UNEDF Workshop
Pack Forest, Washington
June 22 – 25, 2009

Optimization Strategies for Complex UNEDF Simulations

Jorge Moré

Joint work with Jason Sarich and Stefan Wild

Mathematics and Computer Science Division
Argonne National Laboratory

Argonne NATIONAL LABORATORY

Complex Simulations

Simulation-based optimization problems are of the form

$$\min\left\{\mathcal{F}[s(x)] : x_L \leq x \leq x_U\right\},$$

where the mapping $s : \mathbb{R}^n \mapsto \mathbb{R}^m$ describes the simulation as a function of the parameters (or controls) x.

Challenges

- Expensive evaluations of $f(x) = \mathcal{F}[s(x)]$
- Noisy function evaluations
- Lack of derivatives with respect to parameters
- Possibly several minima
- Limited computational budget

Parameter Estimation Problems in UNEDF

- $x \in \mathbb{R}^n$ is the vector of parameters
- m is the number of nuclei.
- $f_k(x)$ is the vector of observables for the k-th nucleus.
- y_k is the data vector associated with the k-th nucleus.

$$x \quad \Longrightarrow \quad \boxed{\text{HFBTHO, MFD, \ldots}} \quad \Longrightarrow \quad f_k(x)$$

The least-squares approach (χ^2) requires the minimization of

$$f(x) = \frac{1}{2} \sum_{k=1}^{m} \sigma_k \| f_k(x) - y_k \|^2$$

where $\| \cdot \|$ is the l_2 norm and σ_k is a set of weights.

Research Issues

- What is the best algorithm for optimization?
- How do we recognize a minimizer?
- What is the sensitivity of the solution?
- Can we find a global minimizer?
- How do we validate the model?

Derivative-Free Optimization Algorithms

Stochastic Methods

- ◇ Simulated annealing algorithms
- ◇ Genetic algorithms
- ◇ Particle swarm algorithms

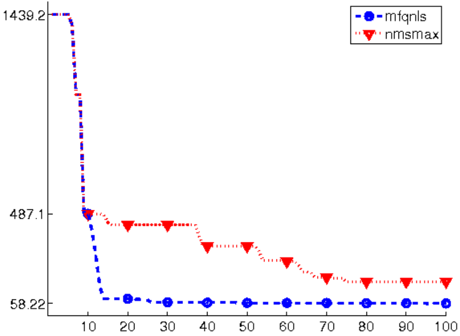
Direct Search Methods

- ◇ Nelder-Mead (nmsmax, nelder, fminsearch)
- ◇ Pattern search (appspack, mdsmax, sid-psm, nomad)

Model-Based Methods

- ◇ Quadratic models (uobyqa, newuoa, dfo)
- ◇ Radial-basis models (orbit, boosters)
- ◇ Quasi-Newton models(imfil, . . .)

Parameter Estimation with HFBTHO Code



HFBTHO code

63 nuclei

64 processors

$\mathrm{cost}\{f(x)\} = 12$ minutes

MFQNLS produces acceptable solutions after 2.6 hours, while the Nelder-Mead code NMSMAX has not converged after 20 hours.

Argonne NATIONAL LABORATORY

How do we recognize a minimizer?

The standard answer requires $\|\nabla f(x)\|$.

Standard Options

- Hand-coded gradients
- Difference approximations
- Automatic differentiation

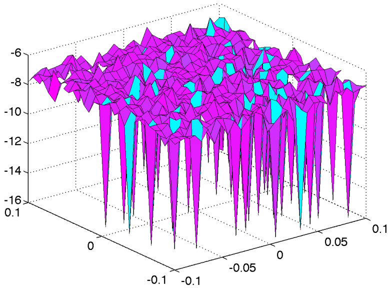
Approximating gradients can be difficult if f is noisy

Noisy Computations

The computed function $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is of the form

$$f(x) = f_s(x) + \varepsilon(x), \qquad x \in \Omega,$$

where $f_s : \mathbb{R}^n \mapsto \mathbb{R}^n$ is smooth and $\varepsilon : \mathbb{R}^n \mapsto \mathbb{R}^n$ is the noise.



