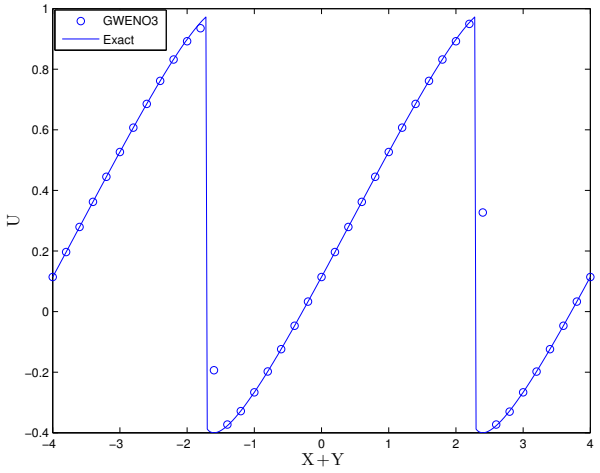
WLS generalizes interpolation, with more flexibility and better stability.



GFD delivers higher order convergence than other state-of-art methods.

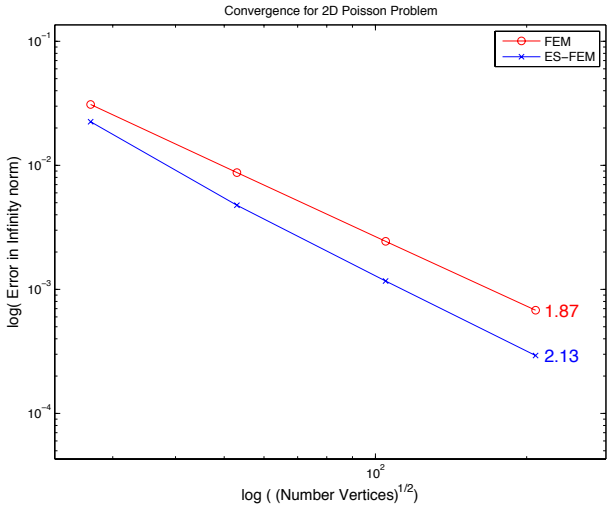
Highlights 2: Advanced Numerical Methods for Complex Geometries

- **LS-WENO**: WLS-based generalization of WENO (weighted essentially non-oscillatory) schemes to solver hyperbolic PDEs over unstructured meshes (student: H. Liu)



Example with LS-WENO for Berger's equation on tetrahedral meshes

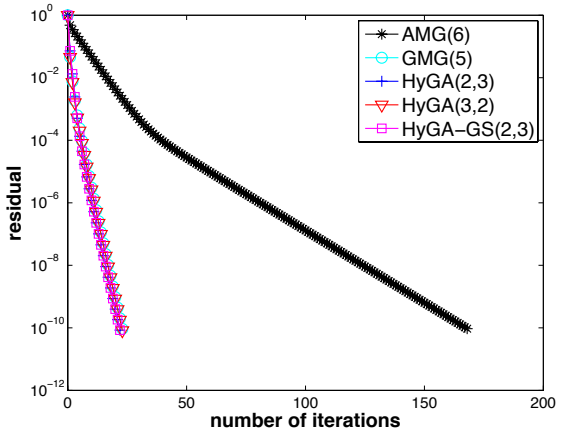
- **AES-FEM**: WLS-based generalization of finite-element methods for more accurate and stable solutions even on poor-quality meshes (student participants: Tristan Delaney and R. Conley)



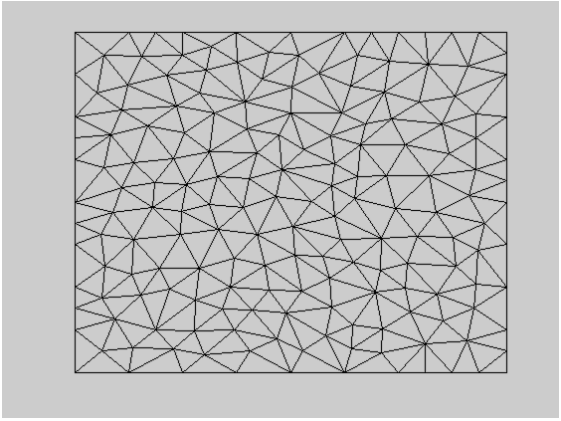
AES-FEM delivers higher accuracy than standard FEM

Highlights 3: Hierarchical Meshes, Adaptive Mesh Refinement, and Multigrid

- **AHF**: Array-based half-facet data structure for unstructured meshes integrated into MOAB at Argonne National Lab (students: Xinglin Zhao et al.)
- **HyGA**: Hybrid Geometric+Algebraic multigrid method for efficient solution of linear systems from GFD and FEM discretizations over hierarchical meshes (student: Cao Lu)
- **AMR**: Data structures, algorithms, and theoretical analysis of FEM with adaptive mesh refinement (students: X. Zhao & C. Lu)



