The Three Pillars of Science

Juan Meza
Department Head, Lawrence Berkeley National Laboratory
http://hpcrd.lbl.gov/~meza

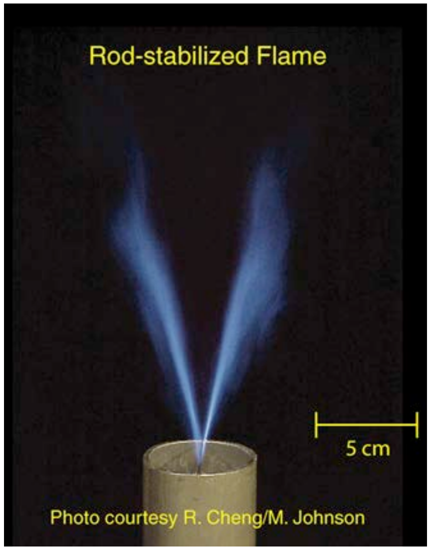
Personal Background

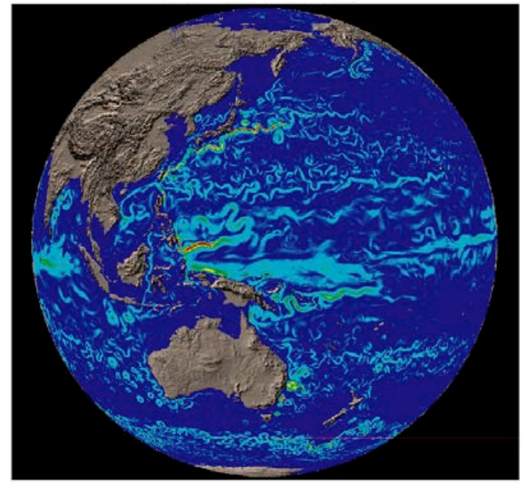
- Born in Houston, Texas
- Lived in Mexico from ages 6-9
- Returned to Houston, graduated from Milby High School
- Attended Rice University, studied EE/CS
- Worked at Amdahl, Exxon for ~ 2 years
- Returned to Rice to study Mathematical Sciences (Computational Mathematics)
- Postdoc for 1 year
- Worked at Sandia for almost 15 years
- Joined Berkeley Lab in 2002 - currently Senior Scientist/Department Head for HPC Research

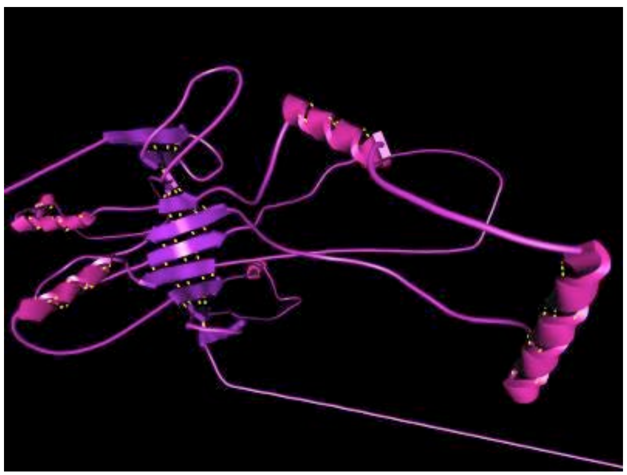
Daily Activities

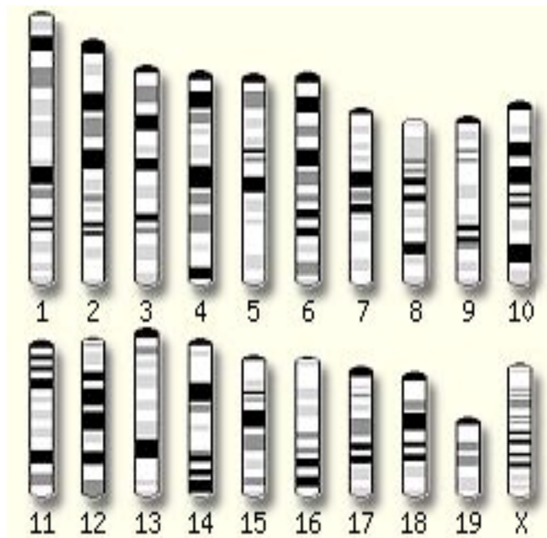
- Most of my time is spent keeping up with these two
- Try to keep both of them interested in math and science
- Watch, play, coach soccer
- Learning new languages, e.g. POS, BRB, BI, ...
- Go to many meetings
- Research in computational science and mathematics

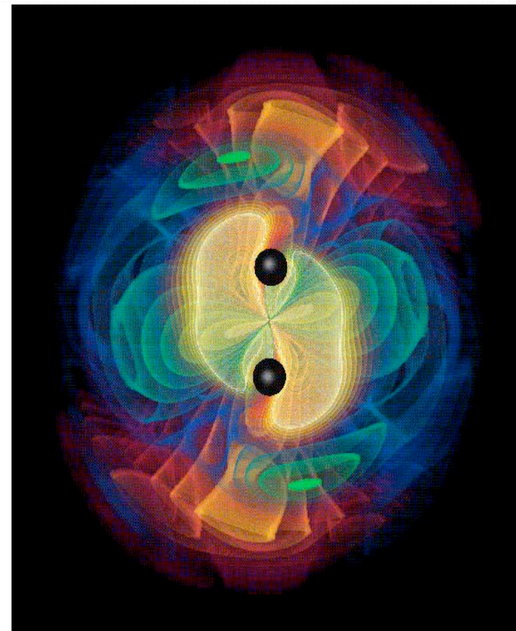
Ever broader use of computational sciences for scientific discovery











COMPUTER SIMULATION: THE THIRD PILLAR OF SCIENCE

"In many cases, theoretical and experimental approaches do not provide sufficient information to understand and predict the behavior of the systems being studied. **Computational modeling and simulation**, which allows a description of the system to be constructed from basic theoretical principles and the available experimental data, **are keys to solving such problems**."

Dr. Raymond L. Orbach, Director, Office of Science, Computation Science: A Research Methodology for the 21st Century, APS Meeting, APS, March 2004

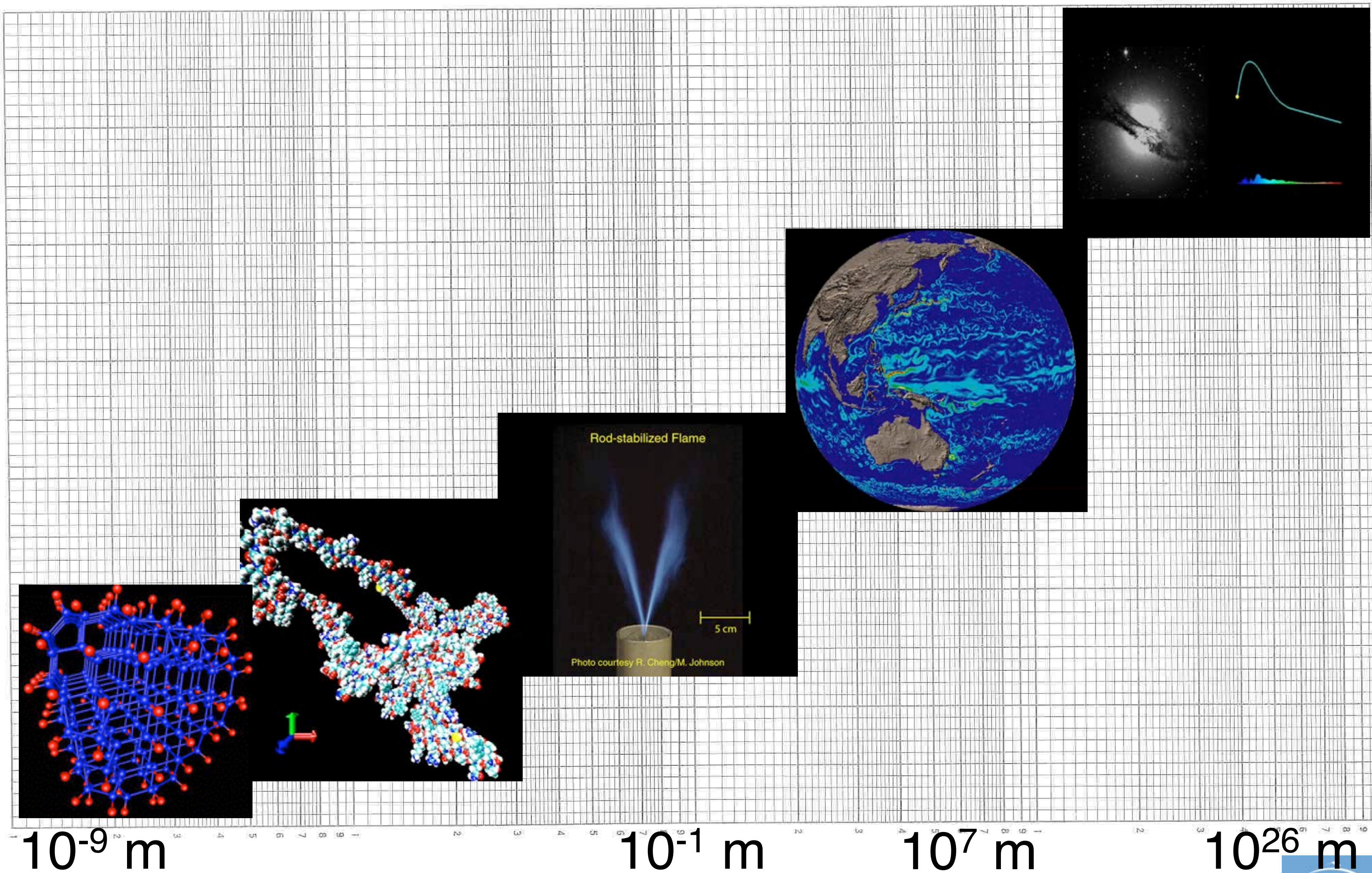
Amplifying the advancement of science and engineering research

Information technology amplifies research in other disciplines in a similar way. As this committee is aware, information technology gave rise to new tools for performing research, computational science techniques. Previously, research was experimental, observational or analytical. Progress in computer and information science and engineering not only advances information technology itself, but leverages advancement of knowledge in other areas. It shares this trait with mathematics. But most other disciplines like astronomy or geology do not offer such leverage.

