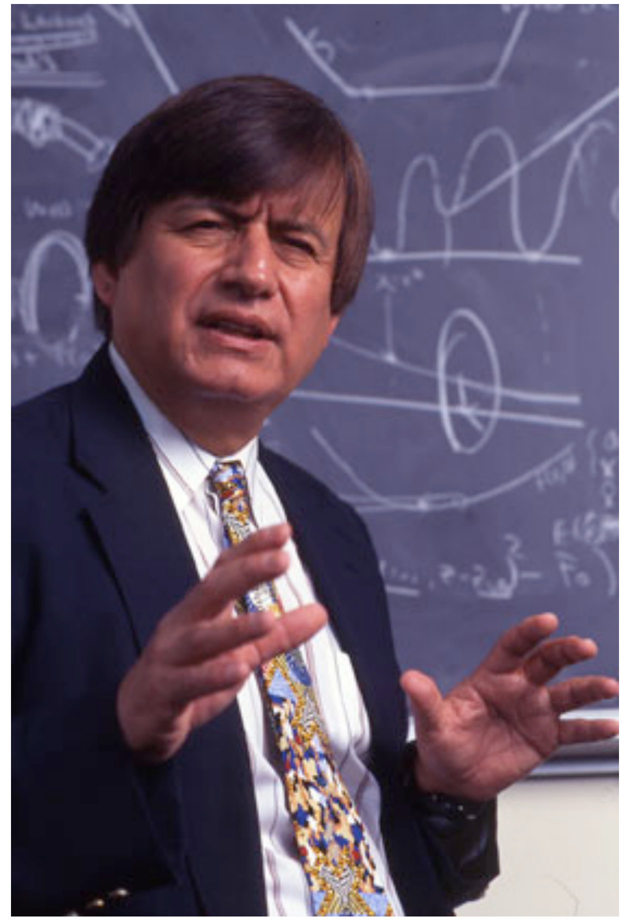
Optimization: The Difference Between Theory and Practice ...

Juan Meza
Senior Scientist and Department Head
High Performance Computing Research
Lawrence Berkeley National Laboratory

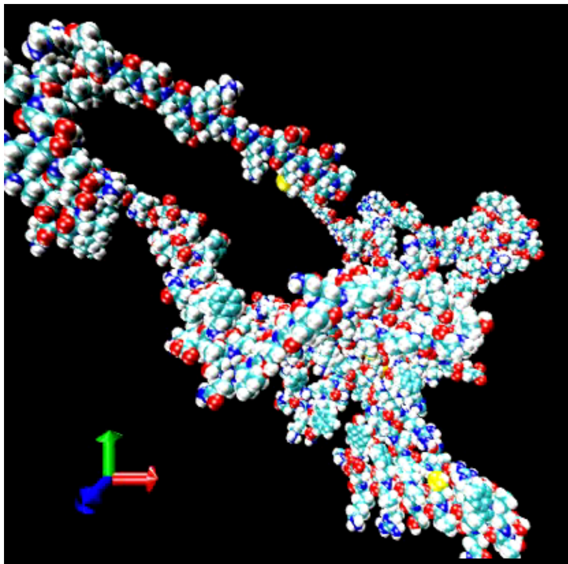
Blackwell-Tapia Conference SAMSI November 14-15, 2008

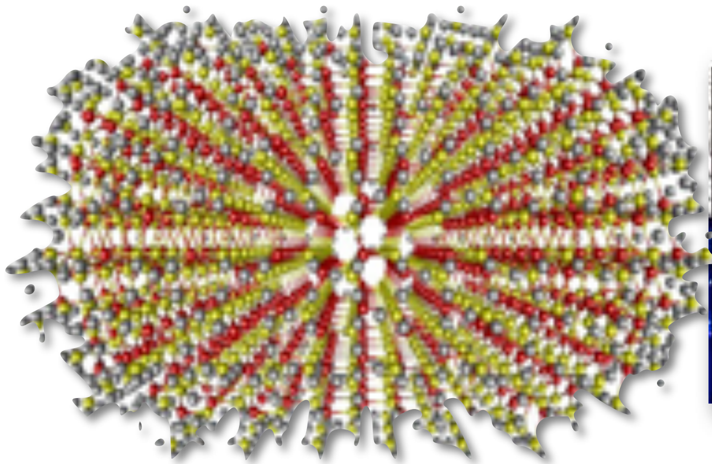
Why we're here

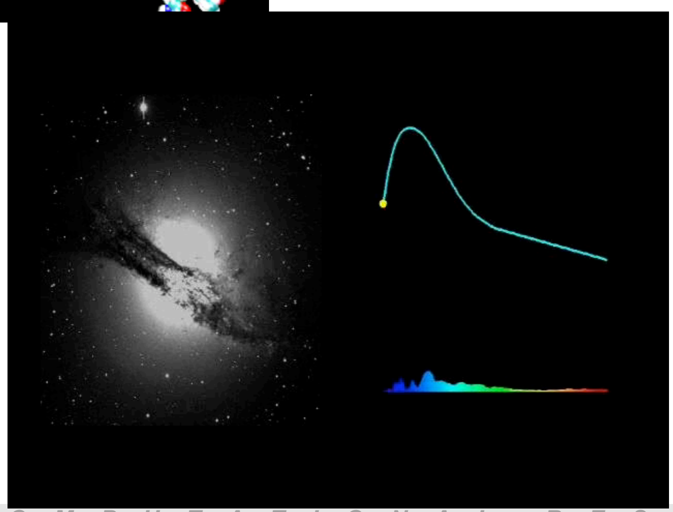


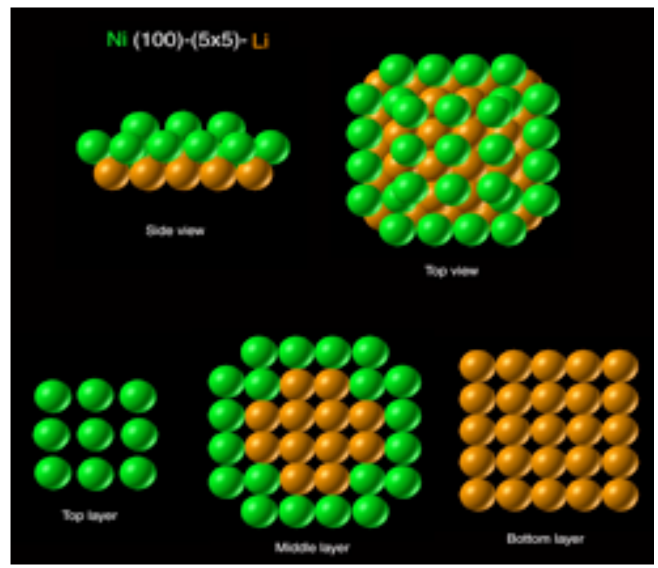
Thank you!!!

What do all of these have in common?