Noise ε

Leading causes of noise

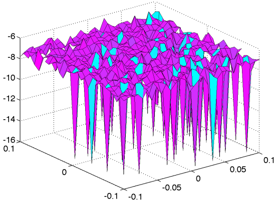
- Petaflops
- Iterative calculations
- Single precision

Partial Trace Eigenvalue Problem

Let A be a symmetric matrix, and define

$$f(x) = \sum_{i=1}^{k} \lambda_i(x)$$

where $\lambda(x)$ are the eigenvalues of $A + \mathrm{diag}(x)$ in increasing order.



Relative error when f is computed with **eigs** and $tol = 10^{-3}$

$n = 243$ and $k = 5$

Argonne NATIONAL LABORATORY

Computing Gradient Norms of Noisy Functions

Choose $t > 0$ and a random direction p.

Compute the directional derivative

$$\frac{f(x+tp) - f(x)}{t} \approx \langle \nabla f(x), p \rangle$$

Use the result that with high probability

$$|\langle \nabla f(x), p \rangle| \approx \|\nabla f(x)\|$$

⚠ The choise of t is delicate

Noisy Functions

Assume that the computed function $f : \mathbb{R} \mapsto \mathbb{R}$ is of the form

$$f(t) = f_s(t) + \varepsilon(t), \qquad t \in [t_a, t_b],$$

where $f : \mathbb{R} \mapsto \mathbb{R}$ is a smooth function, and $\varepsilon : \mathbb{R} \mapsto \mathbb{R}$ is the noise.

Assumption. The random variables $\{\varepsilon(t) : t \in [t_a, t_b]\}$ are independent and identically distributed.

Definition. The *noise level* of f is

$$\varepsilon_f = \left(\mathrm{Var}\{\varepsilon(t)\}\right)^{1/2}.$$

The Main Results

We have algorithms for the following problems:

- Determine the noise with a few function evaluations
- Determine optimal approximations to $\langle \nabla f(x), p \rangle$
- Determine $\|\nabla f(x)\|$

Theorem. If f_s is continuous at t, then

$$\lim_{h \to 0} \mathrm{Var}\left\{ \Delta f(t) \right\} = 2\, \varepsilon_f^2.$$

Computational Results

Case 1. $f(x) = x^T x(1 + \sigma \, \mathrm{randn})$ with $\sigma = 10^{-3}$

n_f	$\mathrm{mean}(\varepsilon_f)$
7	9.780e-04
9	1.024e-03
11	1.001e-03
13	9.888e-04
15	9.923e-04

Case 2. f is the partial trace eigenvalue problem

$$\varepsilon_f = 1.5 \, 10^{-8}$$

Argonne NATIONAL LABORATORY

Isomerization of α-pinene

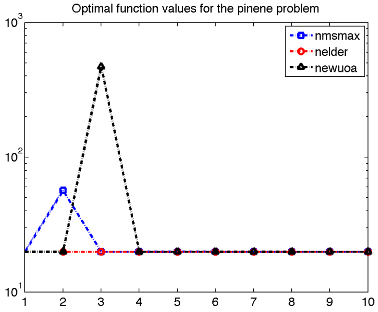
Determine the reaction coefficients in the thermal isomerization of α-pinene from measurements $z_1, \ldots z_8$ by minimizing

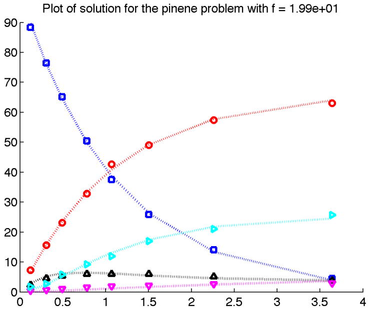
$$\sum_{j=1}^{8} \|y(\tau_j;\theta) - z_j\|^2$$

where $y(\cdot, \theta)$ satisfies

$$\begin{aligned} y_1' &= -(\theta_1 + \theta_2) y_1 \\ y_2' &= \theta_1 y_1 \\ y_3' &= \theta_2 y_1 - (\theta_3 + \theta_4) y_3 + \theta_5 y_5 \\ y_4' &= \theta_3 y_3 \\ y_5' &= \theta_4 y_3 - \theta_5 y_5 \end{aligned}$$