So, my first conclusion is *that investment in the research in computer and information science and engineering has strengthened our economy not just by enabling entirely new products and industries, but by amplifying the efficiency and productivity of almost all other areas of our economy as well as amplifying the advancement of science and engineering research*. It is extraordinarily productive.

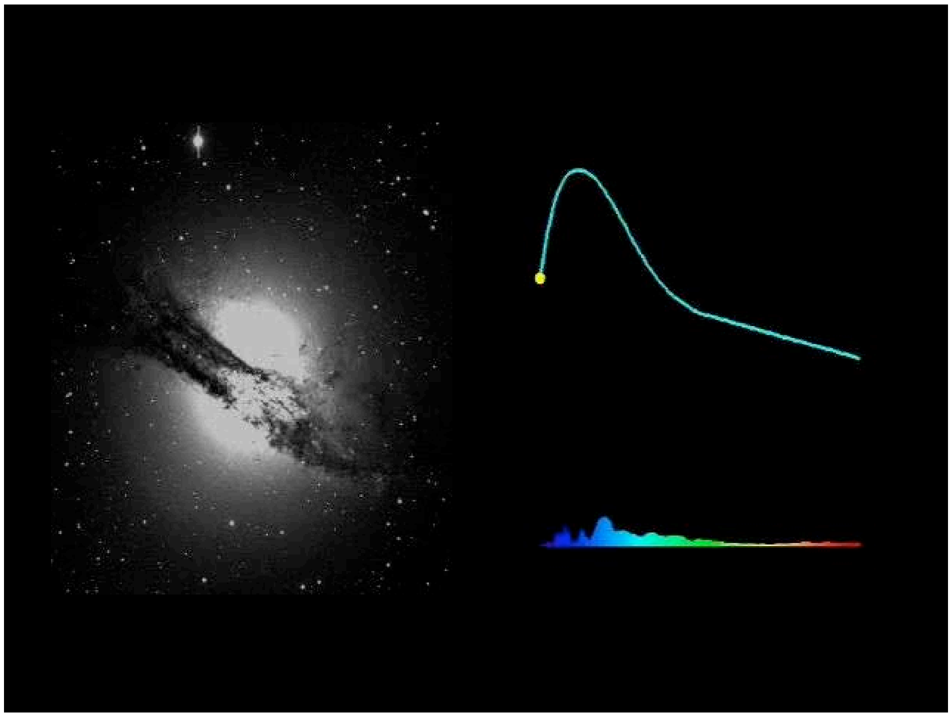
Anita Jones
Quarles Professor of Engineering & Applied Science
University of Virginia
Before the Subcommittee on Research
House Committee on Science
June 16, 2001

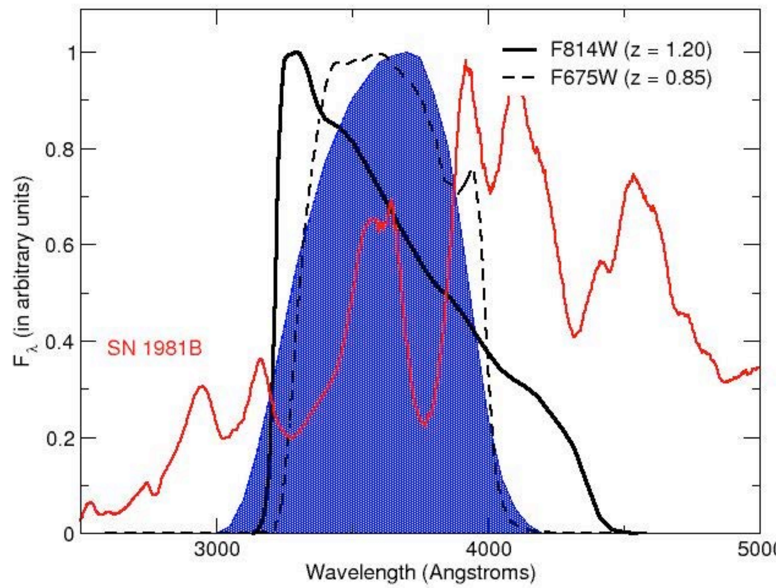
Roadmap - A trip through time and space



Results: Spectral Templates

K-corrections - How we compare supernovae at different redshifts. Especially in the UV…

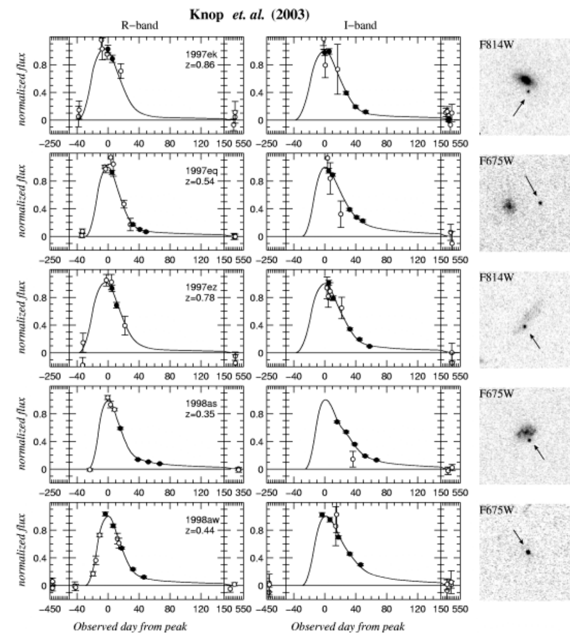




The spectral templates were created by homogenizing IUE and HST observations + modeling to fill in the gaps.

Spectral Templates for Type Ia Cosmology…

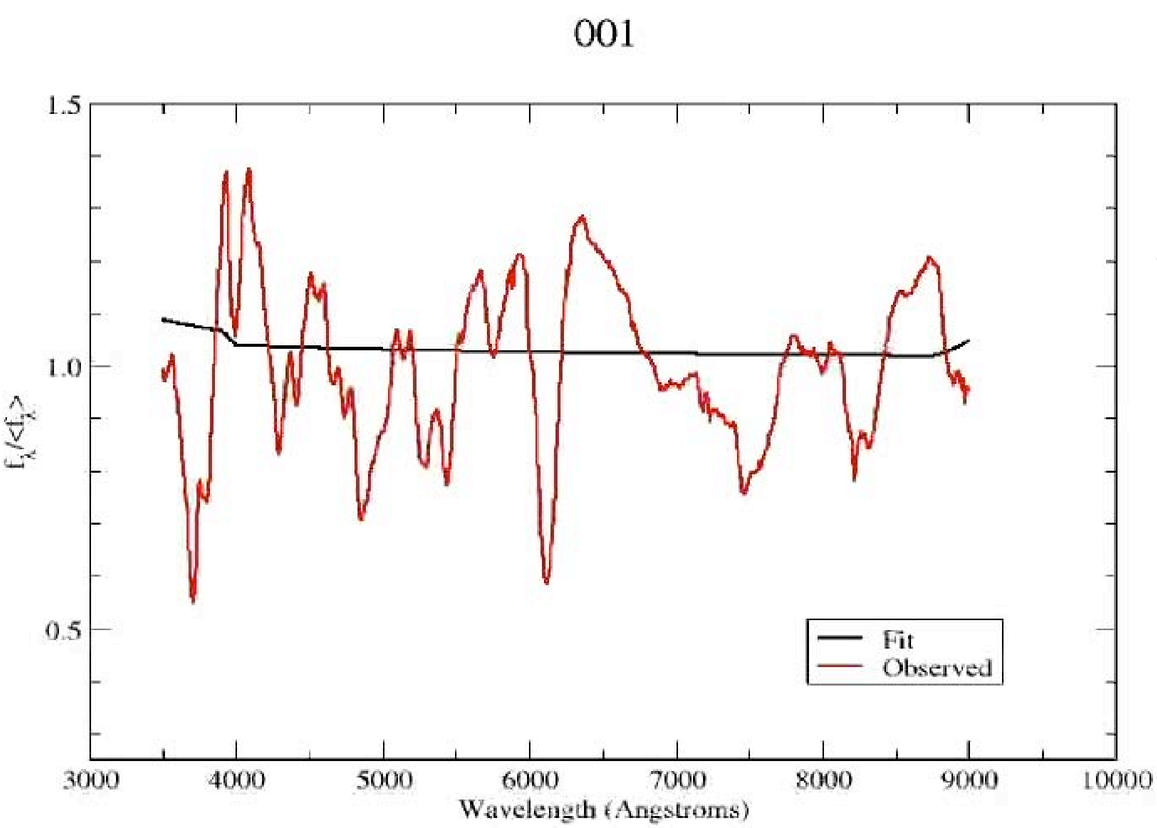
Aphrodite (z~1.3) from the Riess *et al.* GOODs Very High-Z SN Search



