Applications of Nonlinear Optimization

GIAN Short Course on Optimization: Applications, Algorithms, and Computation

Sven Leyffer

Argonne National Laboratory

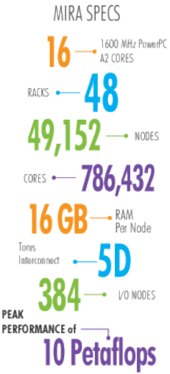
September 12-24, 2016

Outline

1. Introduction to Argonne National laboratory
2. Nonlinear Optimization for Power Grid Applications
   - Introduction to Power-Grid Applications
   - Optimal Unit Commitment for Power-Grid
3. Design of Complex Structures
   - Design of Nano-Structures
   - Topology and Structural Optimization
4. Analysis of Data and Images
   - X-ray Fluorescence Imaging
   - Multimodal Image Analysis
5. A Control Application: Optimal Transition to Clean Energy

Argonne National Laboratory

Department of Energy Laboratory (Open Science)



Two large user-facilities:

1. Advanced Photon Source
   ... ultra-bright X-ray source
2. Advanced Leadership Computing Facility
   ... Mira: 10 petaflops machine

Argonne National Laboratory

Argonne National Laboratory

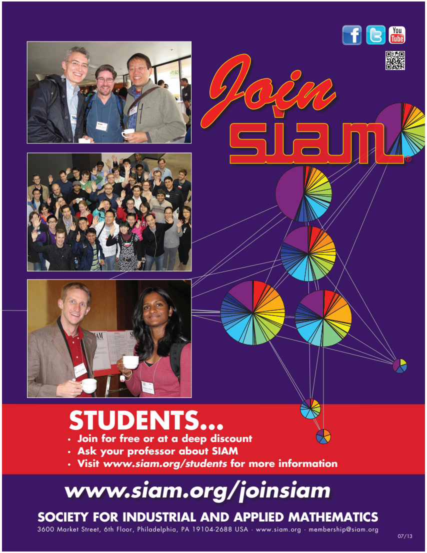
- Situated 25 miles SW of Chicago … site of CP-2
- 3,500 employees working in 14 research divisions
- Conduct basic research relevant to mission of DOE
  … used to be nuclear … now solar, wind, batteries, bio, …
- Mathematics and Computer Science has 110 staff & postdocs
- CS Research: Globus, MPICH, ZeptOS, …
- Math Research: Nek5000, NEOS, PETSc, …
- Student opportunities: summer interns (apply in January)
- Postdoc opportunities: e.g. Wilkinson Fellowship, late 2017

Math and CS at Argonne

Do you recognize this picture?

Math and CS at Argonne

UIC, Northwestern, Chicago SIAM student chapters visiting MCS Division

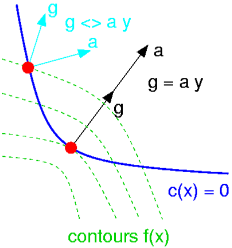


Create your own
SIAM Student Chapter
… and visit us too!

Today's Problem: Nonlinearly Constrained Optimization

Nonlinear Optimization Problems (NLPs)

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to } & c(x) \leq 0 \\ & x \in \mathcal{X}, \end{aligned}$$



where

- $\mathcal{X} \subset \mathbb{R}^n$ compact set, e.g. $\mathcal{X} = \{x \mid l \leq A^T x \leq u\}$ polyhedral
- $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^m$ smooth functions

More general problems ...

- $l_x \leq x \leq u_x$ simple bounds
- $l_c \leq c(x) \leq u_c$ more general constraints
- Classes of variables: binary, integer, semi-definite, ...
- Classes of constraints: DAE, PDE, complementarity, ...

Outline

1. Introduction to Argonne National laboratory
2. Nonlinear Optimization for Power Grid Applications
   - Introduction to Power-Grid Applications
   - Optimal Unit Commitment for Power-Grid
3. Design of Complex Structures
   - Design of Nano-Structures
   - Topology and Structural Optimization
4. Analysis of Data and Images
   - X-ray Fluorescence Imaging
   - Multimodal Image Analysis
5. A Control Application: Optimal Transition to Clean Energy

Challenges in Power Grid Modeling

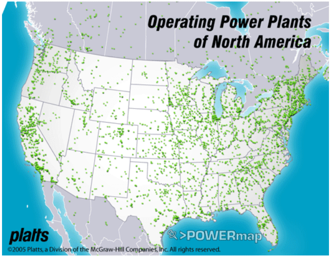
Computational Challenges in Power Grid Modeling

1. **Size:** $\simeq$ 100k lines ... "most complex machine ever built"
2. **Complexity:** nonlinear, hierarchical, and discrete decisions
3. **Uncertainty:** demand and supply (renewable) uncertainties

... many applications combine all three challenges

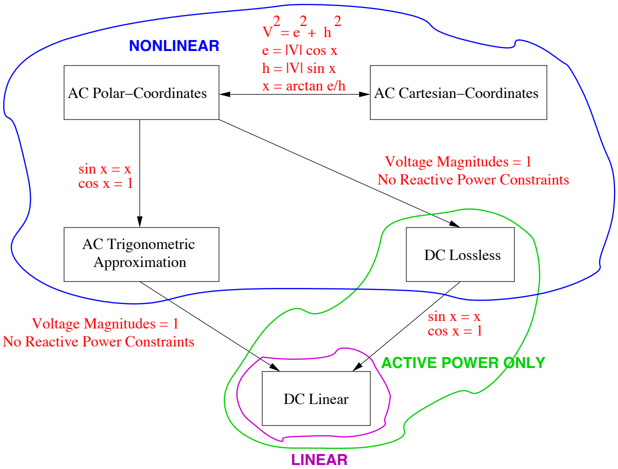
Missing from this talk:

- Big data
- Real-time decisions
- Cyber-security



... all involve modeling and computation

Complexity of Power Grid: Nonlinearities



- Operation & Design: optimal power flow, transmission switching, network expansion
- Challenge: interaction of nonlinearities & discrete decisions

Complexity of Power Grid: Discrete Decisions

- Given existing power grid network and demand forecast
- Design expanded network for secure transmission

Traditional Approach. Simplify nonlinear (AC) power flow model:

$$F(U_k, U_l, \theta_k, \theta_l) := b_{kl} U_k U_l \sin(\theta_k - \theta_l) + g_{kl} U_k^2 - g_{kl} U_k U_l \cos(\theta_k - \theta_l)$$

by setting $\sin(x) \simeq x$ and $\cos(x) \simeq 1$ and $U \simeq 1$

Nonlinear Optimization Approach. Work with nonlinear model

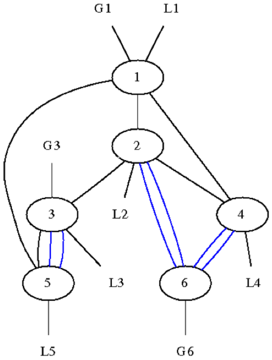
- $-M(1 - z_{k,l}) \leq f_{k,l} - F(U_k, U_l, \theta_k, \theta_l) \leq M(1 - z_{k,l})$
- $z_{k,l} \in \{0, 1\}$ switches lines on/off; $M > 0$ constant

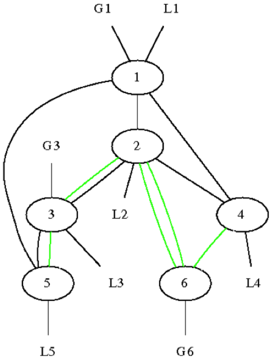
Questions.

Can we solve the nonlinear models? Do nonlinearities it matter?

Power-Grid Transmission Network Expansion

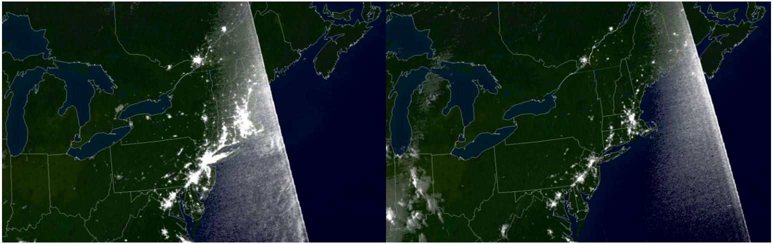
Expansion Results for linear vs. nonlinear power flow models





- Solve realistic AC power flow expansion models on desktop
- Significant difference between DC and AC solution
- Linearized DC model not feasible in AC power flow
- Linear DC model *not* valid when topology changes

Blackout Prevention in National Power Grid



2003 blackout: before and during

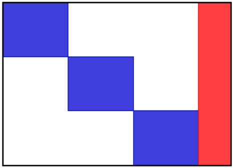
- 2003 blackout cost $4-10 billion and affected 50 million people
- prevent with contingency analysis
  - find least number of transmission lines whose removal results in failure
  - binary variables model removal of lines
  - nonlinearities model power flow
  - results in large integer optimization problem
- current analysis limited to 10s of lines

... similar models arise in many other power-grid applications

Unit Commitment with Wind Power [Cosmin Petra]

Wind uncertainty $\Rightarrow$ stochastic optimization: min. expected cost

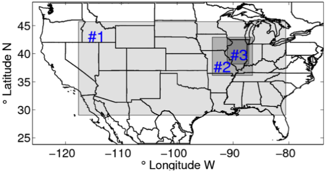
$$\begin{aligned}&\underset{x}{\text{minimize}} \quad f(x) + \mathbb{E}_\omega \left(\min_z \; h(x, z; \omega) \text{ s.t. } g(x, z; \omega) \geq 0 \right) \\ &\text{subject to } c(x) \geq 0\end{aligned}$$



- x — here-and-now decisions
- z — 2nd-stage decisions/scenarios
  … random realizations of wind
- $\omega \in \Omega$ random parameters

Realistic wind scenarios

- Weather Research Forecasting (WRF)
- Real-time grid-nested 24h simulation
- $|\Omega| = 30$ samples of WRF



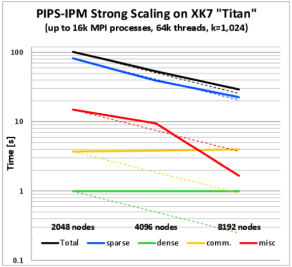
Stochastic Unit Commitment [Cosmin Petra]

PIPS - scalable framework for stochastic optimization problems

- Parallel distributed implementations of interior-point (IPM)
- Block-angular linear systems suitable to parallelization
- Schur complement-based decomposition of linear algebra
- Parallelization bottlenecks: dense linear algebra (first stage)
- Dense matrices can go on GPUs, multicores, or be distributed.