HyGA accelerates algebraic multigrid



Animation of AMR

Rich-Get-Richer in Crowdfunding

Arnout van de Rijt

Projects on www.kickstarter.com

random assignment

Experimental condition

Raised from others:

$294

Control condition

Raised from others:

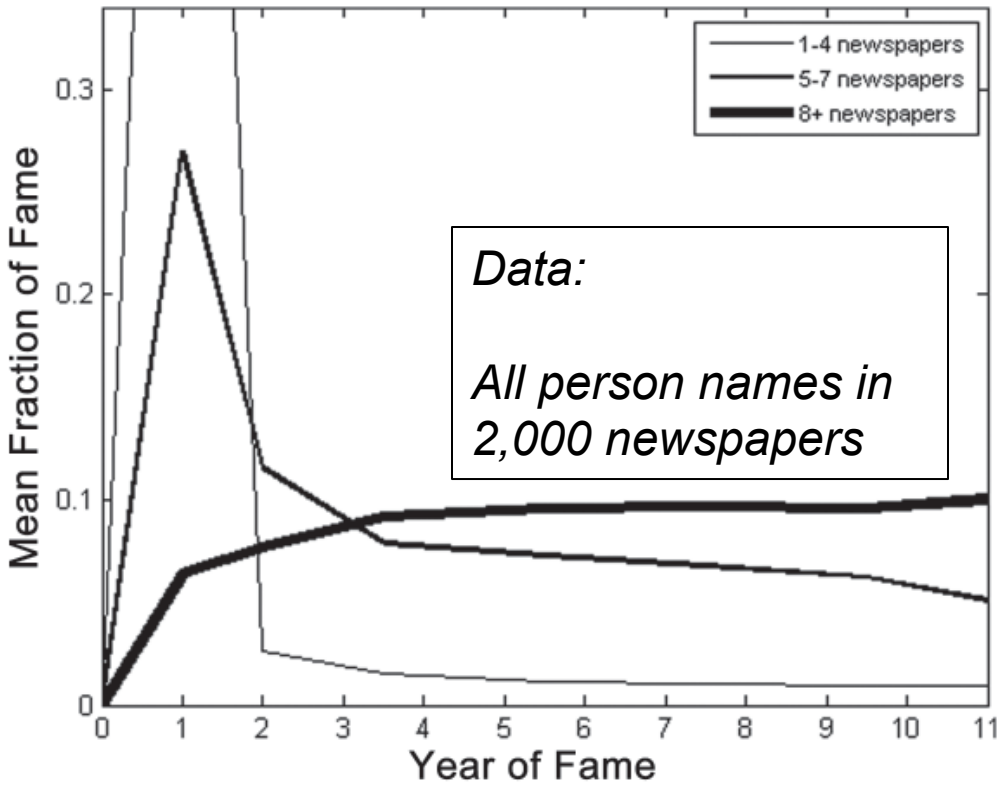
$103

Source: *Van de Rijt et al. in **PNAS** (2014)*

Funding: ***NSF** grants SES-1340122*

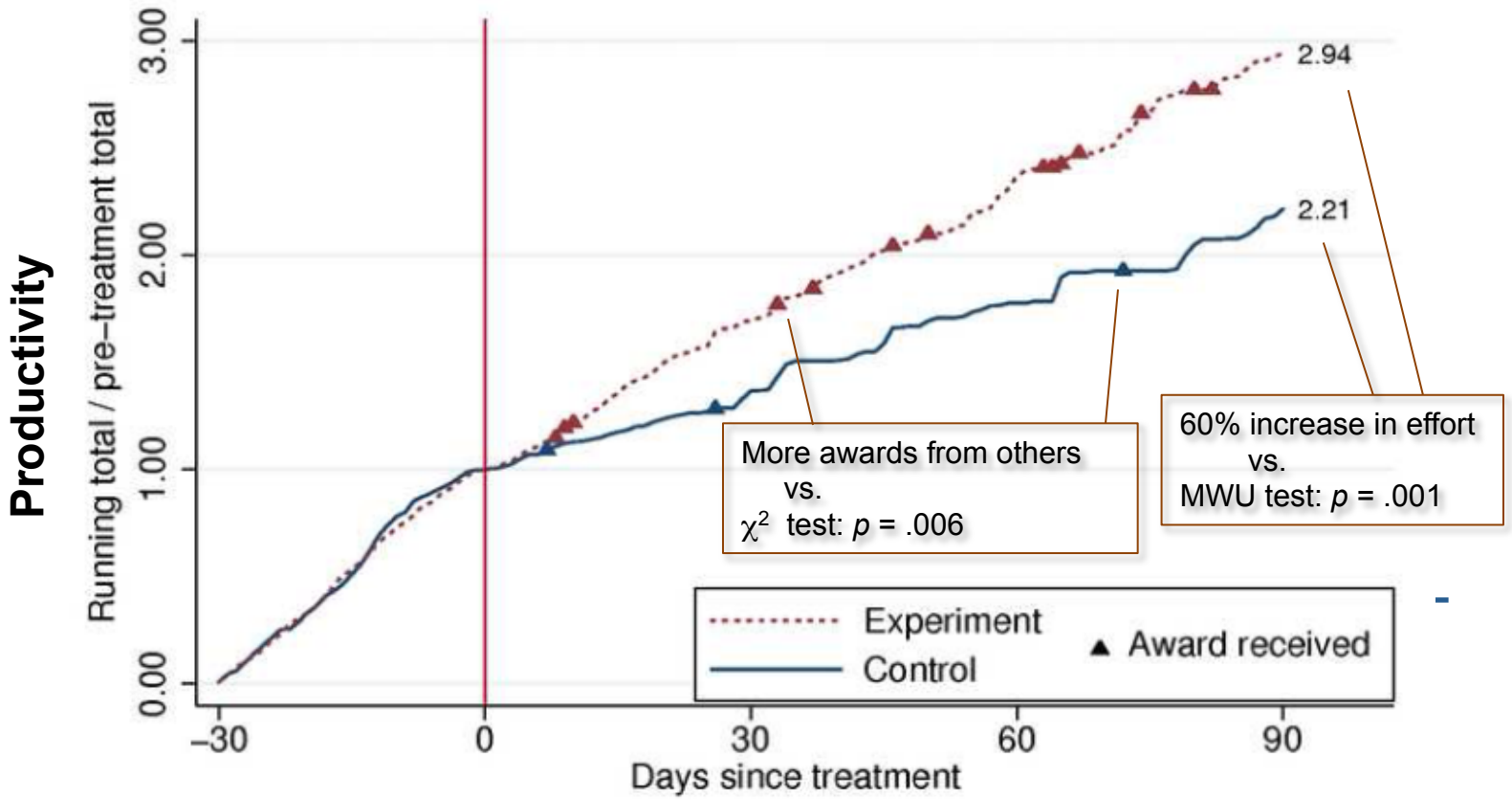
Press: *Economist, Time Magazine, National Geographic, WAMC*