Precision measurements from Knop *et al.* HST data

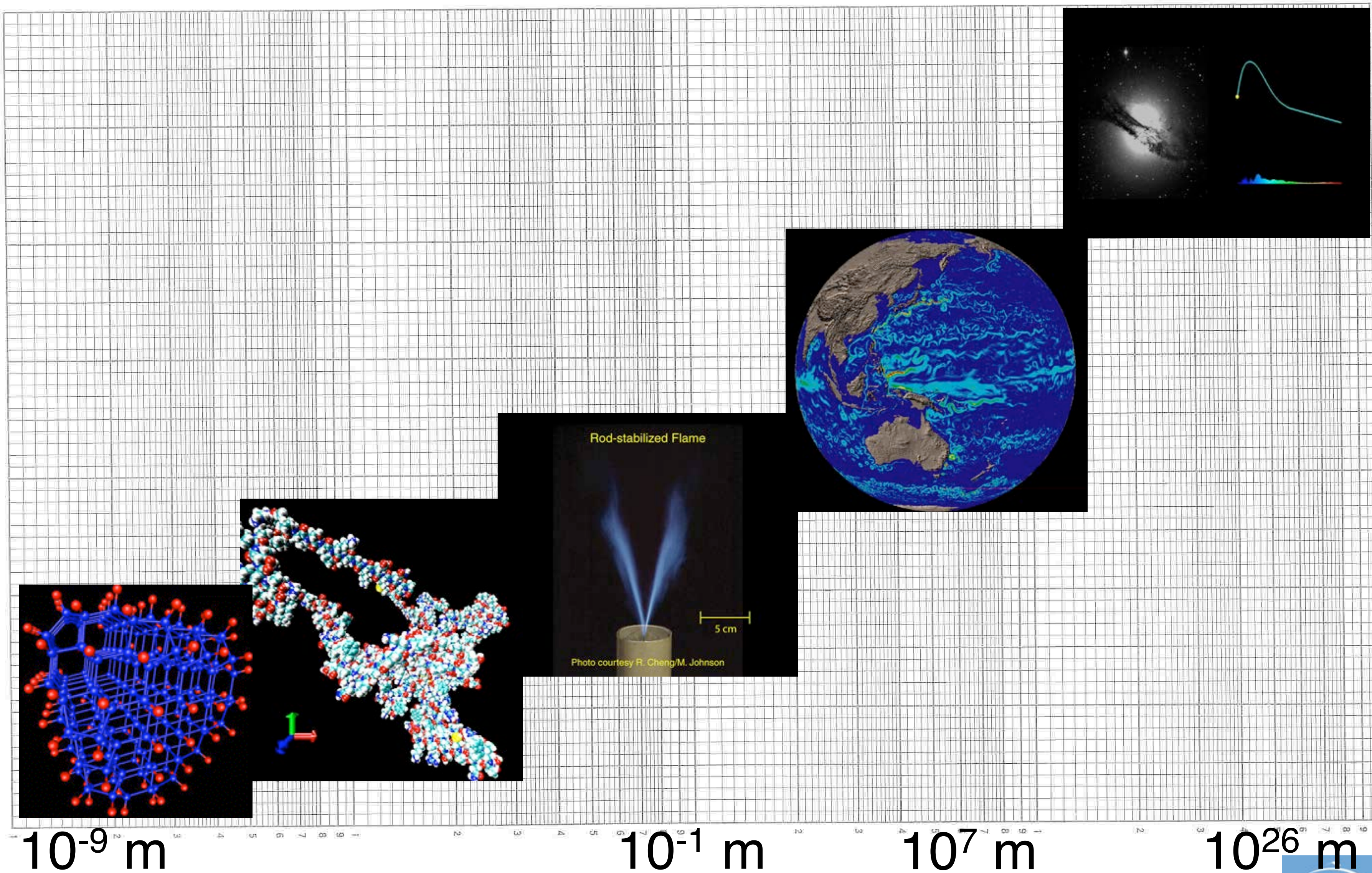
SN 1997ff, still the highest redshift SN Ia observed to date from Riess, Nugent, *et al.*

Real-Time Assessment of Data for SN Factory/JDEM



- **Combines optimization research with spectrum synthesis.**
- **Makes it possible to"objectively" fit a spectrum and determine the parameters of the model atmosphere and their uncertainties.**
- **Allows for real-time assessment of many spectra**

Roadmap - A trip through time and space



Climate extremes

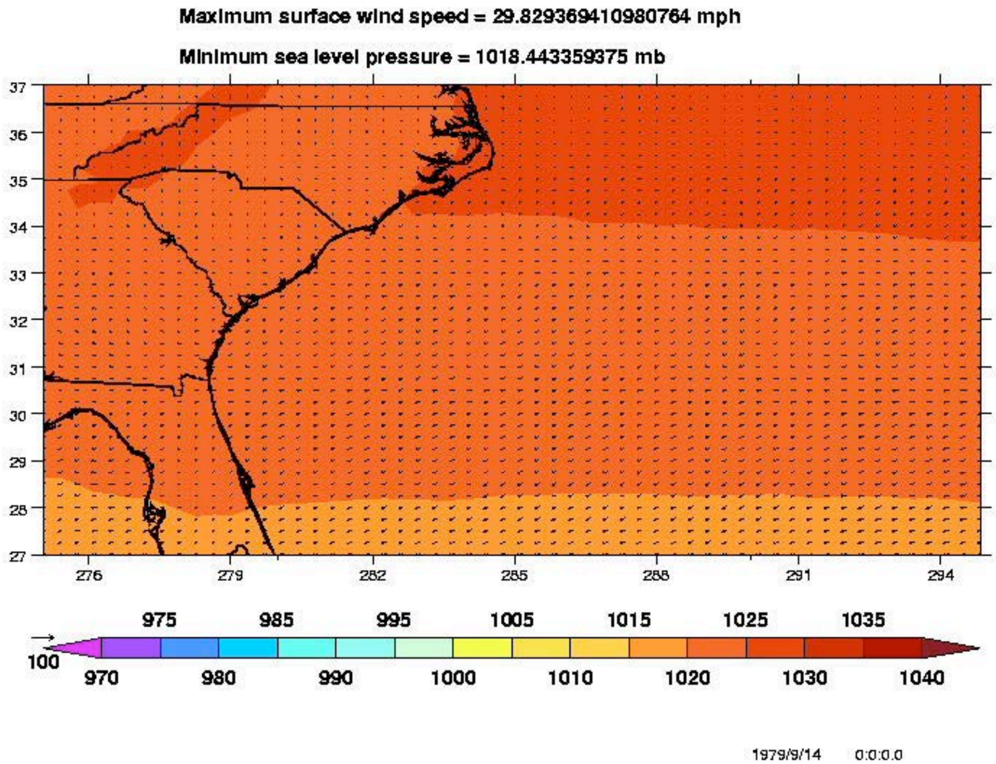
- There is very little doubt that the climate is changing due to human influences.
  - I.e. thermal structure of the atmosphere, Santer, Wehner, et al Science **301** (2003) 479-483
- However, we live in the noise (weather) not in the mean (climate).
- What can we say about recent and future changes in extreme weather?
  - Quantitative statements about extreme weather push the capabilities of current models.

Extreme weather. Computing?

- Droughts, floods, heat waves, cold spells
  - Quantifying change is a signal to noise problem.
  - To extract signals, large ensembles of integrations are required.
- Hurricanes
  - Also a signal to noise problem
  - But much higher spatial resolution is required as well.

- Computing technology is a limiting factor.

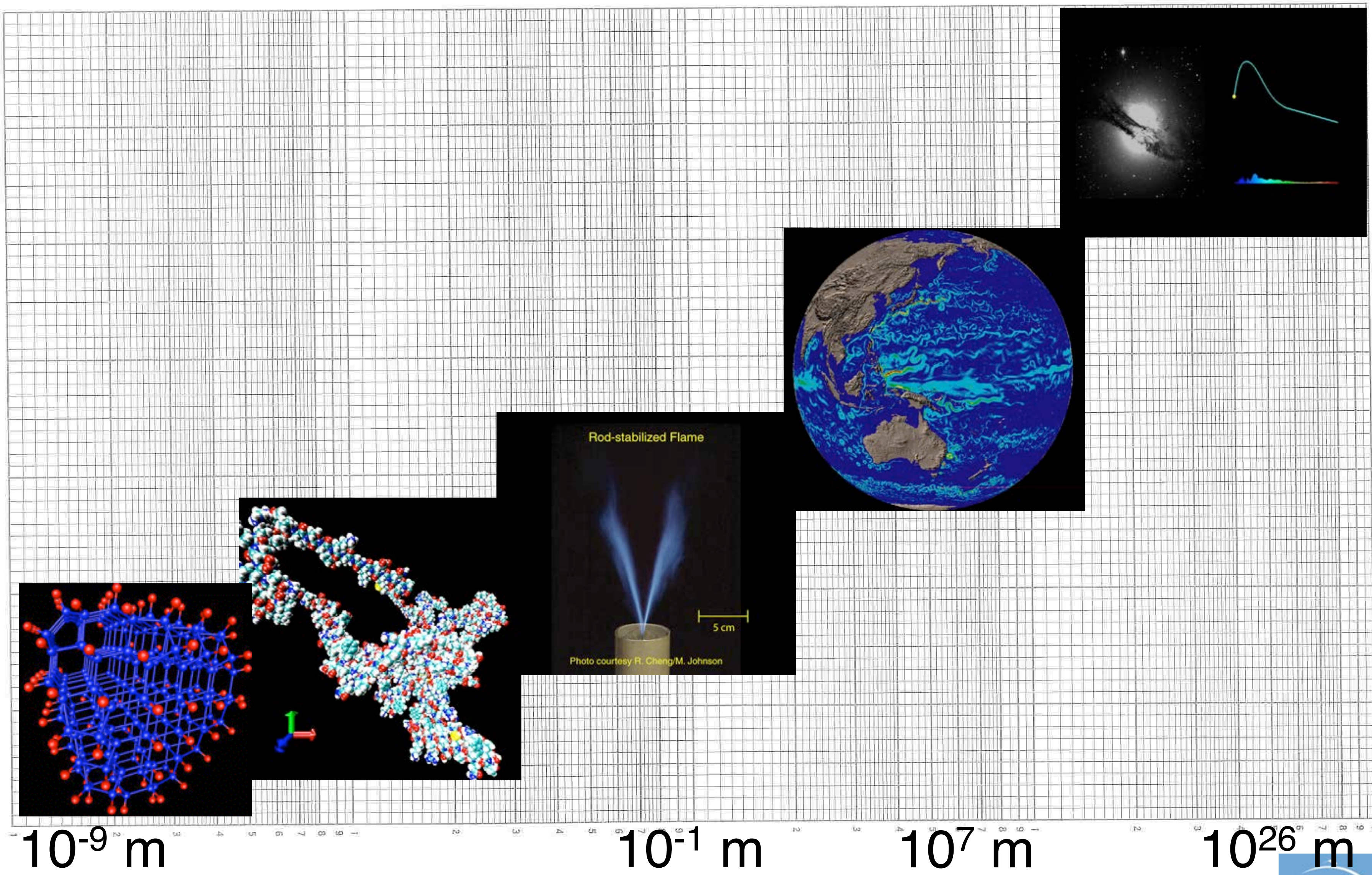
Climate modeling and predicting hurricane patterns