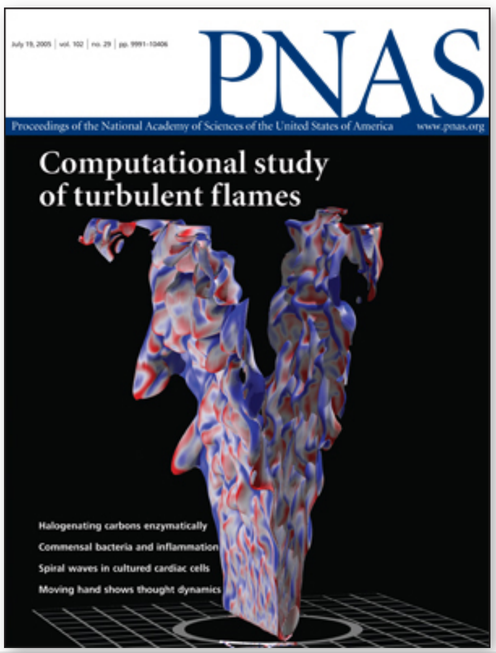
New Role of Computational Science

Experimental science has become increasingly difficult and expensive to do

Theory

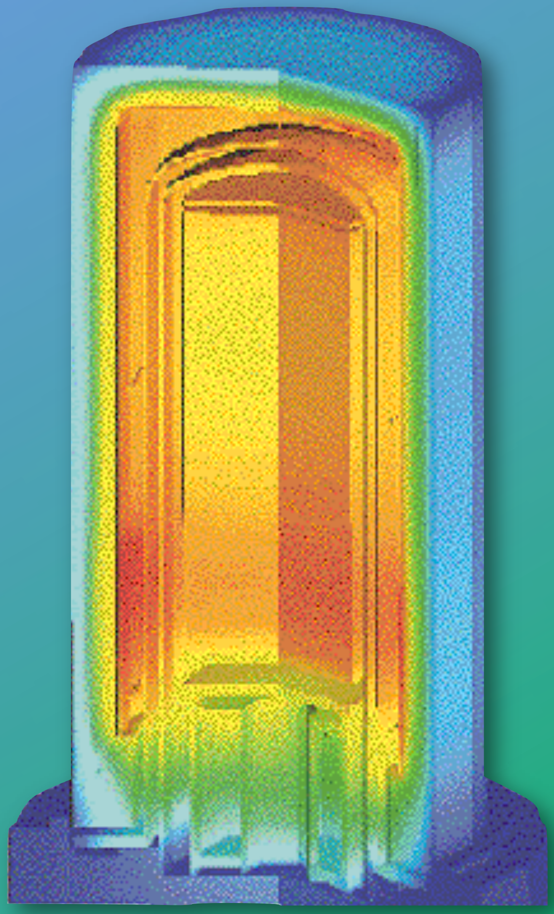
Experiment

Simulation

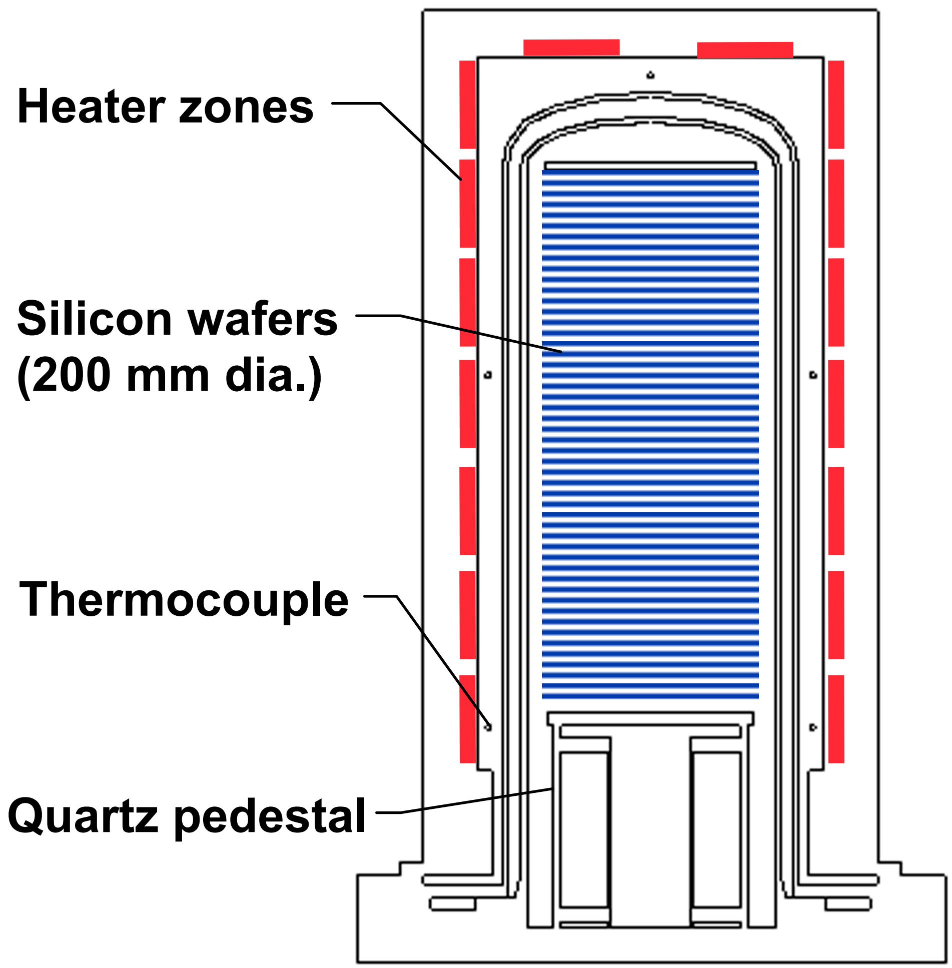


Computing power has increased tremendously

Chemical Vapor Deposition Control

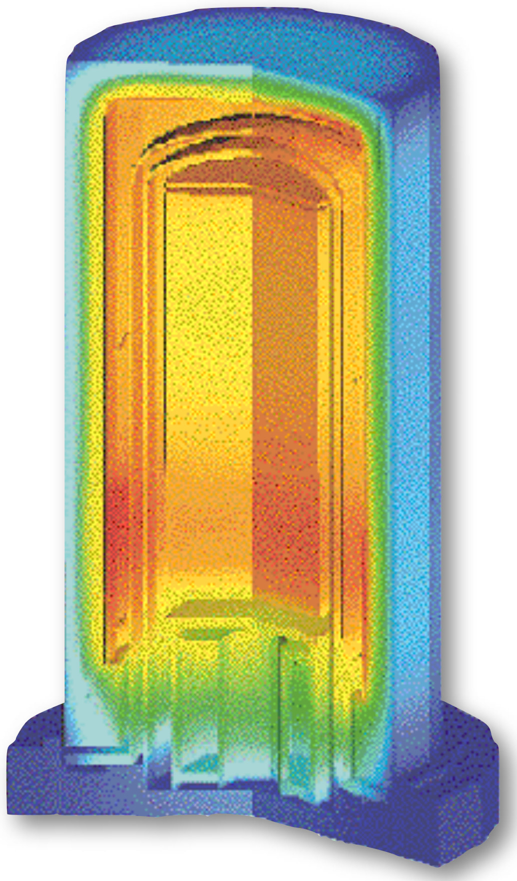


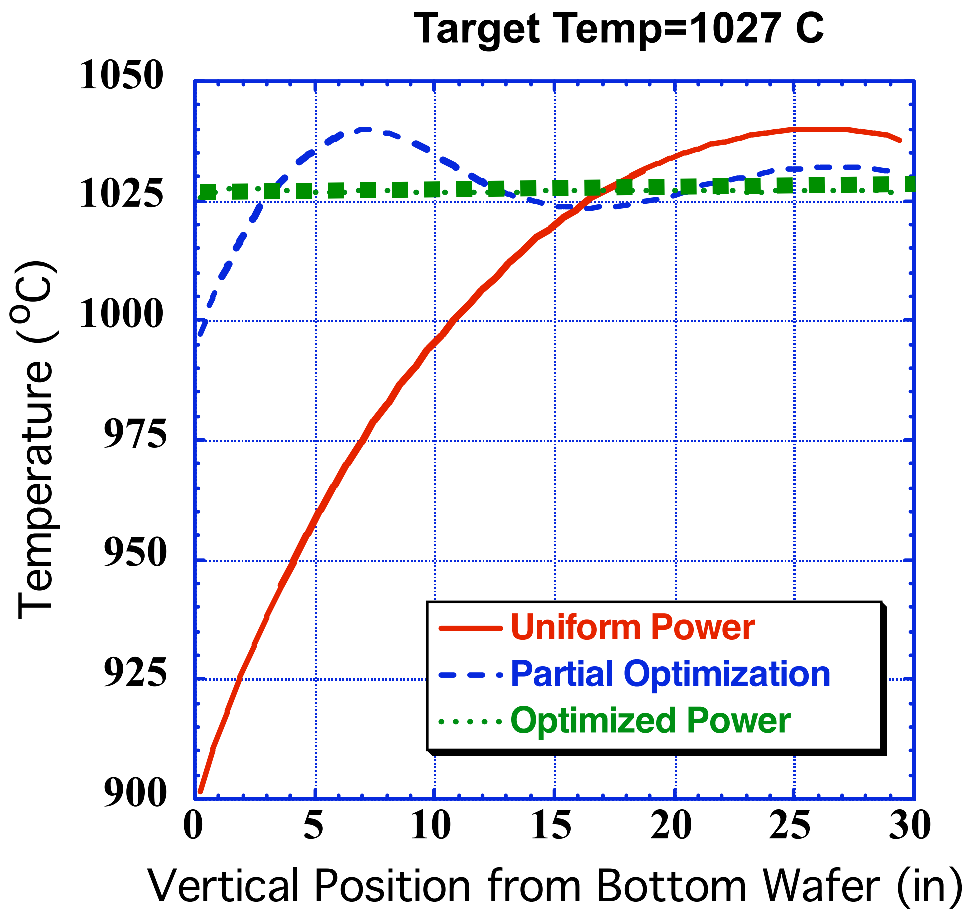
The design of a small-batch fast-ramp LPCVD furnace can be posed as an optimization problem