PIPS-IPM ported to IBM BG/P and BG/Q, Cray XE6, XK7 & XC30

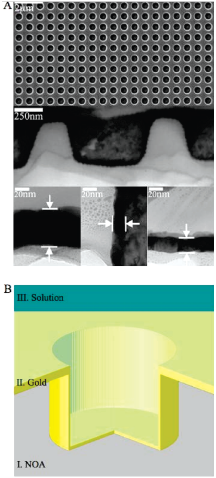
- 32k scenarios
- 4 billion variables and constraints
- 128K cores on BG/P and 64K cores on XK7



Outline

1. Introduction to Argonne National laboratory
2. Nonlinear Optimization for Power Grid Applications
   - Introduction to Power-Grid Applications
   - Optimal Unit Commitment for Power-Grid
3. Design of Complex Structures
   - Design of Nano-Structures
   - Topology and Structural Optimization
4. Analysis of Data and Images
   - X-ray Fluorescence Imaging
   - Multimodal Image Analysis
5. A Control Application: Optimal Transition to Clean Energy

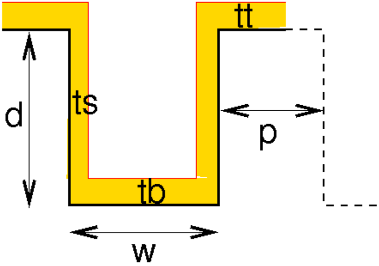
Design of Nano-Photonic Devices



- nano-structures for chemical sensing
  optical response at certain wave-lengths
- Top: scanning electron micrograph
  Middle: cross section of crystal
  Bottom: gold thickness
- 3D FEM analysis simulation; no gradients
  periodicity, gold thickness, depth, & width of nano-wells
- derivative-free optimization
  objective function evaluation takes 12 hrs
  ... on 125 nodes of an Apple G5 X-server
- optimization of a black-box
  ... simulation-based optimization
  ... derivative-free optimization

Inside the Black Box

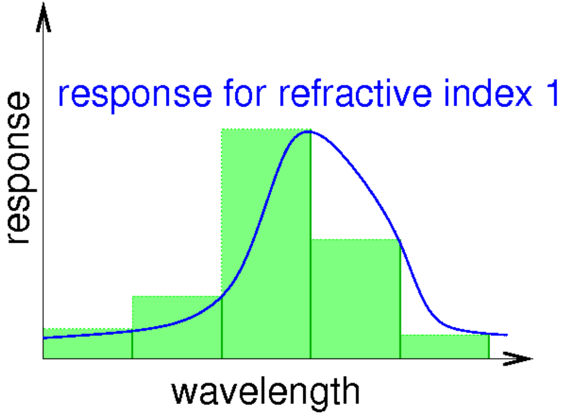
Design Parameters:



- d = depth of nano-well
- p = periodicity of design
- w = width/diameter of nano-well
- t = thickness (side/bottom/top) gold layer

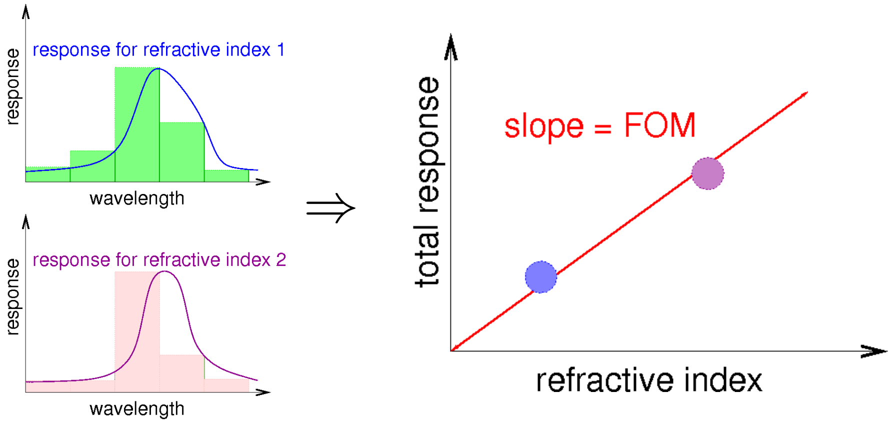
Inside the Black Box

Given values of design parameters (d, p, w, ts, tb, tt)
... perform PDE simulation for refractive index



... get total response

Objective = Figure of Merit (FOM)



... combine responses for different refractive indices

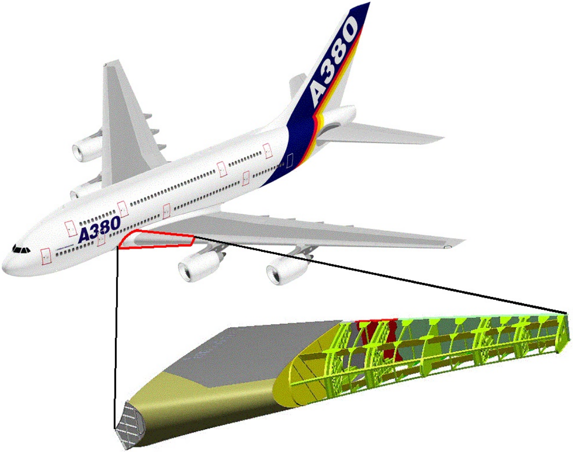
... maximize slope (sensitivity) of design