Michael Wehner, Scientific Computing, LBNL

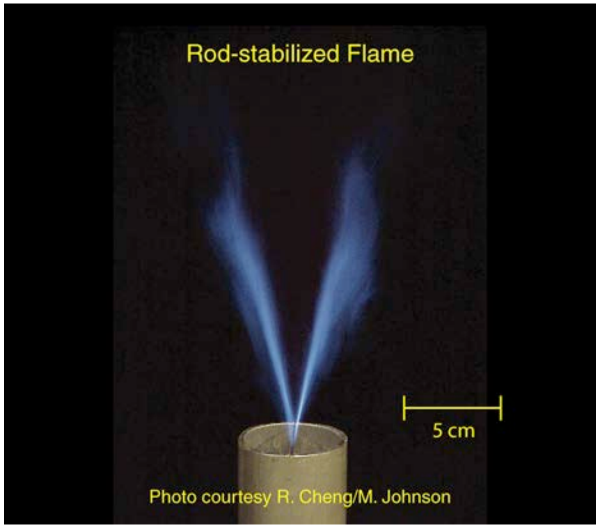
- Tropical cyclones are not generally seen in integrations of global atmospheric general circulation models at climate model resolutions T42 (~300km)
- In CCM3 at T239 (50km), the lowest pressure attained is 995mb. No realistic cyclones are simulated.
- In high resolution simulations of the finite volume dynamics version of CAM2, strong tropical cyclones are common.

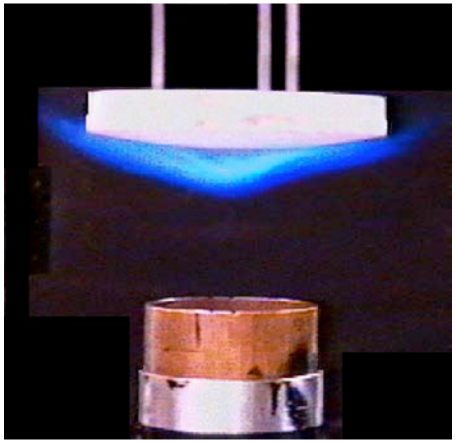
Roadmap - A trip through time and space



Turbulent premixed flames

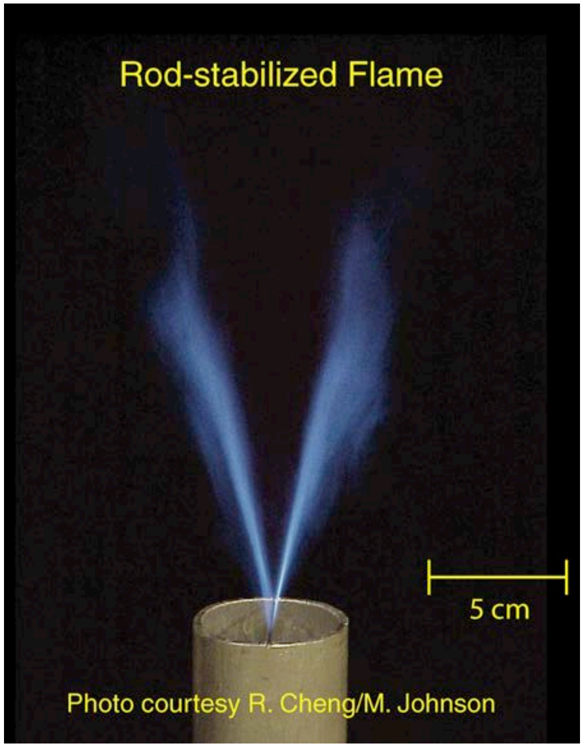
- Most new combustion systems are premixed
  - Lean premixed systems have potentially high-efficiency and low emissions
  - Premixed flames are unstable
  - Systems are turbulent





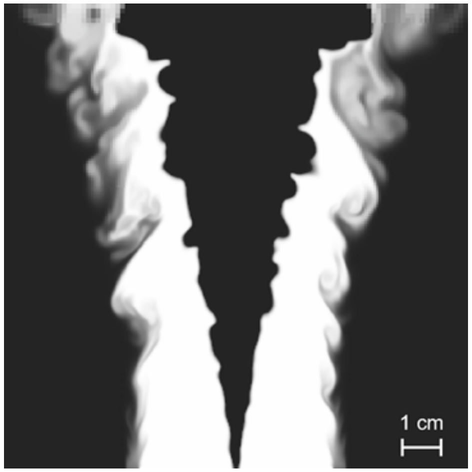
V-flame

- Simulate turbulent V-flame
- Strategy – Independently characterize nozzle and specify boundary conditions at nozzle exit
  - 12 x 12 x 12 cm domain
  - Methane at $\phi = 0.7$
    - DRM 19, 20 species, 84 reactions
    - Mixture model for species diffusion
  - Mean inflow of 3 m/s
  - Turbulent inflow
    - $l_t = 3.5$mm, u' = 0.18 m/sec
    - Estimated $\eta = 220\ \mu$ m
  - No flow condition to model rod
  - Weak co-flow of air