- Temperature uniformity across the wafer stack is critical
- Independently controlled heater zones regulate temperature
- Wafers are radiatively heated
- Design parameters:
  - Number of heater zones
  - Size / position of heater zones
  - Pedestal configuration
  - Wafer pitch
  - Insulation thickness
  - Baseplate cooling

Optimized power distribution enhances wafer temperature uniformity for steady-state operation



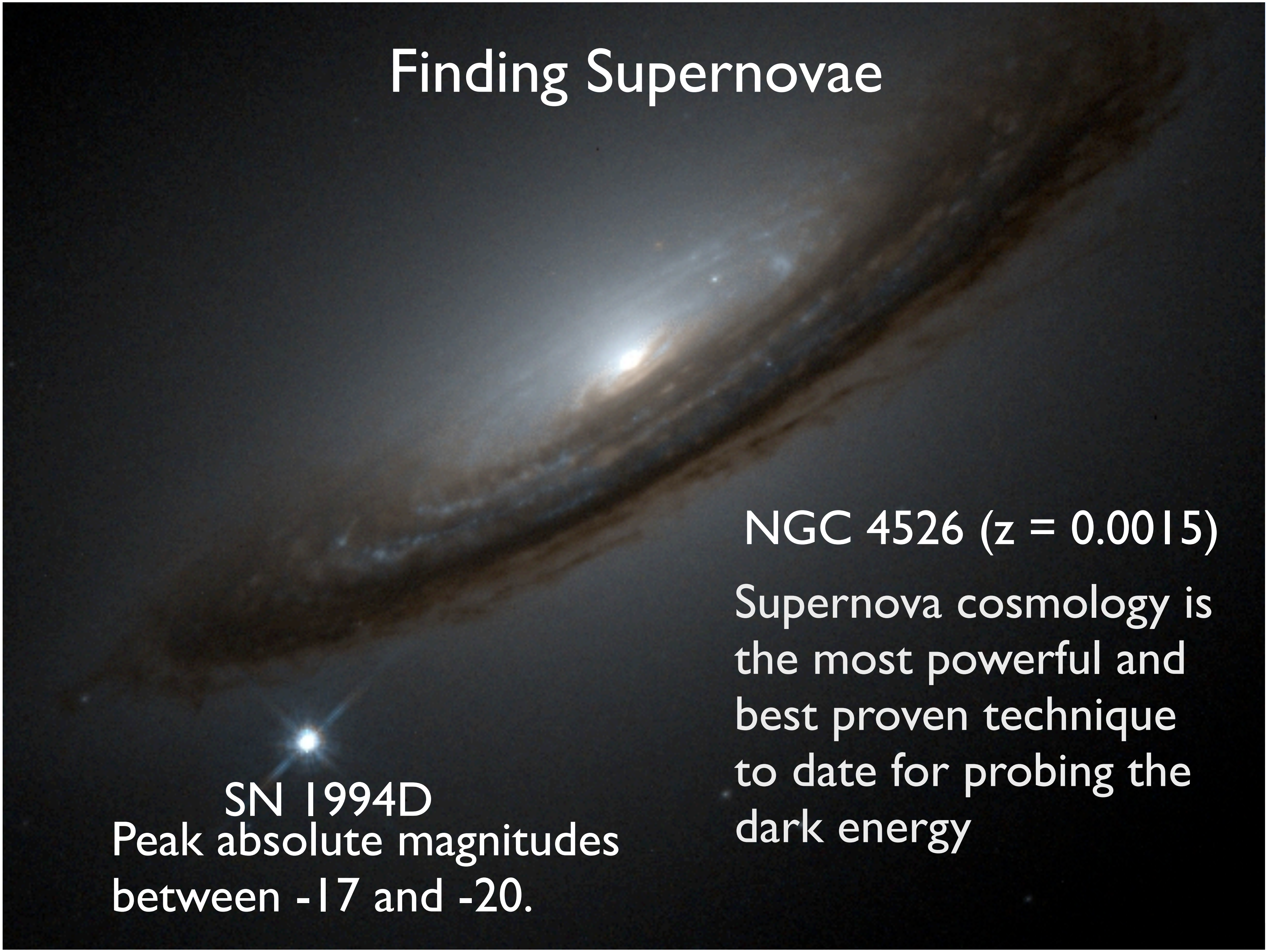


Simulation of Equipment Design Optimization in Microelectronics Manufacturing, **J.C. Meza, C.H. Tong, C.D. Moen, Proc. 30th Annual Simulation Symposium, Atlanta, GA, April 7-9, 1997.**

In Theory You Have Derivatives ...

- But in practice you don't
- Lack of derivatives reduces the choice of optimization methods one can use
- Accurately computing derivatives much harder than one might think at first glance
- Nonetheless, optimization loop turned a week long job into a 30 minute computation and with a solution that was an order of magnitude better

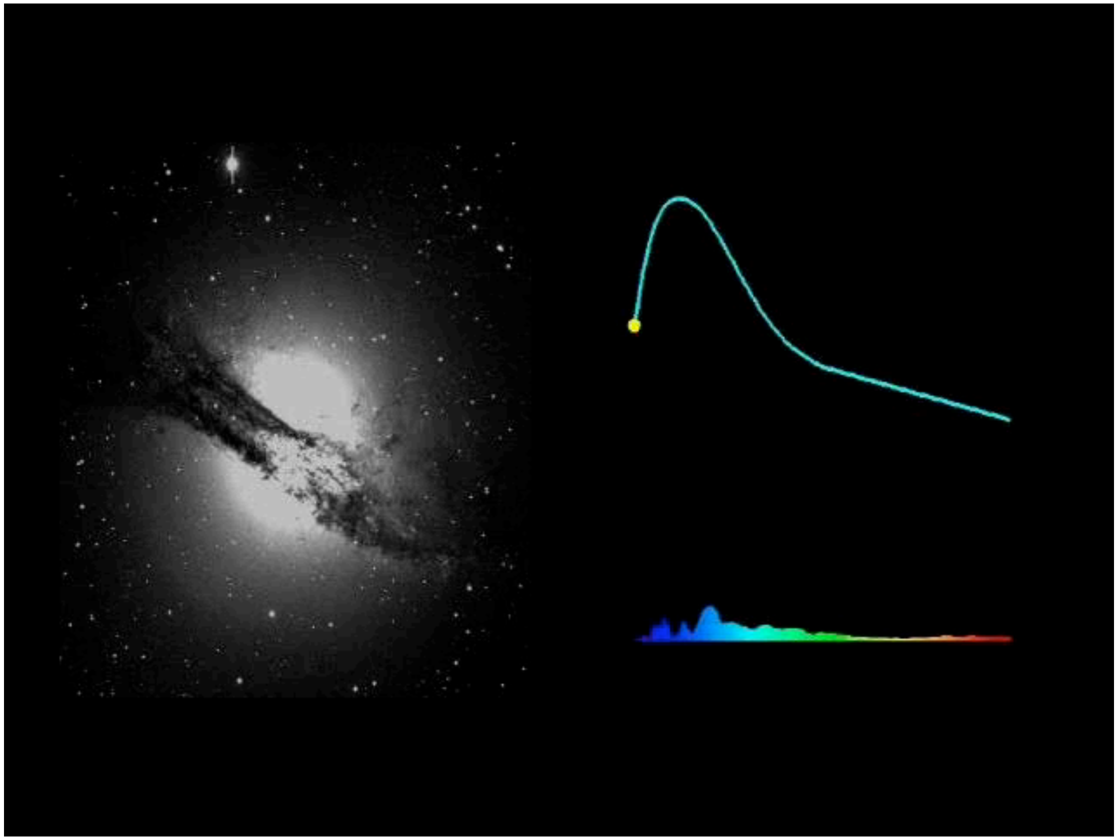
Finding Supernovae



Supernova cosmology is the most powerful and best proven technique to date for probing the dark energy