... requires solution of many PDEs

Outline

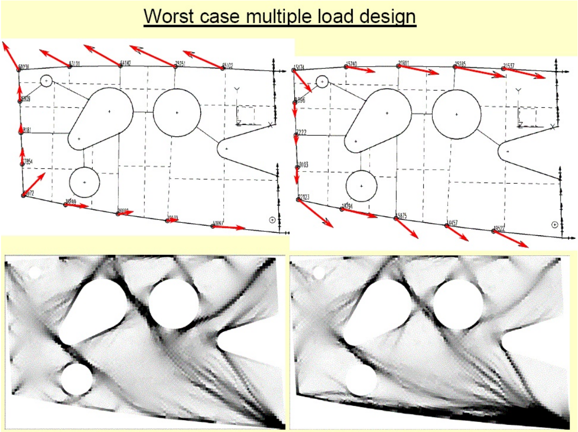
1. Introduction to Argonne National laboratory
2. Nonlinear Optimization for Power Grid Applications
   - Introduction to Power-Grid Applications
   - Optimal Unit Commitment for Power-Grid
3. Design of Complex Structures
   - Design of Nano-Structures
   - Topology and Structural Optimization
4. Analysis of Data and Images
   - X-ray Fluorescence Imaging
   - Multimodal Image Analysis
5. A Control Application: Optimal Transition to Clean Energy

Example: Topology & Structural Optimization [Kocvara]



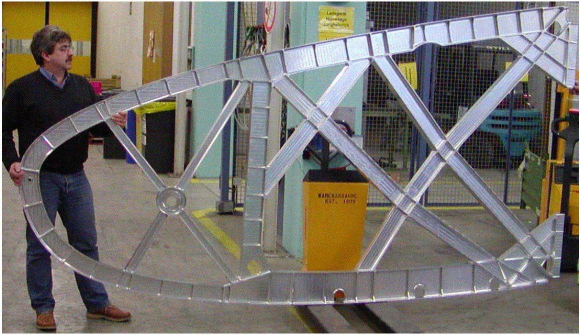
Airbus A-380 inboard inner leading edge ribs

Example: Topology & Structural Optimization [Kocvara]



Minimize weight of structure subject to load & design restrictions

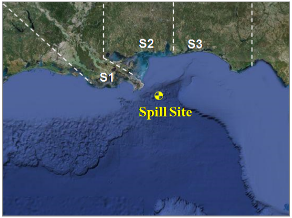
Example: Topology & Structural Optimization [Kocvara]

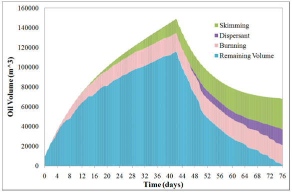


Optimized topology reduces weight by 33%

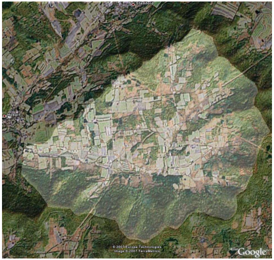
Other Optimization Applications

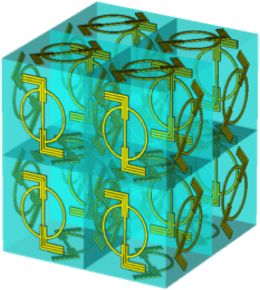
- Optimization of oil-spill response ... MIP control
- Design and control of chemical processes
- Design and operation of building energy systems
- Design of wind farms & dispatch of wind energy
- Image analysis, inverse problems, big data, ...
- Cybersecurity: resource allocation & attacker-defender models
- Fighter intercept optimal control problems

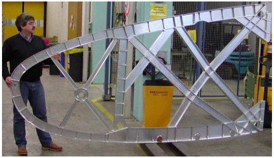




Optimization of Complex Systems







Take-Home Message: More complex optimization problems!

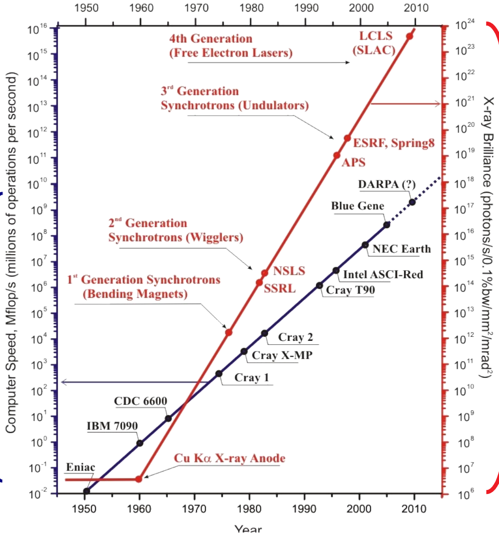
- Complex science & engineering apps drive optimization:
  - Discrete decision give rise to integer variables
  - Complex physics modeled using ODEs, DAEs, or PDEs
  - Uncertainty quantification for robust decisions & design
- Many open problems in need of new methods and ideas!