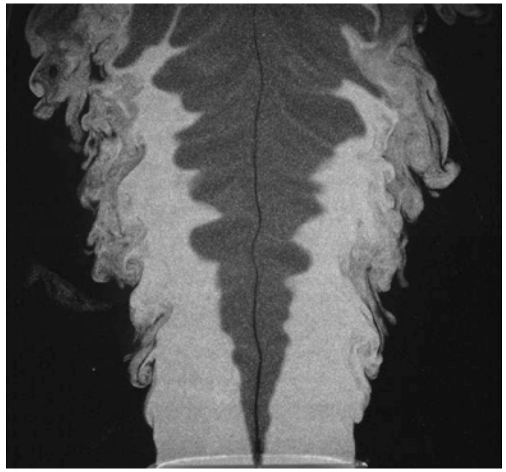


Experimental comparisons

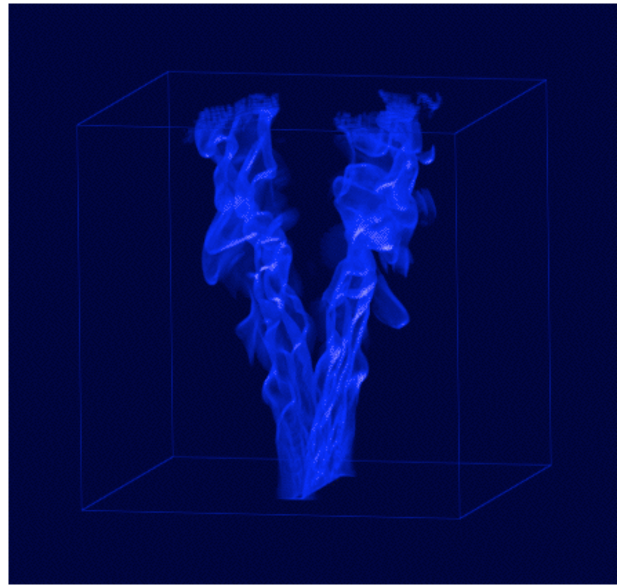
Simulation

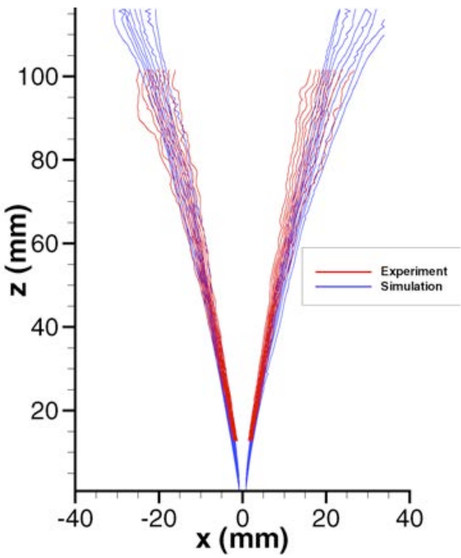


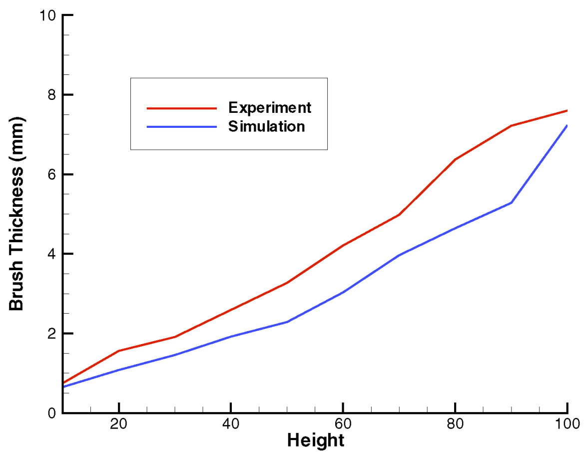
Experiment



Instantaneous flame surface animation





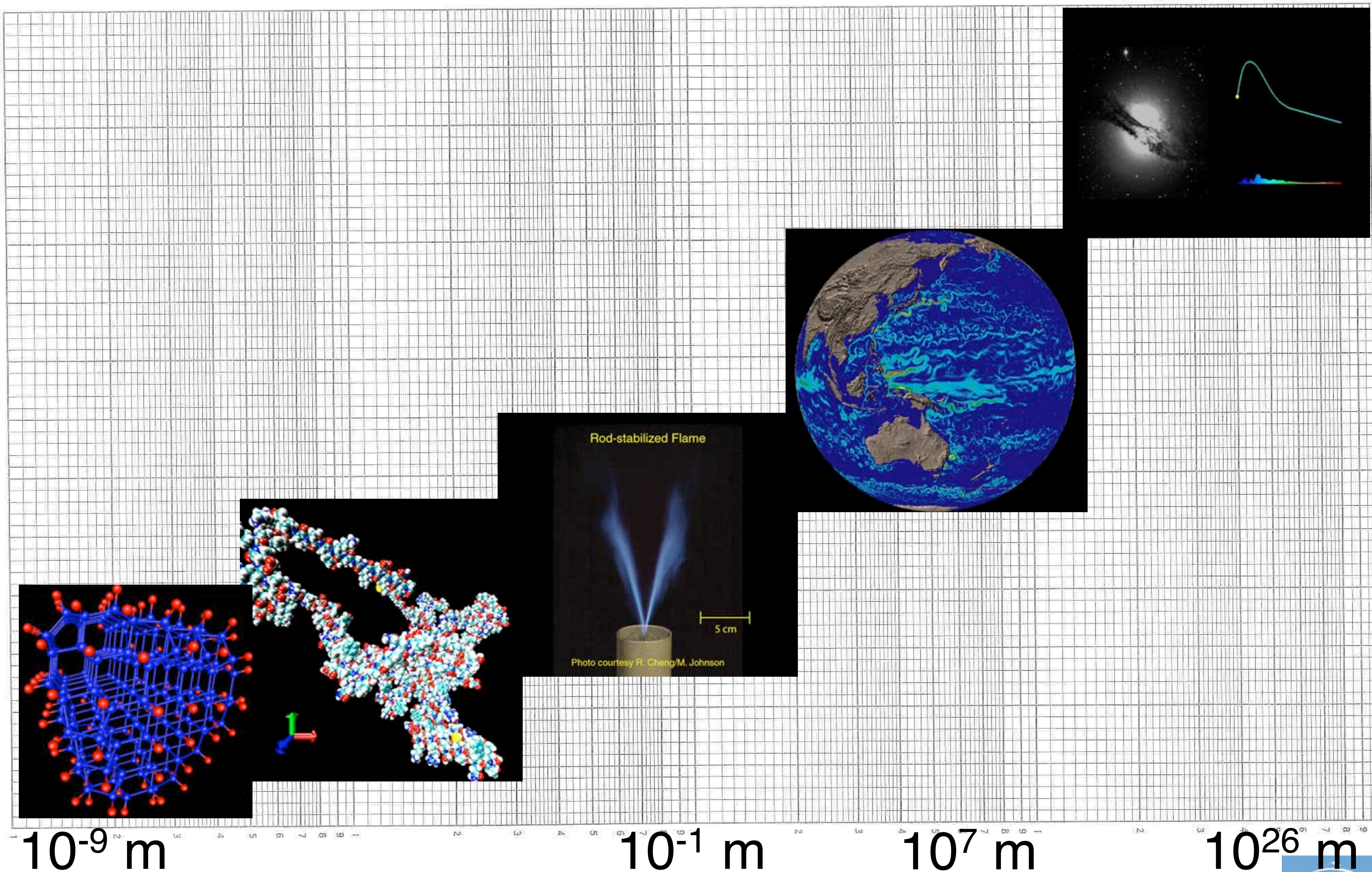


Flame brush comparisons

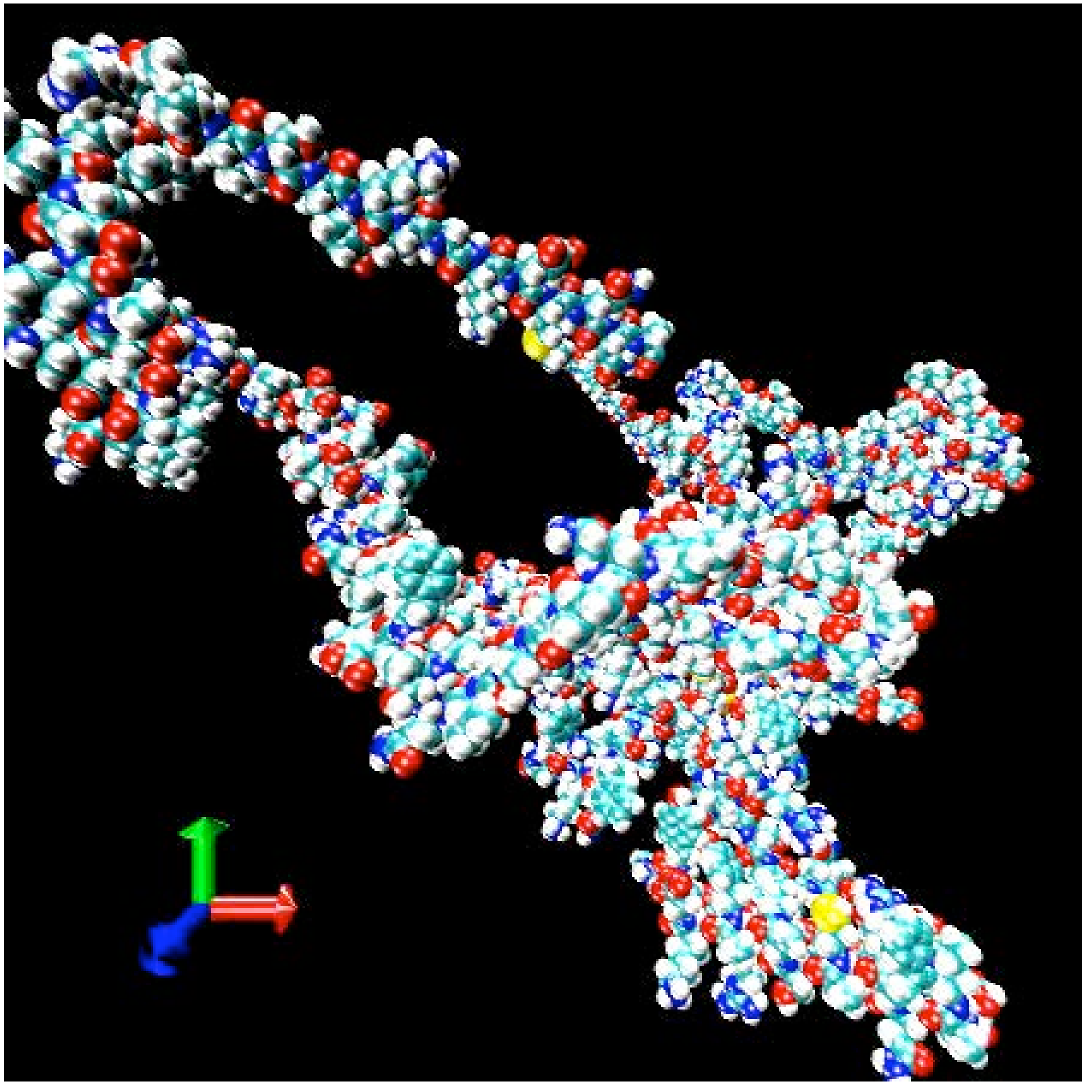
Joint with R. Cheng, M. Johnson and I. Shepherd

Submitted PNAS, 2005

Roadmap - A trip through time and space



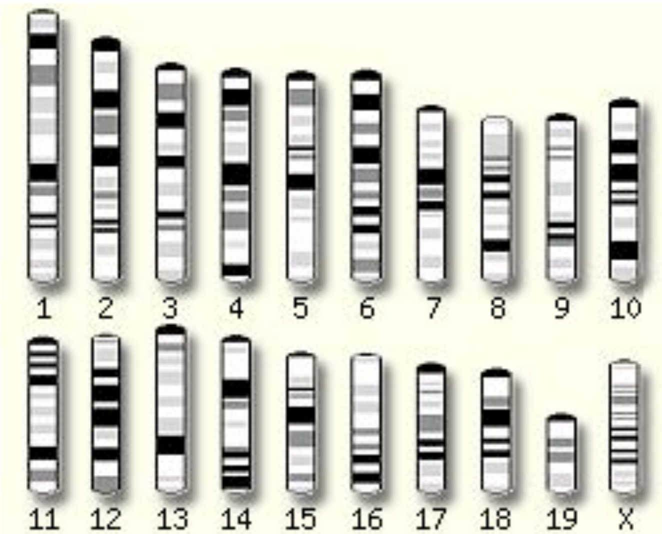
Protein T162 from CASP5



- Initial configuration created using ProteinShop (S. Crivelli)
- Energy minimization computed using OPT++/LBFGS
- Final average RMSD change was 3.9 Å
- Total simulation took approximately 32 hours

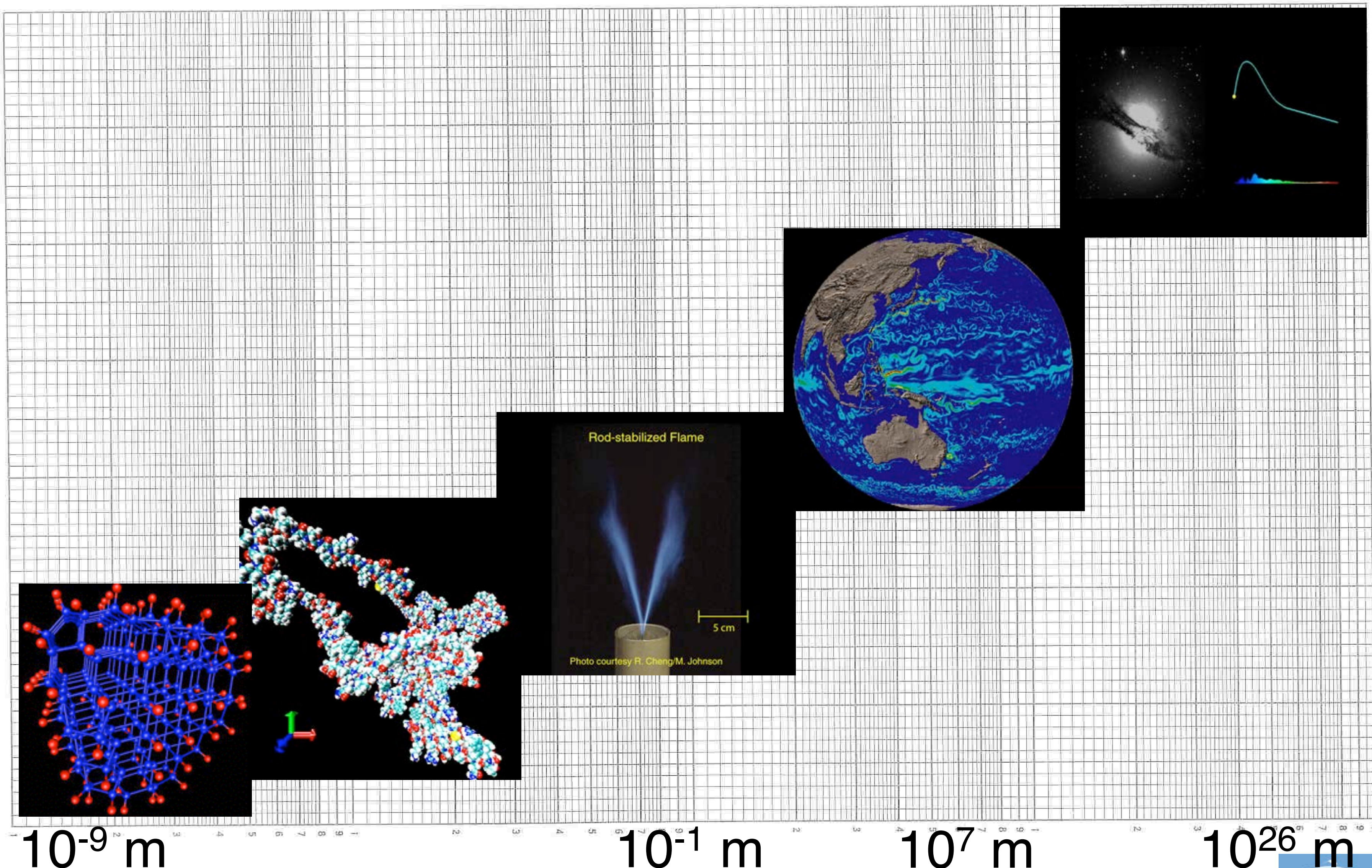
Juan Meza, Ricardo Oliva, Scientific Computing, LBNL