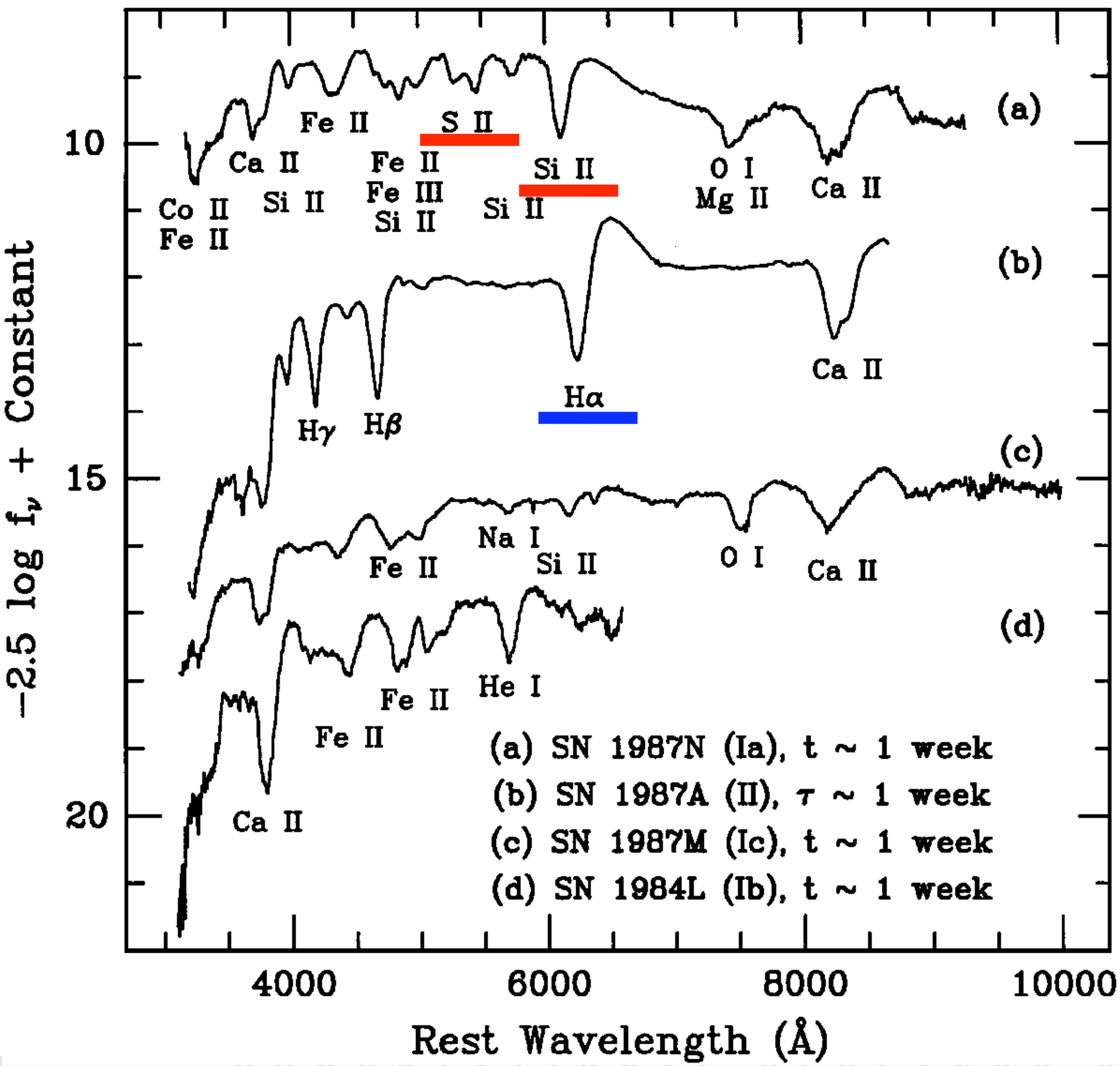
Peak absolute magnitudes between -17 and -20.

Supernova Spectra



- Most SN reach maximum light a few weeks after outburst.
- Light curves evolve on day to week time scales.
- Fade away over months to years.

Spectra



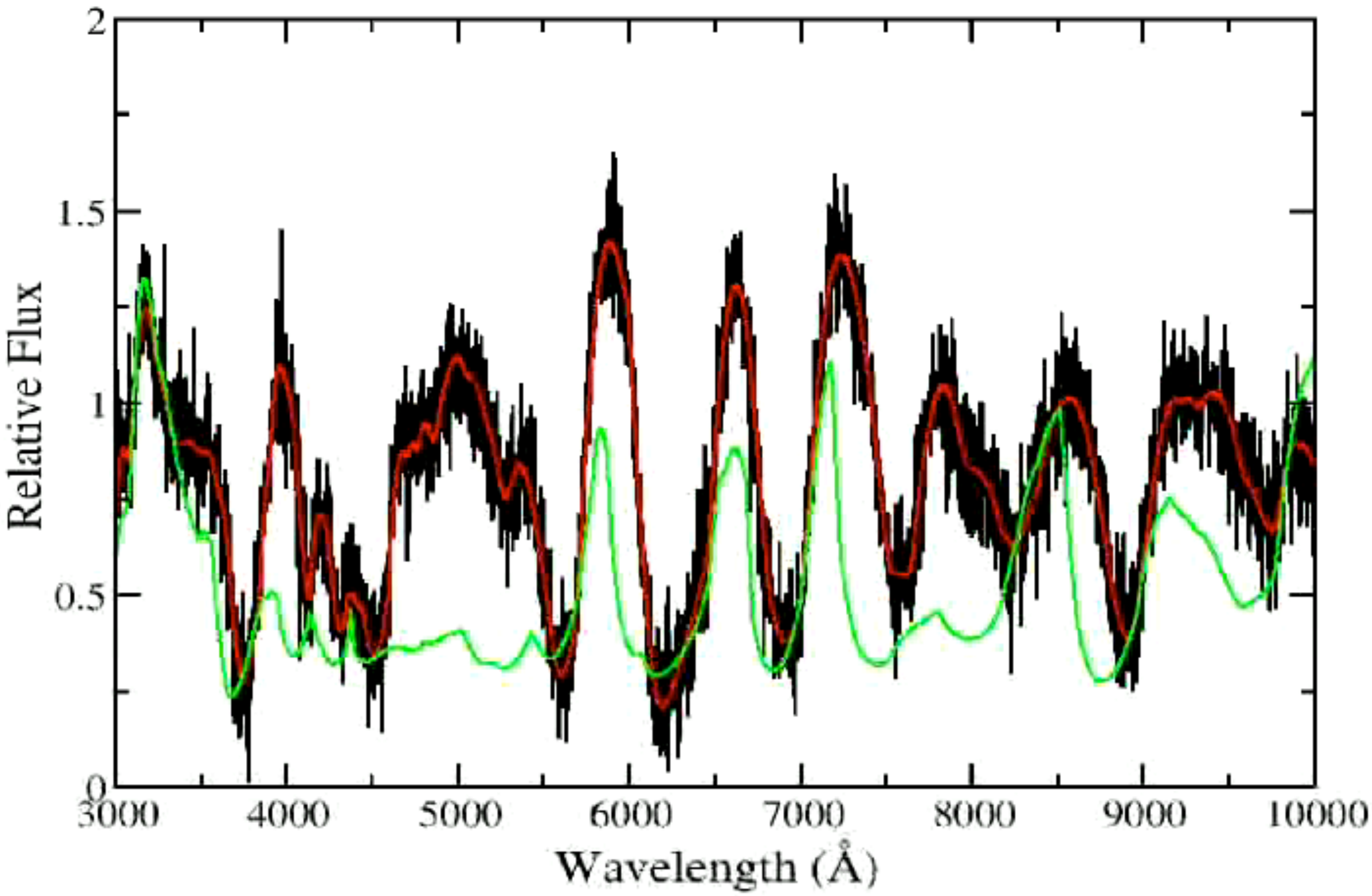
- Composition
- Density
- Temperature
- Energy deposited
- Kinetic energy
- Environment

Spectra indicative of rapid ejecta outflow up to 0.1c.