Outline

1. Introduction to Argonne National laboratory
2. Nonlinear Optimization for Power Grid Applications
   - Introduction to Power-Grid Applications
   - Optimal Unit Commitment for Power-Grid
3. Design of Complex Structures
   - Design of Nano-Structures
   - Topology and Structural Optimization
4. Analysis of Data and Images
   - X-ray Fluorescence Imaging
   - Multimodal Image Analysis
5. A Control Application: Optimal Transition to Clean Energy

Light source data growing faster than Moore's law

Computers: 12 orders of magnitude in 6 decades



Light sources: 18 orders of magnitude in 5 decades

Data Analysis: X-ray Imaging the APS

SL, Stefan Wild, Siwei Wang, & Stefan Vogt (APS)

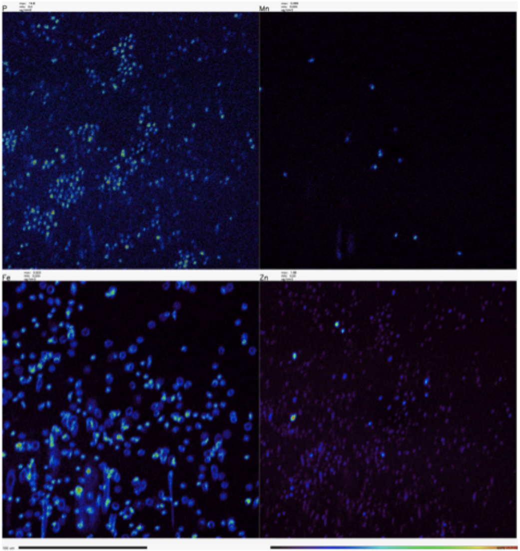
Science Challenges in Nano-Medicine & Theranostics:

- Design new treatment and drugs for targeted drug delivery
  Combine therapy & diagnostics by targeting nano-particles
- Extract efficiency score from multiple sources of data
  MRI of living specimen; X-ray, fluorescent & visible light

Specific Questions:

- Given a sample, how many elements/cells are in the sample?
- Identify abnormal concentrations: cancer vs. healthy cells.
- Identify marker elements (e.g. Ti) for tracking biochemical processes.

Typical Image: Postdoc Marks Cells By Hand

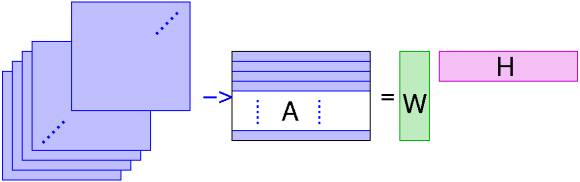


Concentration maps of P, Mn, Fe, Zn ... where's the yeast/algae?

Nonnegative Matrix Factorization

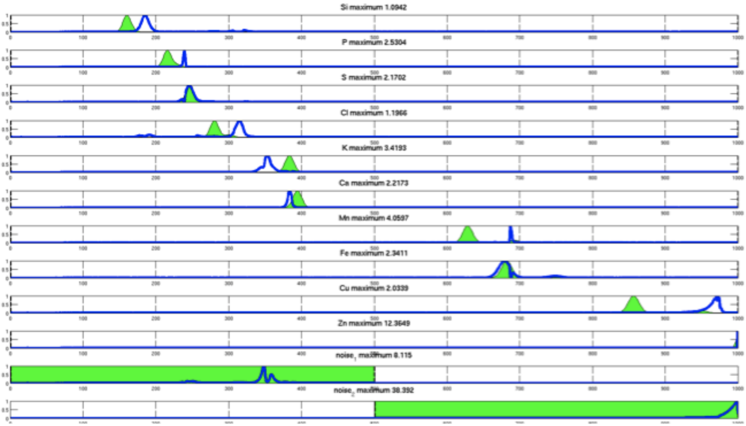
Goal: Find signature masks for cells for image segmentation

- Photon counts are additive $\Rightarrow$ use additive reduction
- $\underset{W,H}{\text{minimize}} \; \|A - WH\|_F^2 \quad$ subject to $W \geq 0, \; H \geq 0$



- W are weight $\simeq$ additive elemental spectra
- H are images $\simeq$ additive elemental maps
- Solve using (cheap) gradient steps

Nonnegative Matrix Factorization



Initialization of W and final value.

… now use H (additive elemental maps) to segment images

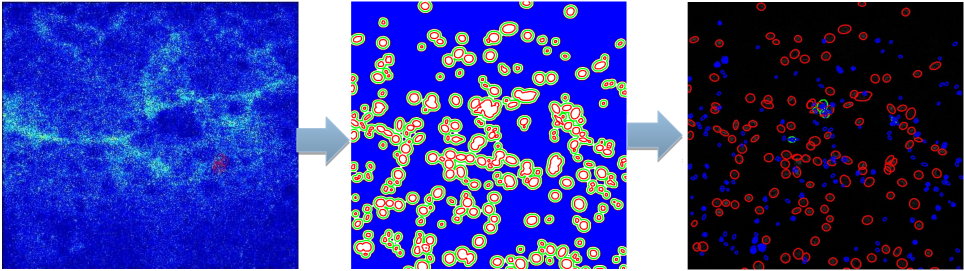
XFM@APS: Automatic Cell Discovery & Classification

Enables *fast* processing of larger samples, reduces error, automatically classifies cells

- Pre-processing with **non-negative matrix factorization**
- Image partitioned using **spectral graph partitioning**
- **Wavelet transforms** delineate curves and allow for overlapping cells
- Incorporate domain-specific knowledge with **large-scale MINLP optimization** approach