Assemby of Fugu genome

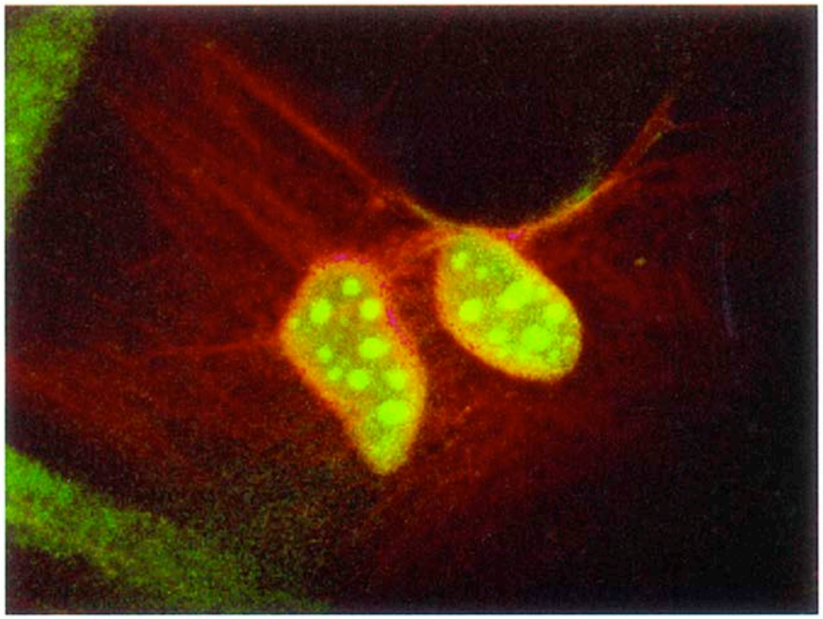


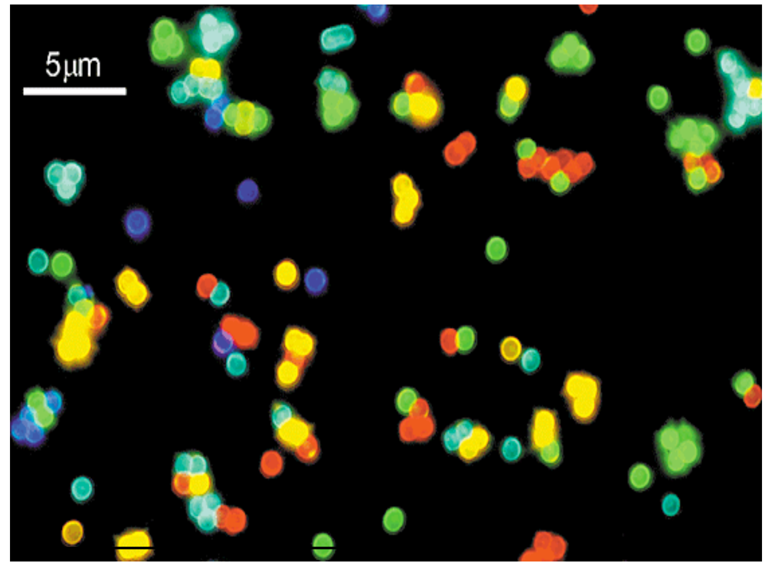
- Assembly of Fugu genome from 3.1 million reads, and initial preparation of mouse genome data.
- 75% of human genes have counterparts in Fugu genome
- Easier to find genes in Fugu because it has fewer noncoding (junk) DNA
- Led to prediction of 961 previously unidentified human genes
- Need new discrete math algorithms to study these problems

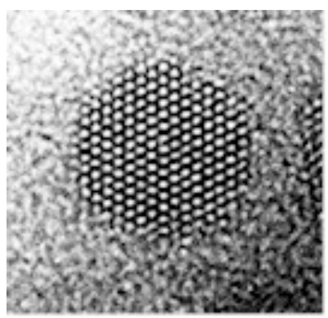
Roadmap - A trip through time and space



Many interesting applications of nanostructures



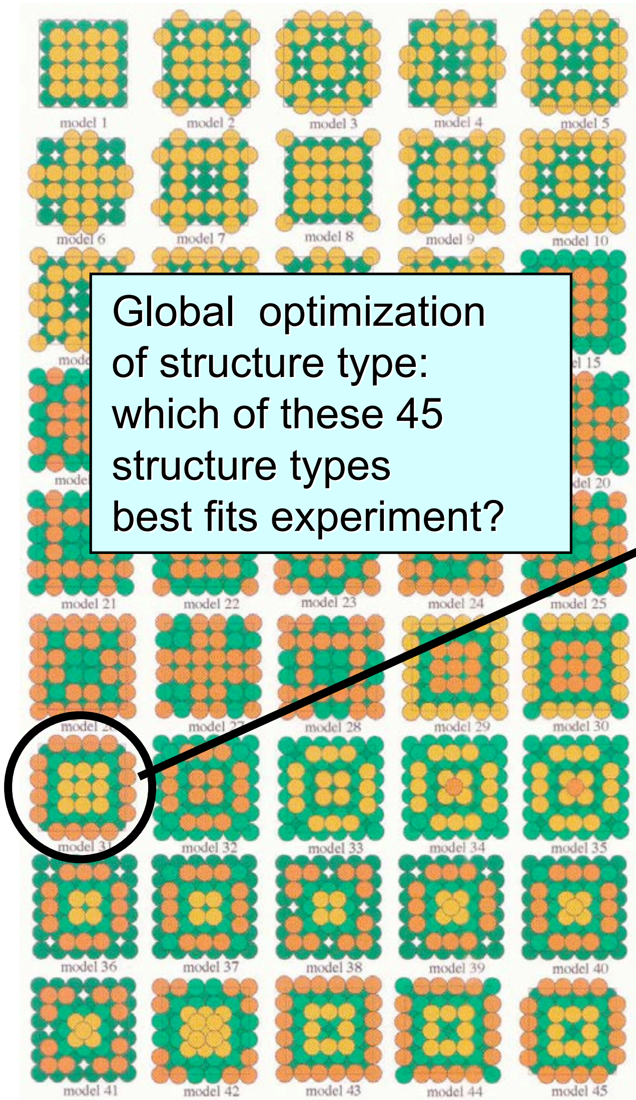




- CdSe quantum dot
- Optically more stable than dye molecules
- Can have multiple colors

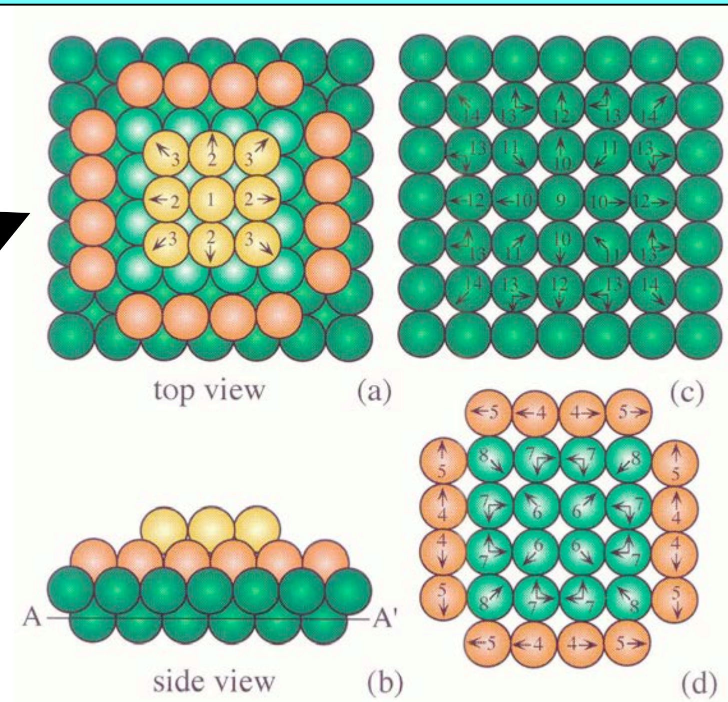
Surface structure determination from experiment

- Electron diffraction determination of atomic positions in a surface:
  - Li atoms on a Ni surface