Problem is a least squares problem

- We have a simulation code that can produce synthetic spectra based on inputs that describe the elements that are present and their energies
- Goal is to fit the synthetic spectra to the observed data by varying the elements
- Must have an solution within 24 hours
- Did I forget to mention … we don't have derivatives

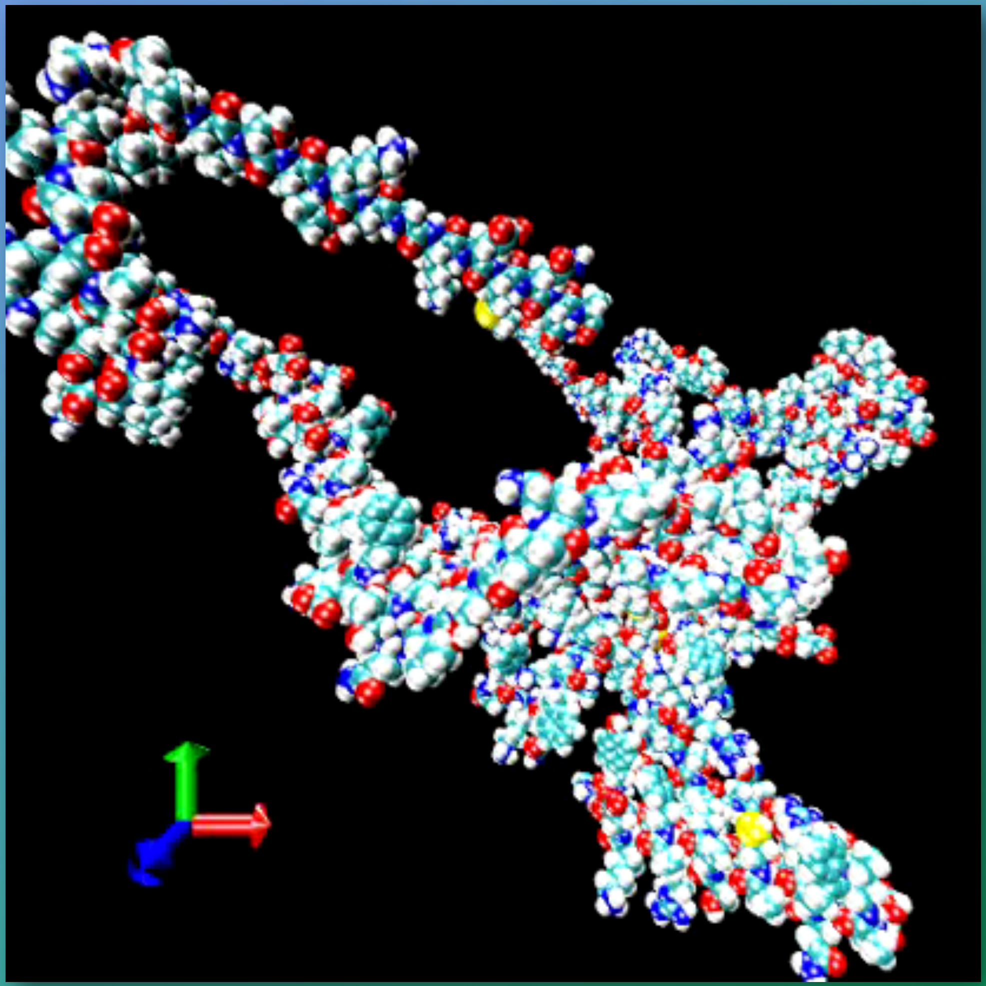
10 ions (52 parameters)



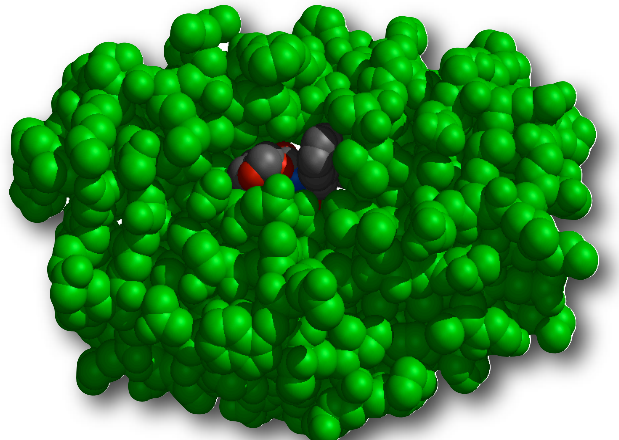
In Theory The Functions are Smooth ...

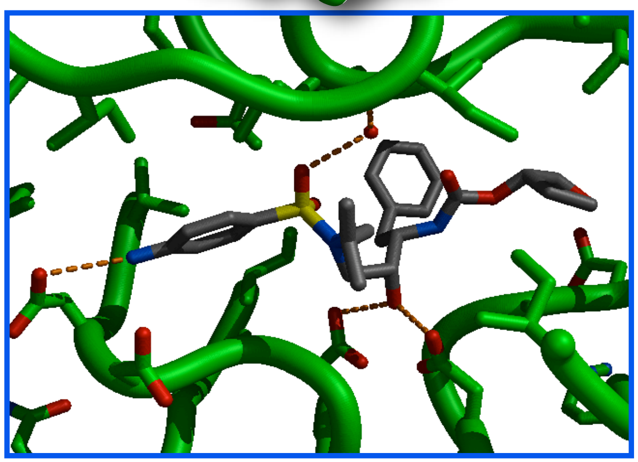
- But in practice the functions and data are noisy
- Need to account for uncertainty in both the model and the data
- “Best fit” can be a hard concept to handle in many scientific problems

Molecular Conformation



Drug design can be viewed as an optimization problem in computational chemistry





HIV-1 Protease Complexed with Vertex drug VX-478

- Formulated as an energy minimization problem
- Typically there are thousands of parameters with thousands for constraints
- There are many (thousands) of local minimum

Novel Applications of Optimization to Molecule Design, T.D. Plantenga, R.S. Judson, J.C. Meza, IMA Series "Large Scale Optimization with Applications, Part III", Vol. 94, 1997, Springer.

Amber Force Field Equations

$$E_{total} = E_{bond} + E_{angle} + E_{dihedral} + E_{nonbond}$$

$$E_{bond} = \sum K_b (r_i - \bar{r_i})^2$$

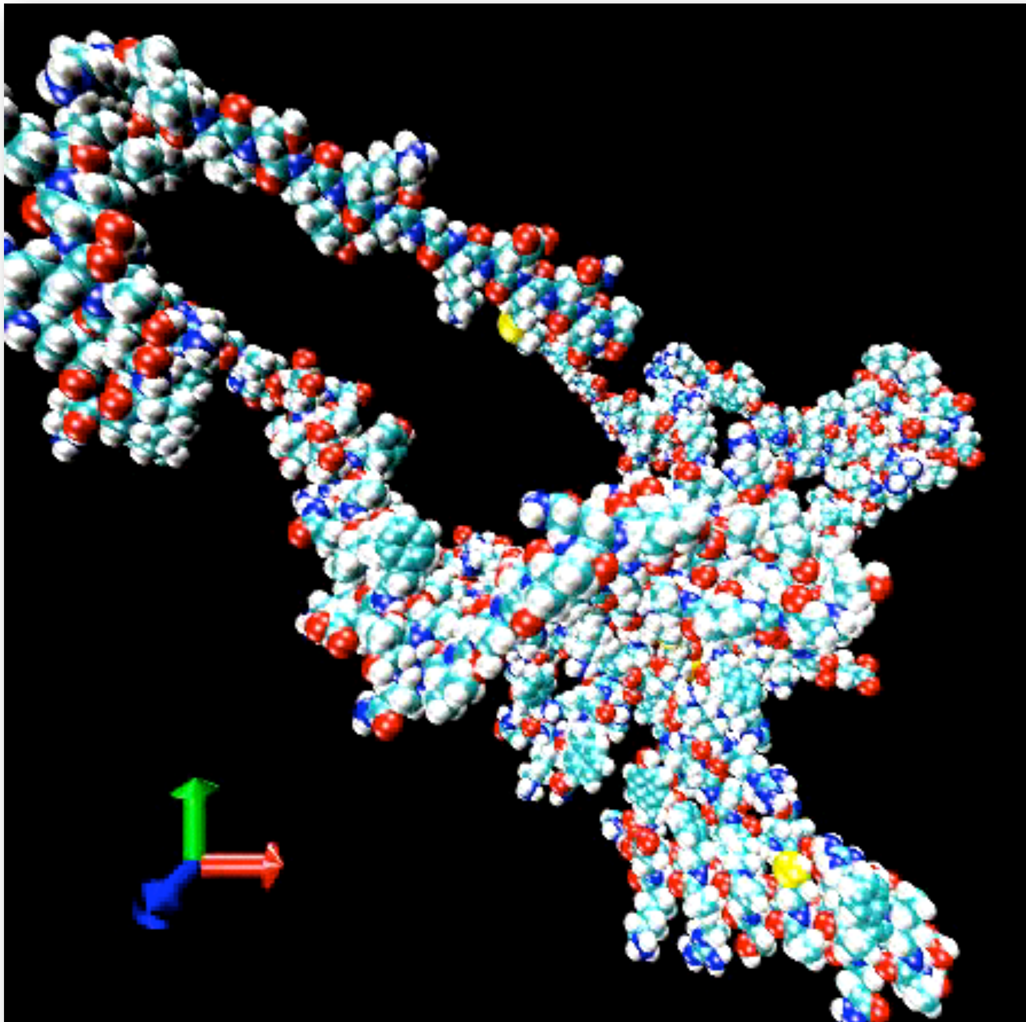
$$E_{angle} = \sum K_\theta (\theta_i - \bar{\theta}_i)^2$$

$$E_{dihedral} = \sum K_\phi (1 + cos(n_i \phi_i - \delta_i))$$

$$E_{nonbond} = \sum_i \sum_{i<j} \left[\epsilon_{ij} \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\epsilon_0 r_{ij}}$$

A Physical Approach to Protein Structure Prediction, Crivelli, et.al. Biophysical Journal, Vol 82, 2002.