Only Fifteen Minutes?



Source: *Van de Rijt et al. in **American Sociological Review** (2013)*
Press: *LA Times, NBC News, Toronto Star, Globe and Mail, Yahoo News, Pacific Standard, Politiken (front page)*

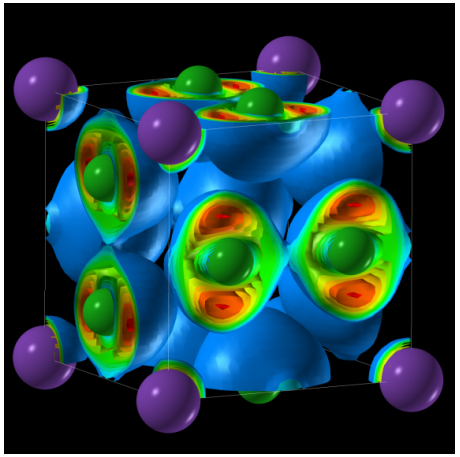
Wikipedia Volunteer Editors Increase Effort after Virtual Awards Given by Researchers



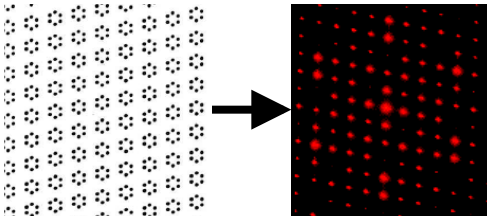
Source: *Restivo & Van de Rijt in **PLOS ONE** (2012)*