Noisy, Multichannel Data

Yeast, **Algae**, **Red Blood** Cells

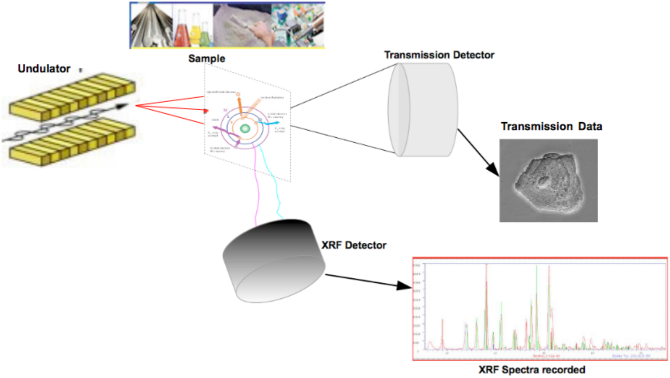


Outline

1. Introduction to Argonne National laboratory
2. Nonlinear Optimization for Power Grid Applications
   - Introduction to Power-Grid Applications
   - Optimal Unit Commitment for Power-Grid
3. Design of Complex Structures
   - Design of Nano-Structures
   - Topology and Structural Optimization
4. Analysis of Data and Images
   - X-ray Fluorescence Imaging
   - Multimodal Image Analysis
5. A Control Application: Optimal Transition to Clean Energy

Multimodal Image Analysis

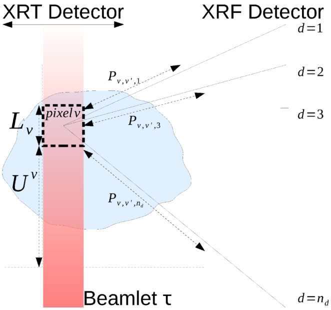
Next-generation accelerators capture multiple data modalities:
e.g. transmission (XRT) and fluorescence (XRF) in single shot

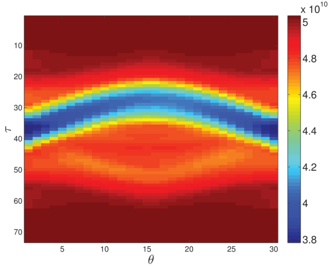


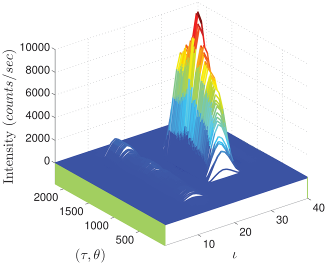
Goal:

Optimization for joint inversion from multiple data modalities.

Multimodal Image Analysis: Explosion of Data







Transmission

Fluorescence

Joint XRT & XRF Reconstruction (JRT)

More challenging nonlinear optimization:

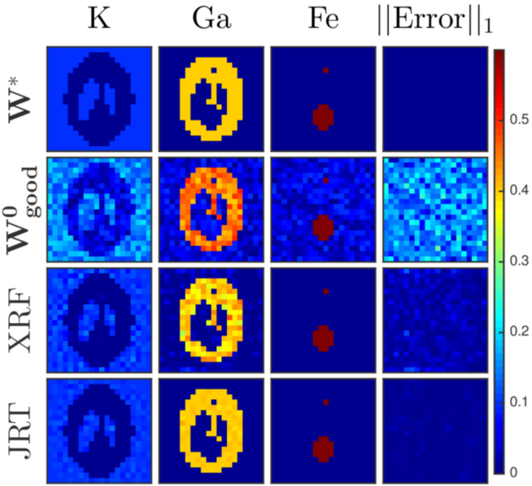
- Find weights $\mathbf{W}$: $\mathbf{F}^{\mathfrak{R}}_{\boldsymbol{\theta},\boldsymbol{\tau},\mathbf{E}}(\mathbf{W}) = \mathbf{D}^{\mathfrak{R}}_{\boldsymbol{\theta},\boldsymbol{\tau},\mathbf{E}}$ and $F^{\mathfrak{T}}_{\theta,\tau,E}(\mathbf{W}) = D^{\mathfrak{T}}_{\theta,\tau,E}$
- Single mode reconstruction is under-determined

$$\min_{\mathbf{W}} \sum_{\theta,\tau} \left(\frac{1}{2} \left\| \mathbf{F}^{\mathfrak{R}}_{\boldsymbol{\theta},\boldsymbol{\tau},\mathbf{E}}(\mathbf{W}) - \mathbf{D}^{\mathfrak{R}}_{\boldsymbol{\theta},\boldsymbol{\tau},\mathbf{E}} \right\|^2 + \frac{\beta}{2} \left\| F^{\mathfrak{T}}_{\theta,\tau,E}(\mathbf{W}) - D^{\mathfrak{T}}_{\theta,\tau,E} \right\|^2 \right)$$

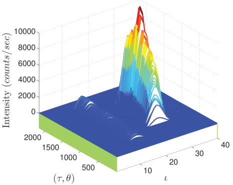
where

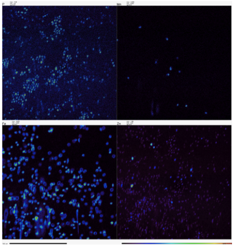
- $\mathbf{D}^{\mathfrak{R}}_{\boldsymbol{\theta},\boldsymbol{\tau},\mathbf{E}} \in \mathbb{R}^{n_E}$: Fluorescence data
- $D^{\mathfrak{T}}_{\theta,\tau,E} \in \mathbb{R}$: Transmission data
- $\mathbf{W} \geq 0$: constraint on voxel contents

JRT versus Single XRF Reconstruction



Optimization for Large Data and Image Analysis





Take-Home Message: Many interesting problems!