Simulation of Bio-Molecule Folding



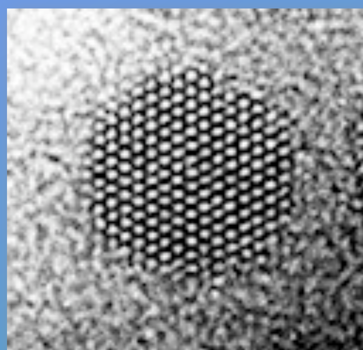
- Energy minimization computed using a large-scale optimization method
- Solution matched experimental data to within 3.9 Å
- Total simulation took approximately 32 hours on a workstation

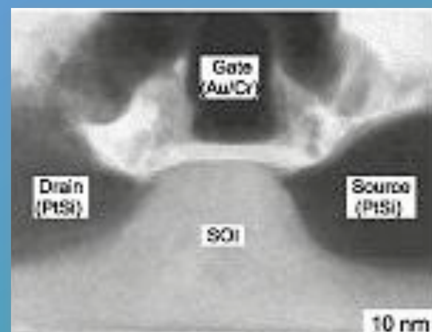
R. Oliva, S. Crivelli and J. Meza, *Protein-Folding Via Divide and Conquer Optimization." SIAM, Conference on the Life Sciences,* Portland, OR, July 11-14, LBNL-55869, 2004.

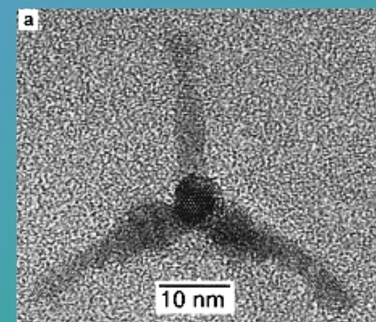
In Theory You Look for 1 or 2 Minima ...

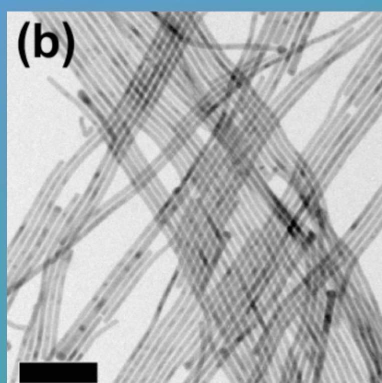
- But in practice problems may have many minima
- Users want "global" minima
- Sometimes only require descent

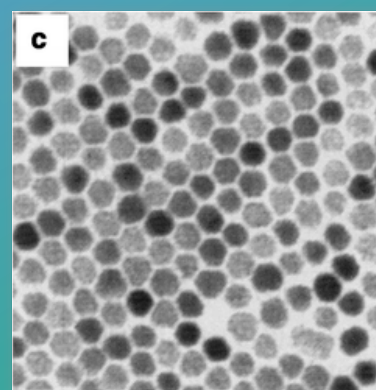
Computational Nanoscience

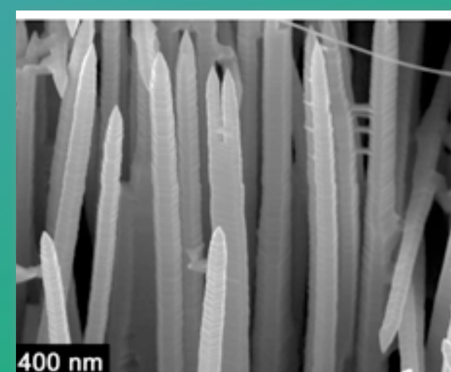




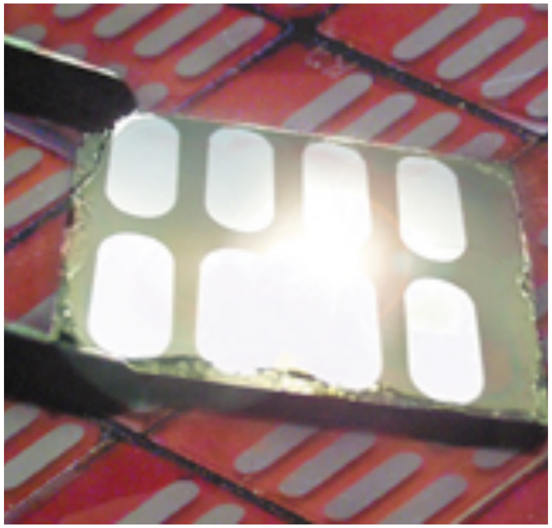


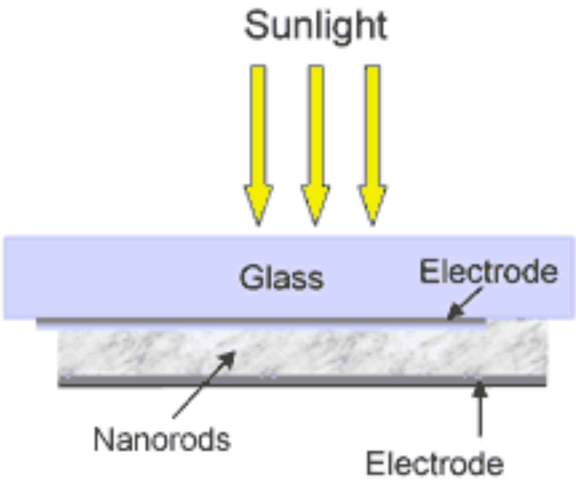






Photovoltaic Solar Cells





- Solar cells based on inorganic nanorods and semiconducting polymers
- Nanorods can be made of CdSe, a semiconducting material
- Nanorods act like wires, absorbing light and generating hole-electron pairs
- Biggest challenge is cost, ~30 cents/kWh

Density Functional Theory and the Kohn-Sham equations

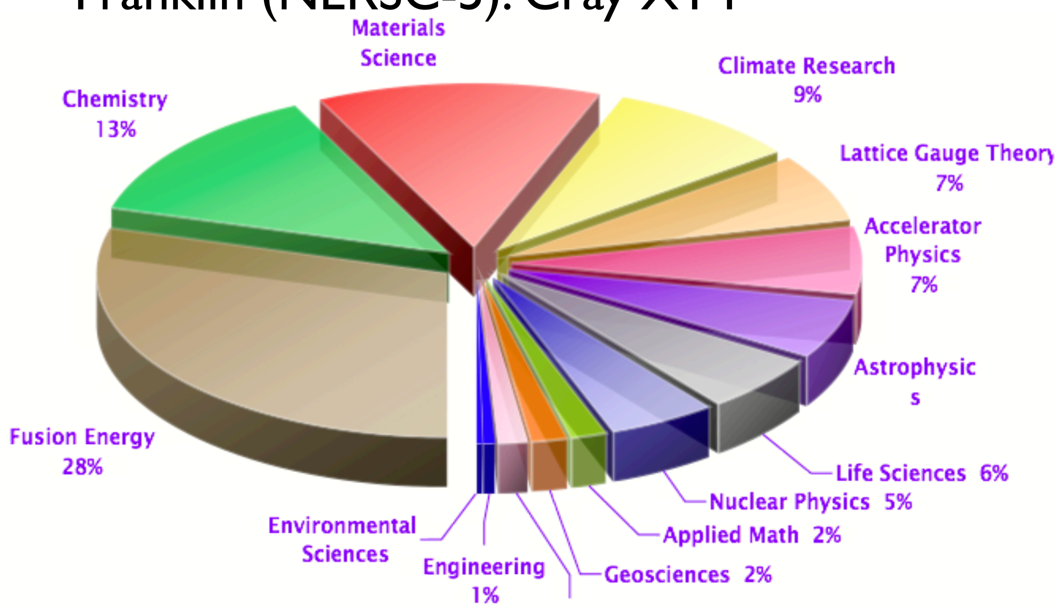
$$E_{total}[\{\psi_i\}] = \frac{1}{2}\sum_{i=1}^{n_e}\int_\Omega |\nabla\psi_i|^2 + \int_\Omega V_{ext}\rho + \frac{1}{2}\int_\Omega \frac{\rho(r)\rho(r^{'})}{|r-r^{'}|}drdr^{'} + E_{xc}(\rho),$$