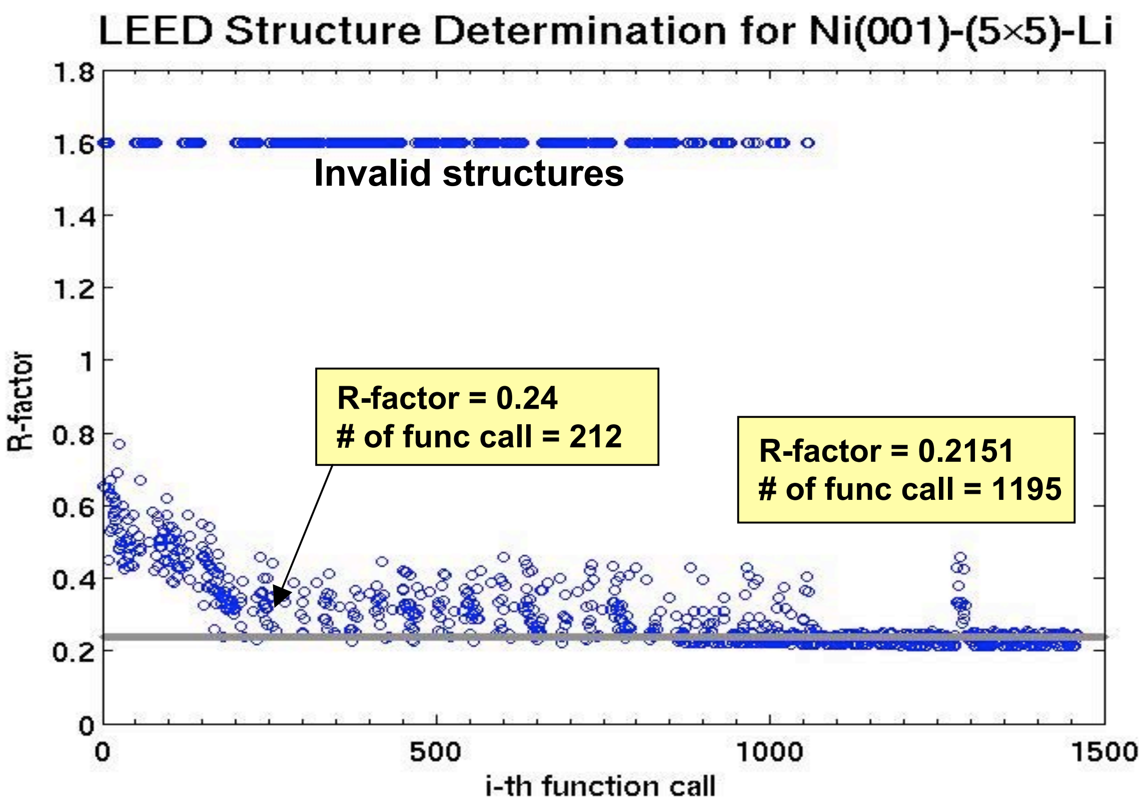
Ni(001)-Li-(5x5) structure models

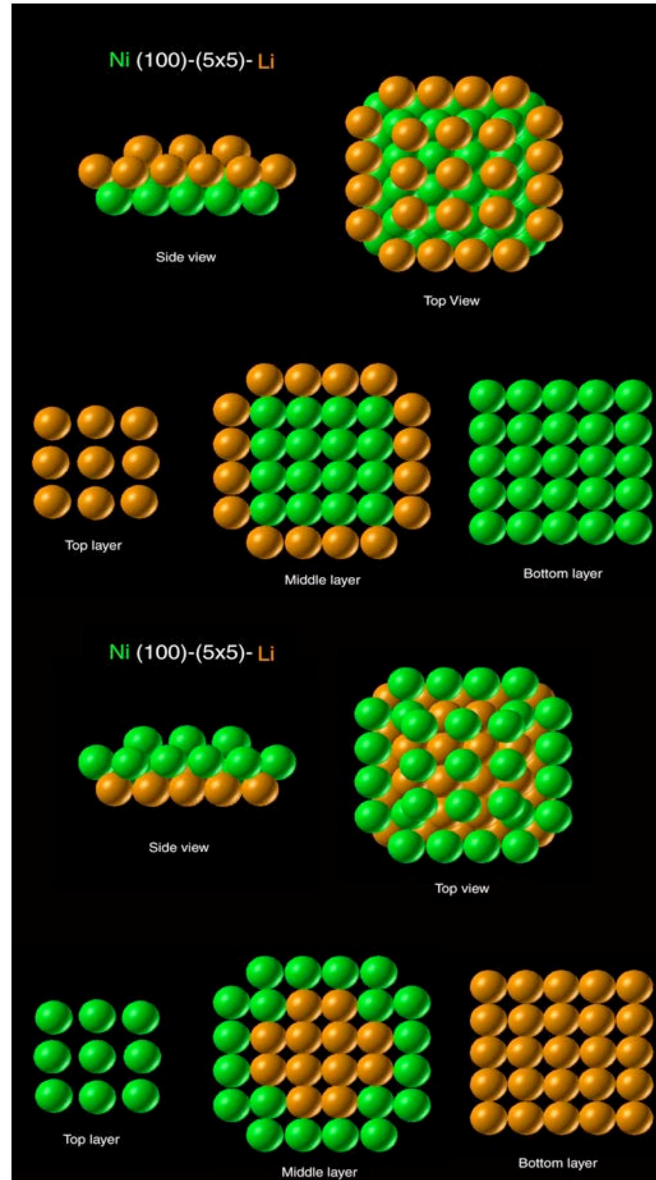
Global optimization of structure type: which of these 45 structure types best fits experiment?



Local optimization of structure parameters: which are the best interatomic distances and angles?

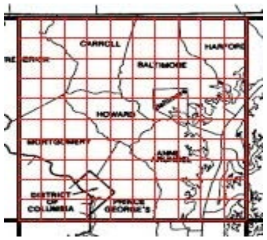
Minimization with respect to both types of variables removes coordinate constraints



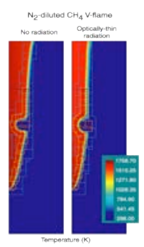


Previous best known solution R-factor = 0.24
New solution found with R-factor = .2151
Final (global) solution with R-factor = .1184

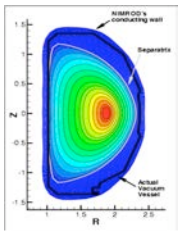
WHAT HAPPENS AT 50 TFLOPS SUSTAINED SPEED?



The earth simulator works on a grid 10 km on a side for climate models, while U.S. computers do no better than 100 km on a side. U.S. simulations average over microclimates — mountains and coastal effects, river flow, cloud and storm systems, or hurricane storms. Averaging means that our models cannot predict much about regional scale and perhaps higher frequency fluctuations, critical for drought and flood predictions. Higher resolution does mean we can do better science, and faster machines will help us with higher resolution runs -- but the science in the models needs to be improved too before we can be too confident in the models at any resolution



We can use simulation to design coal burning boilers that dramatically could reduce NOx emissions by almost one million tons per year in the United States alone.. Simulations are needed with increased fidelity in both chemical kinetics and spatial resolution.



We can develop simulations of burning plasmas, including ITER, that include electron dynamics and will enable us to predict and avoid major disruptions. Realistic simulations of disruptions in ITER would require several orders of magnitude increase in the number of space and time points for the calculation as well as additional physical properties, most importantly a model of the plasma–wall (divertor) interaction.

For evaluation of a design alternative for the purpose of optimization of a jet engine design, GE would require 3.1 x 1018 floating point operations, or 3.6 days of sustained speeds of 10 Tflops. 100 Tflops of sustained speed would require "only" 8.6 hours. This is to be compared with millions of dollars, several years, and designs and re-designs for physical prototyping.

Dr. Raymond L. Orbach, Director, Office of Science, Computation Science: A Research Methodology for the 21st Century, APS Meeting, APS, March 2004

Summary

- Computational science is increasingly being used to aid in the scientific process
  - PDEs, ODEs, FFTs
  - Linear Algebra/Eigenvalues
  - Nonlinear equations and optimization
- Mathematical hurdles must be overcome to solve real world problems
- Many new areas are cropping up
  - Data mining
  - Discrete math and combinatorics
  - Complex systems

Lessons Learned

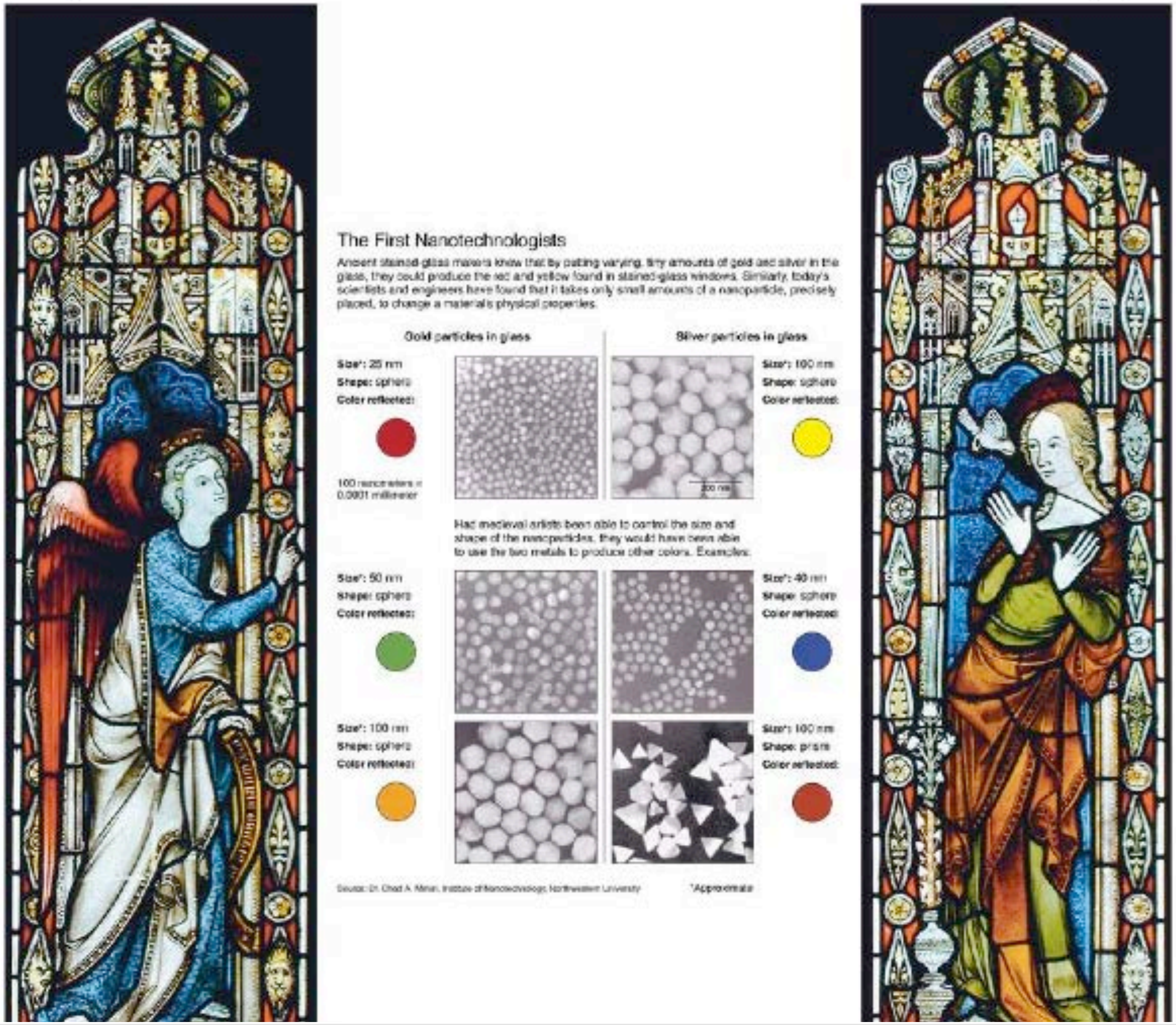
- Always worked on a (multidisciplinary) team
- Learning each other's jargon was usually the first and biggest hurdle
- Projects averaged 2-3 years
- Connections between many of the problems