Isomerization of α-pinene





Argonne NATIONAL LABORATORY

Incompressible Elastic Rods

The shape of a rod, clamped at the origin and acted on by a vertical force α, a horizontal force β, and torque γ is described by

$$
\begin{aligned}
x_1'(s) &= \cos[\theta(s)] \\
x_2'(s) &= \sin[\theta(s)] \\
\theta'(s) &= \alpha x_1(s) - \beta x_2(s) + \gamma,
\end{aligned}
$$

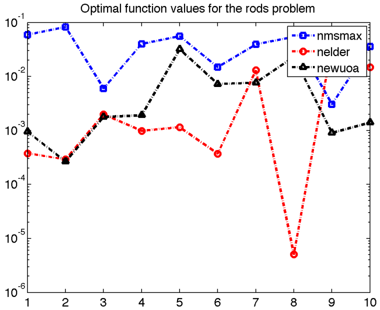
subject to the boundary conditions $x_1(0) = x_2(0) = \theta(0) = 0$, where θ is the angle of inclination, and s is arc length.

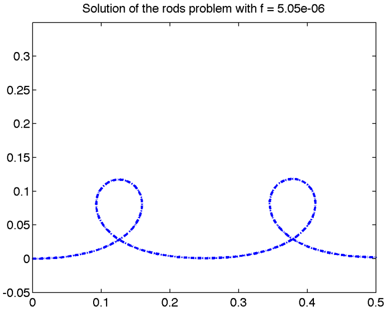
Determine the shape of the rod such that

$$x_1(1) = a, \quad x_2(1) = b, \quad \theta(1) = c,$$

for specified values of a, b and c.

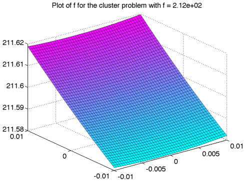
Incompressible Elastic Rods



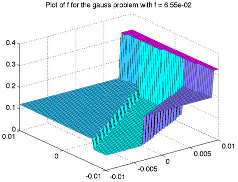


Argonne NATIONAL LABORATORY

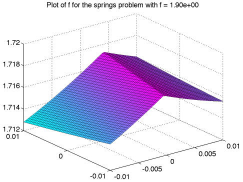
Slices



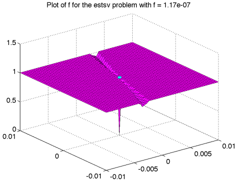
Plot of f for the cluster problem with f = 2.12e+02
211.62
211.61
211.6
211.59
211.58
0.01
0
-0.01
-0.01
-0.005
0
0.005
0.01



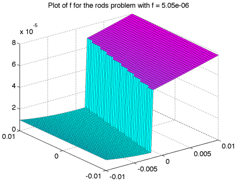
Plot of f for the gauss problem with f = 6.55e-02
0.4
0.3
0.2
0.1
0
0.01
0
-0.01
-0.01
-0.005
0
0.005
0.01



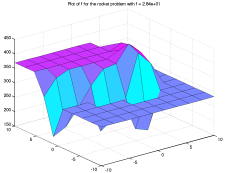
Plot of f for the springs problem with f = 1.90e+00
1.72
1.718
1.716
1.714
1.712
0.01
0
-0.01
-0.01
-0.005
0
0.005
0.01



Plot of f for the estsv problem with f = 1.17e-07
1.5
1
0.5
0
0.01
0
-0.01
-0.01
-0.005
0
0.005
0.01



Plot of f for the rods problem with f = 5.05e-06
x 10^-5
8
6
4
2
0
0.01
0
-0.01
-0.01
-0.005
0
0.005
0.01



Plot of f for the rocket problem with f = 2.84e+01
450
400
350
300
250
200
150
10
5
0
-5
-10
-10
-5
0
5
10

Future Work

Year 4

- Model and geometry-based optimization algorithms
- Open-source implementation of model-based algorithms
- Investigation of performance on new UNEDF functionals

Year 5

- Performance, evaluation, and validation of DFT functional
- Algorithms for noisy and constrained calculations
- Fission pathways