$$\rho = \sum_{i=1}^{n_e} |\psi_i(r)|^2, \int_\Omega \psi_i\psi_j = \delta_{i,j}$$

$$\left[-\frac{1}{2}\nabla^2 + V_{ext}(r) + \int \frac{\rho}{|r-r^{'}|} + V_{xc}(\rho)\right]\psi_i = \epsilon_i\psi_i$$

DFT codes play a major role in computational science

Franklin (NERSC-5): Cray XT4



- DFT methods account for 75% of the materials sciences simulations at NERSC/LBNL
- Estimated total: 200 million hours of computer time in 2008
- 9,660 nodes; 38,640 2.3 GHz cores
- 356 TFlops/s peak

Solving the Kohn-Sham equations

- Self-Consistent Field (SCF) iteration
  - view as a (large) linear eigenvalue problem
  - usually used with other acceleration techniques to improve (ensure) convergence
  - convergence theory sparse
- Direct Constrained Minimization
  - minimize the total energy directly
  - also requires globalization techniques
  - convergence theory still difficult, but can use existing frameworks

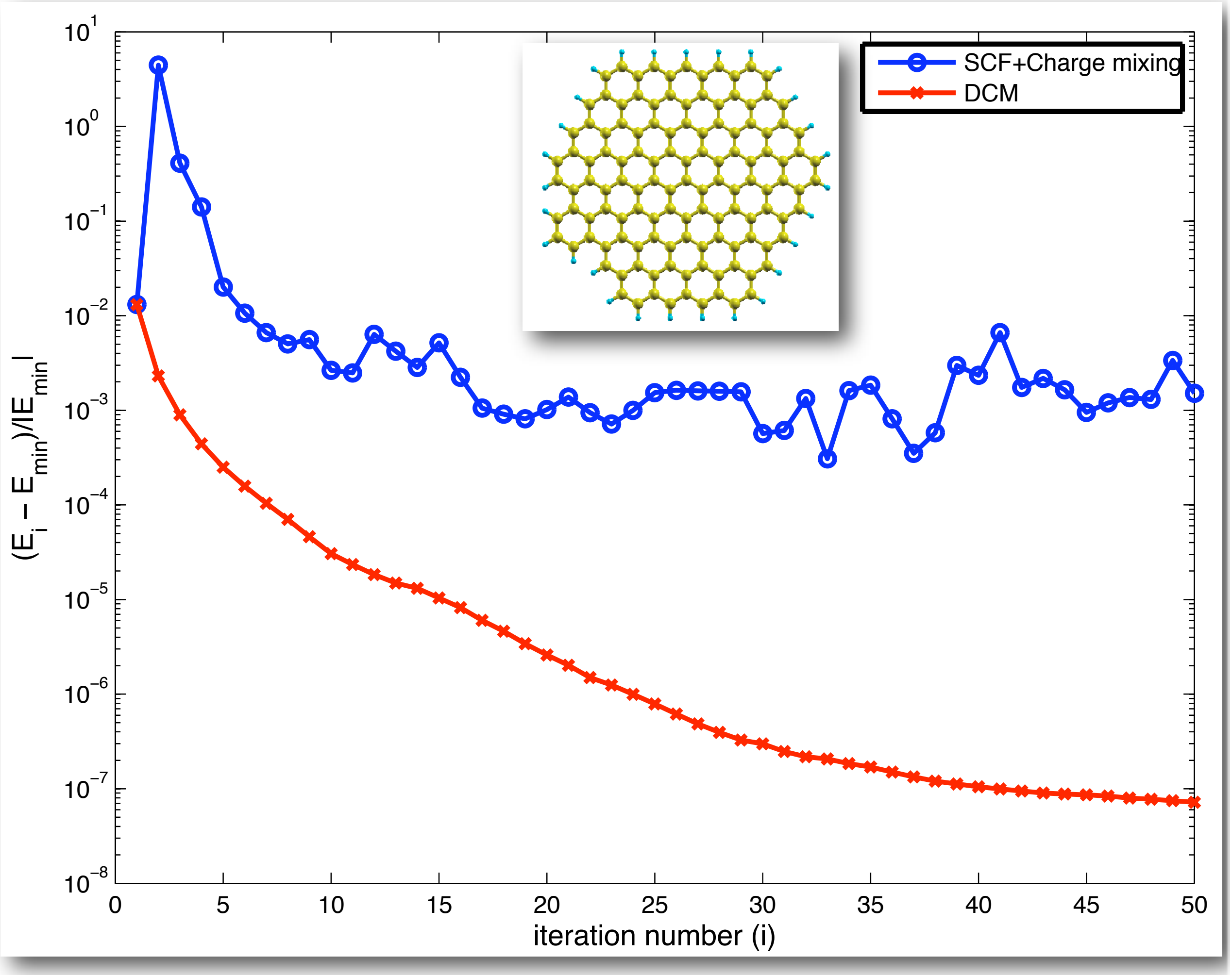
Direct Constrained Minimization

- Trick is to solve a sequence of smaller nonlinear eigenvalue problem
- Minimize energy in a particular subspace
- Construct minimization problem so that constraints are satisfied automatically
- Add trust region to ensure robustness
- Resulting method is more robust and faster (especially for larger systems)

C. Yang, J. Meza, L. Wang, A Constrained Optimization Algorithm for Total Energy Minimization in Electronic Structure Calculation, J. Comp. Phy., 217 709-721 (2006)

Example 2: Graphene

- sampling grid:
  - 114 x 114 x 15
- 10 PCG iterations / SCF outer iteration
- 5 inner SCF iteration / DCM outer iteration
- supercell:
  - 40 x 40 x 5



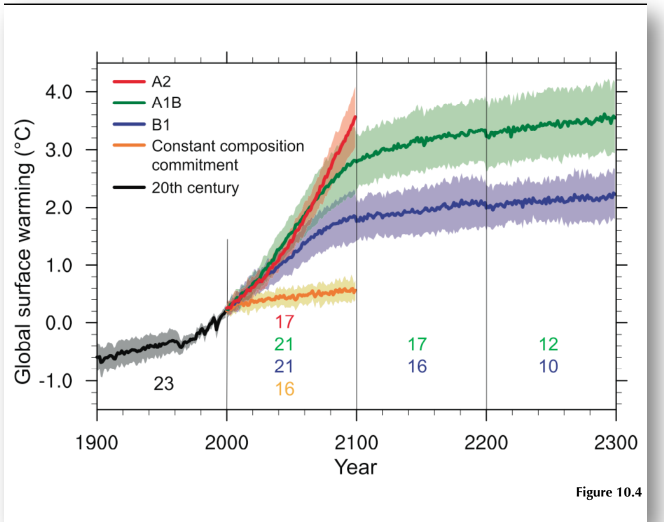
In Theory The Functions are Inexpensive ...