Specifics of a particular discipline are not as important as the general concepts for understanding and communication

Questions

The New York Times

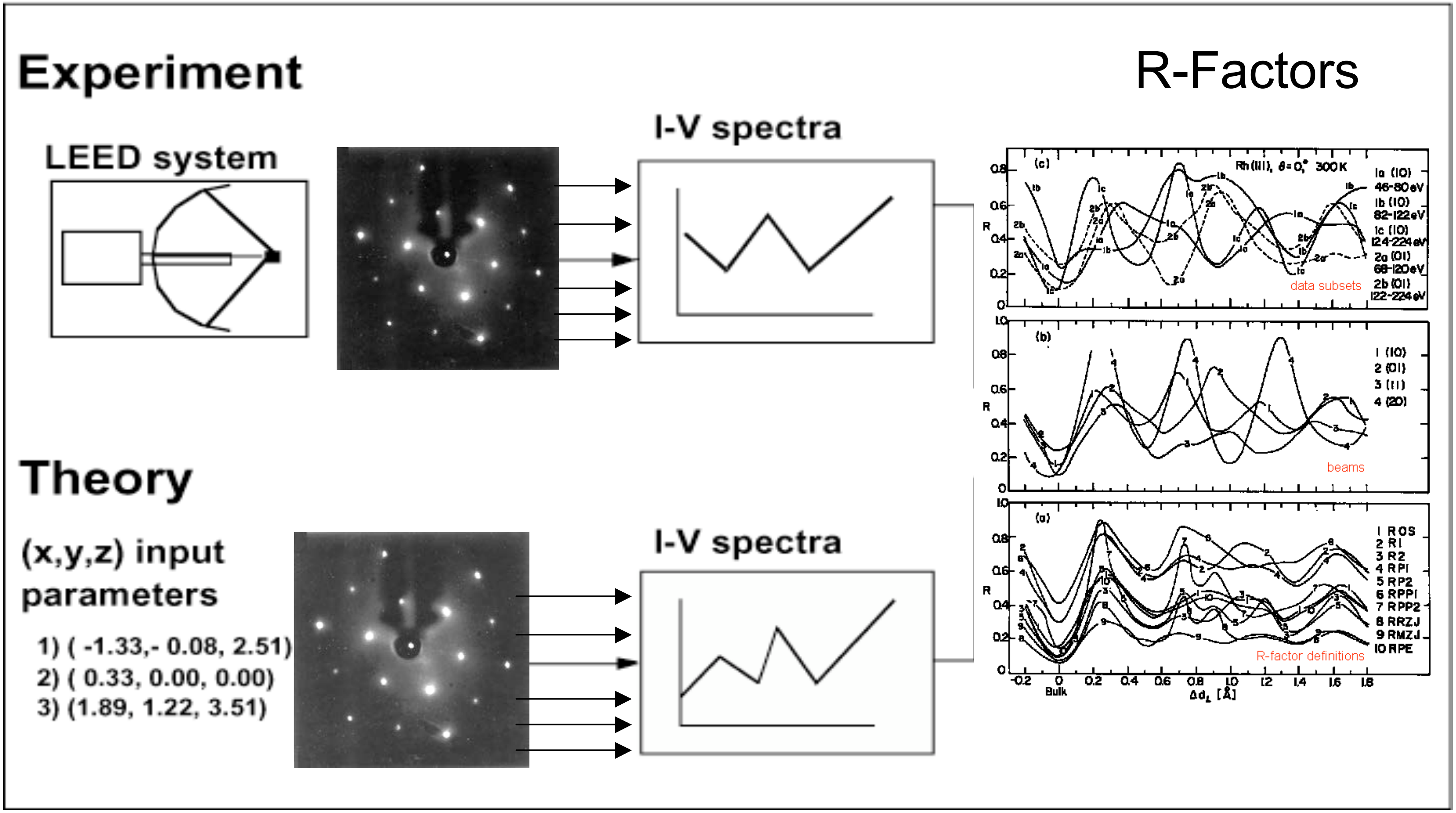
February 21, 2005



Thank you

Q&A

Low Energy Electron Diffraction

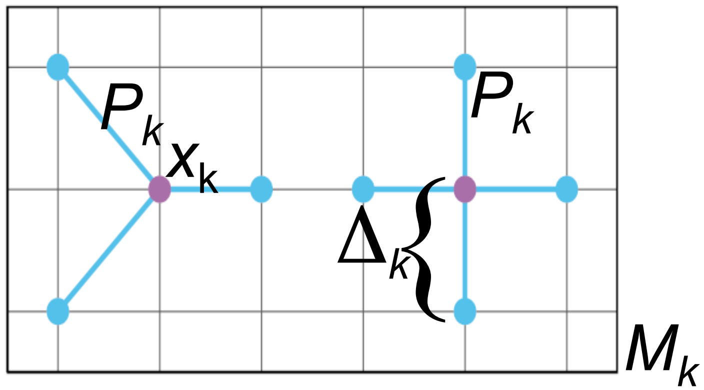


Generalized Pattern Search Framework

1. Initialization: Given Δ_0 , x_0 , M_0, P_0
2. For k = 0, 1, …
   a) SEARCH: Evaluate f on a finite subset of trial points on the mesh M_k
   b) POLL: Evaluate f on the frame P_k
3. If successful - mesh expansion:
   a) $x_{k+1} = x_k + \Delta_k d_k$
4. Otherwise contract mesh

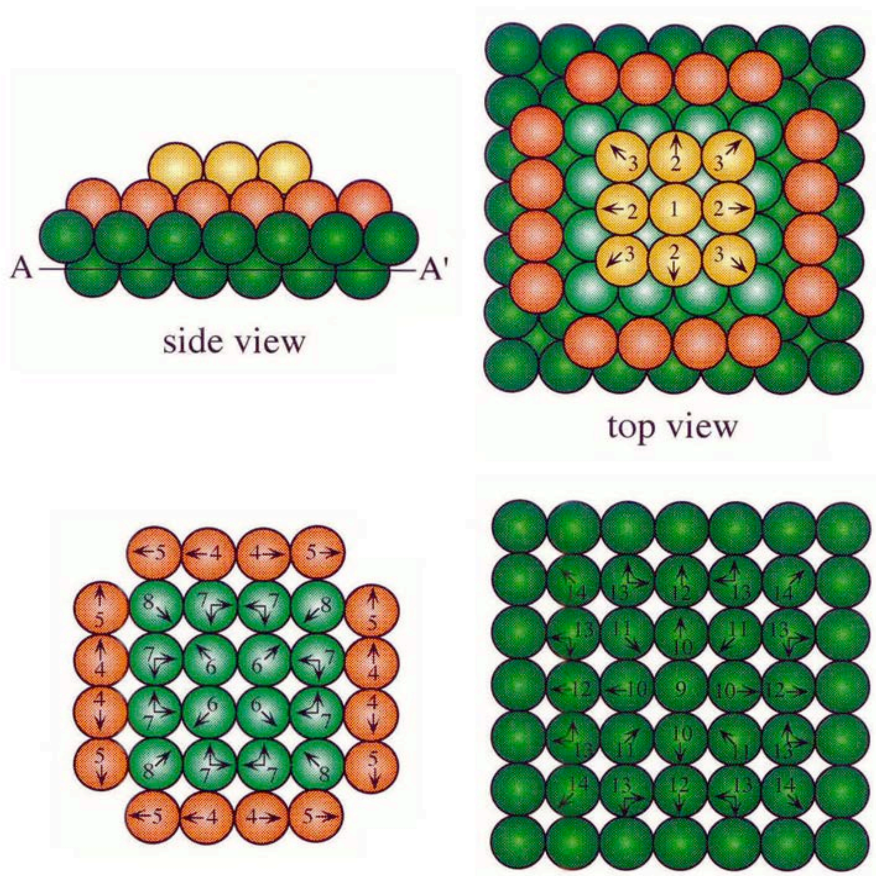
Global phase can include user heuristics or surrogate functions

Local phase more rigid, but necessary to ensure convergence



Test problem

Ni(100)-(5x5)-Li



Model 31 from set of TLEED model problems

- Model contains three layers of atoms
- Using symmetry considerations we can reduce the problem to 14 atoms
  - 14 categorical variables
  - 42 continuous variables
- Positions of atoms constrained to lie within a box
- Best known previous solution had R-factor = .24