- Increased data rates drive optimization
- Huge scope for new optimization models and methods
  - Need to work with large data
    $\Rightarrow$ traditional techniques not applicable
  - Exploit domain knowledge whenever possible!

Outline

1. Introduction to Argonne National laboratory
2. Nonlinear Optimization for Power Grid Applications
   - Introduction to Power-Grid Applications
   - Optimal Unit Commitment for Power-Grid
3. Design of Complex Structures
   - Design of Nano-Structures
   - Topology and Structural Optimization
4. Analysis of Data and Images
   - X-ray Fluorescence Imaging
   - Multimodal Image Analysis
5. A Control Application: Optimal Transition to Clean Energy

Optimal Reduced-Carbon Technology Penetration

What is the optimal path to reduced-carbon economy?

Don Hanson, Steve Kryukov, Sven Leyffer, and Todd Munson

Argonne NATIONAL LABORATORY
... for a brighter future
U.S. Department of Energy
UChicago Argonne LLC

Modeling Optimal Penetration of Low-Carbon Technology

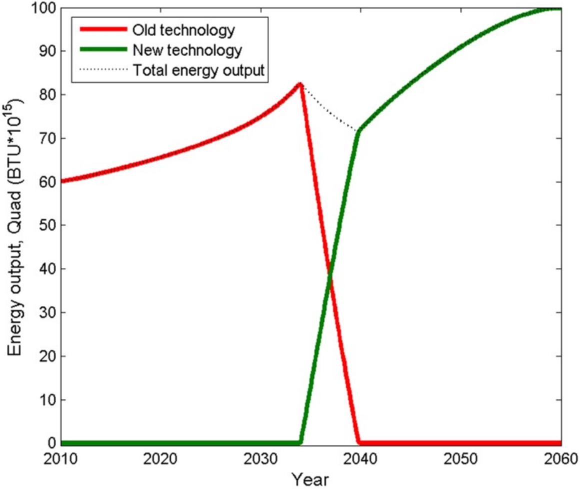
- Optimize energy production schedule and transition from old to new reduced-carbon technology to meet carbon targets
  - Maximize social welfare subject to GHG target in 50 years
  - Reduced-carbon technology subject to learning effects
    i.e. reduced unit cost as new technology becomes widespread
  - Includes adjustment costs (penalize rapid change in energy schedule)

- Social planner's problem … government in charge of everything

- Reasonable assumptions on GHG emissions, demand growth, energy costs, …

- Optimal control problem
  - Formulated using modeling language (AMPL)
  - Solved with nonlinear optimizers (KNITRO)
  - 50-year horizon solves within seconds on desktop

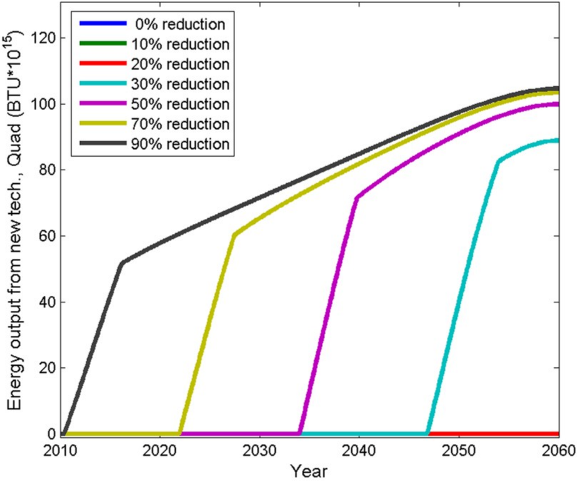
50% GHG Reduction: Optimal Energy Production Schedule



Transition period 2034-2040:
old conventional technology is gradually replaced by new low-emission technology

Effect of Different GHG Reduction Targets

- Plot shows low-emission energy production for different GHG targets.
- Transition to new technology is not necessary until 30% reduction target.



Technology Transition Model Description

- **Time** models dynamics over continuous time $t \in [0, T]$
  Functions of time $x(t)$ with derivatives

$$\dot{x}(t) = \frac{dx(t)}{dt}.$$

- **Energy Output** of old/new technologies at time t are $q^o(t),\ q^n(t)$.
  Define total output $Q(t) = q^o(t) + q^n(t)$.
- **Demand and Consumer Surplus** model benefit of energy to society:

$$\tilde{S}(Q,t) = e^{bt} S(Qe^{-bt}),$$

where $b > 0$ growth rate of demand.
Derived from constant elasticity of substitution (CES) utility.

$$S(Q) = \begin{cases} S(Q) = S_0 \ln Q, & \text{if } \sigma = 1 \\ \frac{S_0}{1-\sigma} Q^{1-\sigma}, & \text{otherwise,} \end{cases}$$

where $\sigma > 0$ is demand parameter.