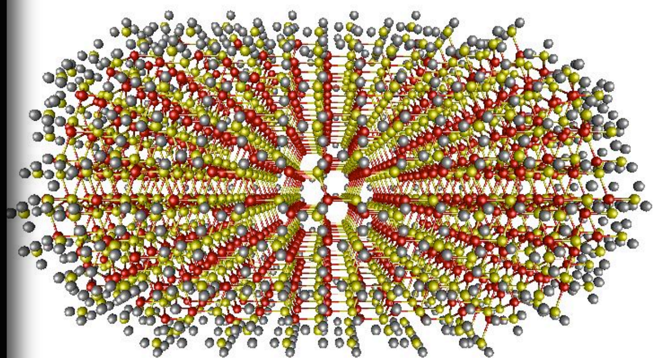
- But in practice the functions can dominate the total cost of the solution
- May not matter how expensive the optimization is as long as you can minimize the total number of function calls
- Sometimes convergence is dictated by how big your computer budget is

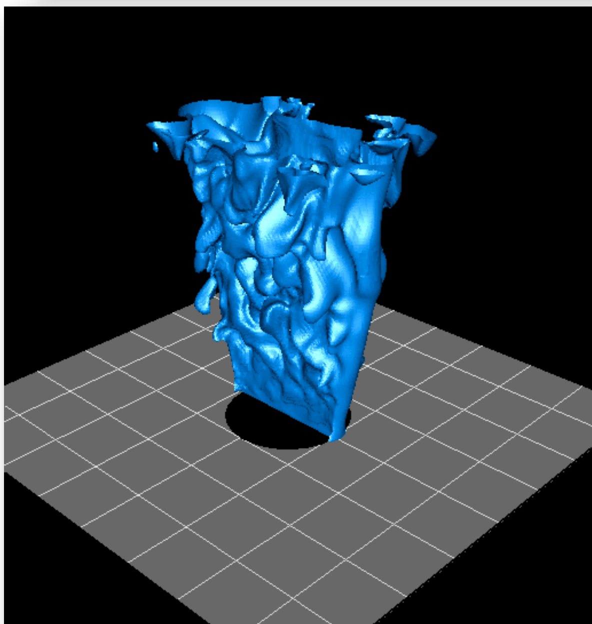
Summary - The Difference Between Theory and Practice ...

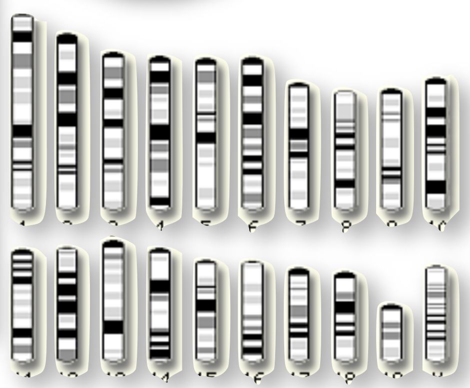
- Theory gives us the framework for analyzing our problems and guide our solutions
- Practice gives us the experience of real-world problems and helps us improve on the theory
- The combination of the two gives us unprecedented power for solving problems
- We are in a unique position to contribute solutions to problems in energy, health, climate, and many others

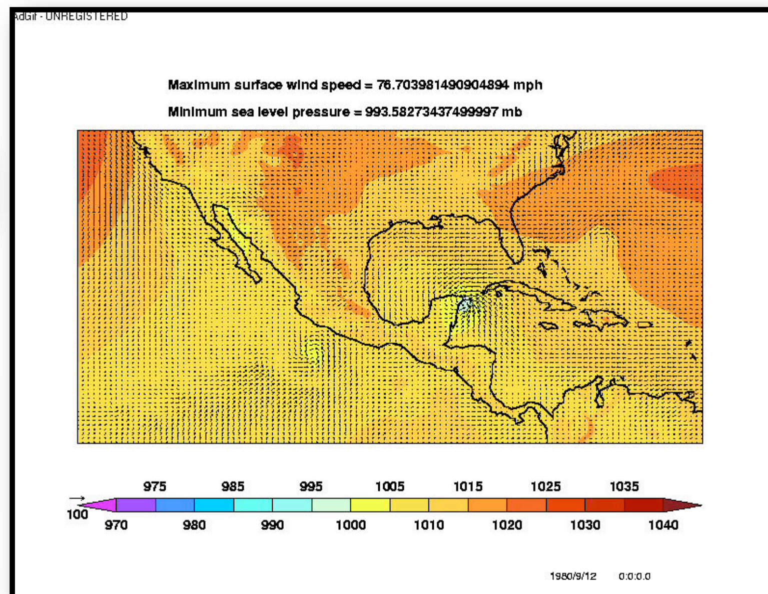
Future Directions


A2
A1B
B1
Constant composition commitment
20th century
Global surface warming (°C)
4.0
3.0
2.0
1.0
0.0
-1.0
23
17
21
21
16
17
16
12
10
1900
2000
2100
2200
2300
Year
Figure 10.4



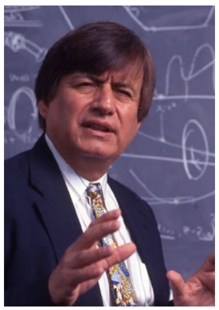





AdGif - UNREGISTERED
Maximum surface wind speed = 76.703981490904894 mph
Minimum sea level pressure = 993.58273437499997 mb
975
985
995
1005
1015
1025
1035
100
970
980
990
1000
1010
1020
1030
1040
1980/9/12
0:0:0.0

MSRI Undergraduate Program 2007

Formula for success

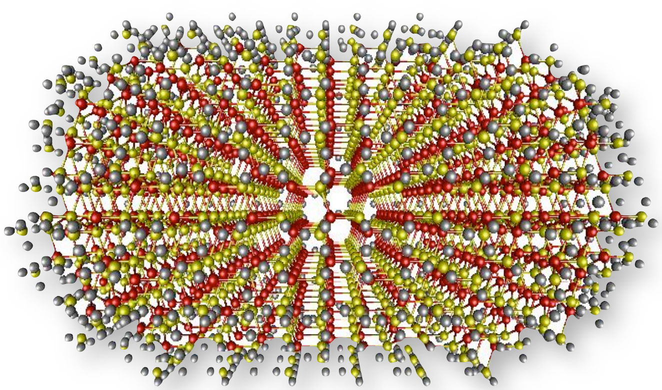


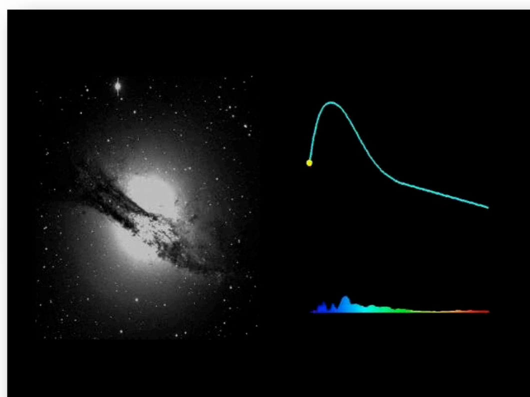
+

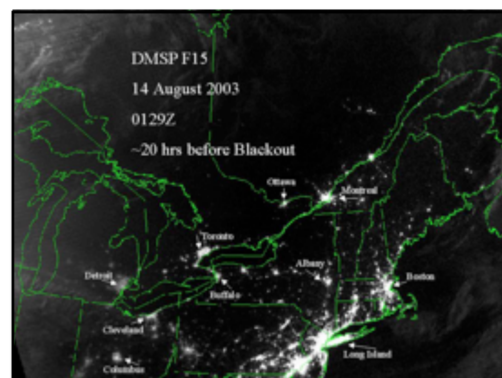
+

=

Solutions to …





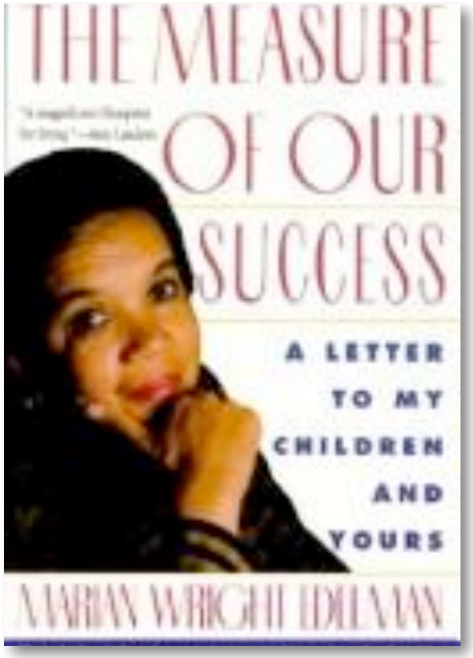


Thank you

The Measure of Our Success

With unprecedented economic competition from abroad and changing patterns of production at home that demand higher basic educational skills, America cannot wait another minute to do whatever is needed to ensure that today's and tomorrow's workers are well prepared rather than useless and alienated — whatever their color

Marian Wright Edelman, 1993



Questions