Predicting “forbidden” chemistry and novel materials with the USPEX method/code

Artem R. Oganov

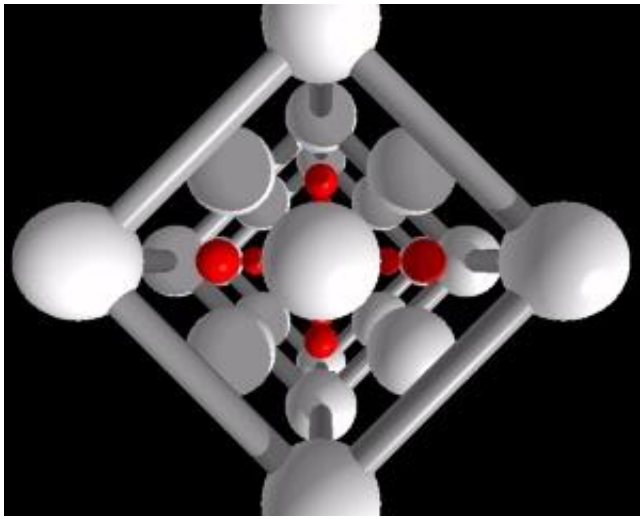


Can Periodic Systems change at extreme conditions?
What is the chemical formula of sodium chloride?
What is the most inert element?
What is the cleanest fuel material?
Why does some dust cause lung cancer?



Structure **Diffraction**

Zincblende ZnS.
One of the first structures solved by Braggs in 1913.



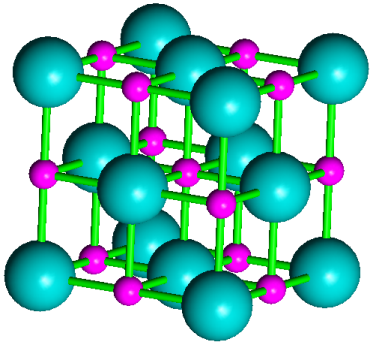
Example: “Crazy” sodium chlorides

Na-Cl

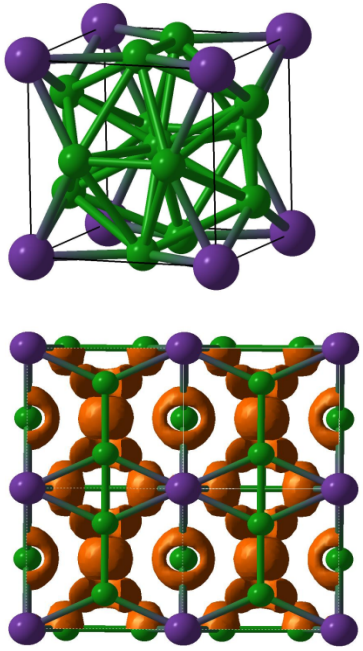
Salt as we know it:

Peculiar Na-Cl compounds:

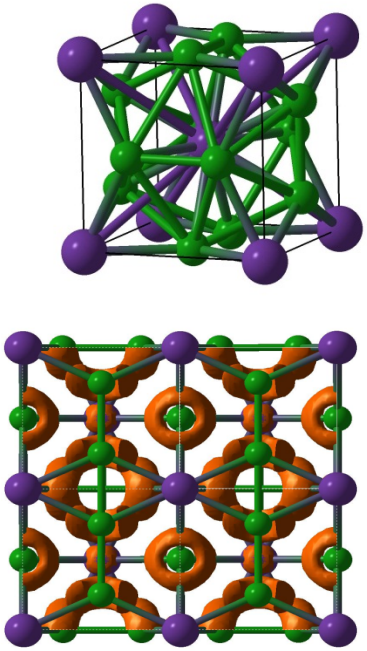
- NaCl7: some Cl atoms have POSITIVE Bader charge (+0.07).
- Na3Cl: 2D-metal



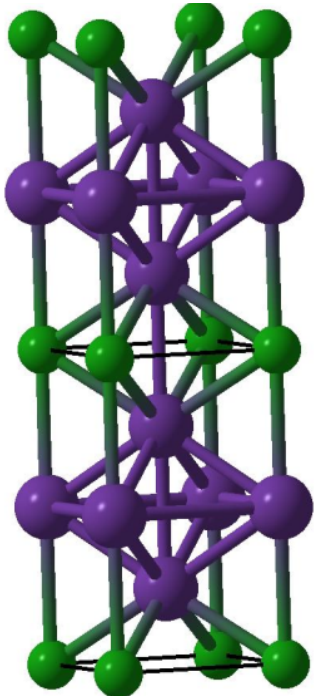
NaCl



$NaCl_7$



$NaCl_3$



Na_3Cl

[Zhang, ARO, et al., Science (2013)]

The USPEX (Universal Structure Prediction: Evolutionary Xtallography) project

http://uspex.stonybrook.edu

- The most popular code for computational materials design in the world (>1700 users)
- The largest, the most versatile, the fastest and the most reliable code in this field.

Many of its capabilities are unique. 3D-, 2D-, 1D-, 0D- systems can be treated

- THE CODE IS FREE
- Effort of ~50 man-years
- ~200 publications, 2 US patents