Technology Transition Model Description

- **Production Costs.** Assume constant marginal costs.
  - Old technology unit cost: c_o
  - New technology unit cost subject to learning-by-doing: $c_n(x(t))$ decreasing function of cumulative output:

$$x(t) = \int_0^t q^n(\tau)d\tau.$$

  - Following [**?**], we let

$$c_n(x) = c_n^0 \left[\frac{x}{\bar{X}} + 1\right]^{\log_2 \gamma},$$

    where $\bar{X}$ and γ are parameters.
- **Greenhouse Gases Emissions**
  - Old technology emissions rate: $b_o > 0$
  - New technology emissions rate: $b_n \in (0, b_o)$
  - Cumulative (discounted) emissions at end period:

$$\int_0^T e^{-at}\big(b_o q^o(t) + b_n q^n(t)\big)dt \leq z_T$$

A Basic Model for Technology Transition

- Maximize social welfare subject to emissions caps.
- Control variables: Energy output $q^o(t), q^n(t)$
- State variables: Experience $x(t)$ and Emissions $z(t)$

$$
\begin{aligned}
\underset{\{q^o,q^n,x,z\}(t)}{\text{maximize}} \quad & \int_0^T e^{-rt}\left[\tilde{S}(q^o(t)+q^n(t),t) - c_o q^o(t) - c_n(x(t))q^n(t)\right] dt \\
\text{subject to} \quad & \dot{x}(t) = q^n(t), \qquad x(0) = x_0 = 0 \\
& \dot{z}(t) = e^{-at}\left(b_o q^o(t) + b_n q^n(t)\right), \qquad z(0) = z_0 = 0 \\
& z(T) \le z_T \\
& q^o(t) \ge 0, \qquad q^n(t) \ge 0.
\end{aligned}
$$

AMPL Model for Optimal Technology Penetration

Write an AMPL model of the basic technology transition problem

$$\begin{aligned}
\underset{\{q^o,q^n,x,z\}(t)}{\text{maximize}} \quad & \int_0^T e^{-rt}\left[\tilde{S}(q^o(t)+q^n(t),t) - c_o q^o(t) - c_n(x(t))q^n(t)\right] dt \\
\text{subject to} \quad & \dot{x}(t) = q^n(t), \qquad x(0) = x_0 = 0 \\
& \dot{z}(t) = e^{-at}\left(b_o q^o(t) + b_n q^n(t)\right), \qquad z(0) = z_0 = 0 \\
& z(T) \le z_T \\
& q^o(t) \ge 0, \qquad q^n(t) \ge 0.
\end{aligned}$$

where

$$\tilde{S}(Q,t) = e^{bt} S(Qe^{-bt}),$$

where $b > 0$ growth rate of demand, and

$$S(Q) = \begin{cases} S(Q) = S_0 \ln Q, & \text{if } \sigma = 1 \\ \frac{S_0}{1-\sigma} Q^{1-\sigma}, & \text{otherwise,} \end{cases}$$

AMPL Model for Optimal Technology Penetration

Parameter	Unit	Notation	Value
Discount rate	-	r	0.05
Demand exponent	-	σ	2.0
Demand scale	\$B	S_0	98,000
Demand growth rate	-	b	0.015
Environmental rate	-	a	0.02
Emissions, old tech.	tC/mBTU	b_o	0.02
Emissions, new tech.	tC/mBTU	b_n	0.001
Unconstrained emissions	BtC	Z_{max}	61.9358
Emission reduction %	-	ζ	0.5
Emissions target	BtC	$z_T = \zeta Z_{\mathrm{max}}$	
Production cost, old tech.	\$/mBTU	c_o	20
Starting cost, new tech.	\$/mBTU	c_n^0	50
Learning rate	-	γ	0.85
Initial experience	quad	x_0	0
Experience unit size	quad	$\bar{X}$	300

AMPL Model for Optimal Technology Penetration

Write an AMPL model of the basic technology transition problem

$$\begin{aligned}
\underset{\{q^o,q^n,x,z\}(t)}{\text{maximize}} \quad & \int_0^T e^{-rt}\left[\tilde{S}(q^o(t)+q^n(t),t) - c_o q^o(t) - c_n(x(t))q^n(t)\right] dt \\
\text{subject to} \quad & \dot{x}(t) = q^n(t), \qquad x(0) = x_0 = 0 \\
& \dot{z}(t) = e^{-at}\left(b_o q^o(t) + b_n q^n(t)\right), \qquad z(0) = z_0 = 0 \\
& z(T) \le z_T \\
& q^o(t) \ge 0, \qquad q^n(t) \ge 0.
\end{aligned}$$

1. Solve the model for different time steps, h
2. Experiment with different discretization schemes, e.g. explicit or implicit Euler and Trapezoidal rule
3. Use AMPL's `fprintf` to output your results & plot the transition paths in Matlab.

Optimal Technology Penetration: Hints & AMPL Tricks

Attributes & Defined Variables

```
param Ntime > 0, integer, default 200;  # ... time-steps

set T := 0..Ntime;

var qo{T} >= 0, := 10;        # ... old technology
var qn{T} >= 0, := 1;         # ... new technology

var z{T};                     # ... emissions

var dzdt{t in T: t<Ntime} = (z[t+1] - z[t])/h;

subject to

 # ... ODE for z(t)
 InitZ: z[0] = z0;                # ... initial cond.
 DiffEqnZ{t in T: t<Ntime}:      # ... disc. ODE
  dzdt[t] = exp(-a*t*Tend/Ntime) * (bo*qo[t] + bn*qn[t]);
```

Optimal Technology Penetration: Hints & AMPL Tricks

AMPL Defined Variables

```
var dzdt{t in T: t<Ntime} = (z[t+1] - z[t])/h;
```

- Acts like a constraints, but substituted out of model.
- Note "=" assigns constraints, and ":=" initial variable value
- All Variables/expressions on RHS must be defined previously

Check out effect of defined variables

- To show variables passed to solver, use
  `display _svarname, _sconname;`
- To show variables in model, use
  `display _varname, _conname;`

Conclusion: Optimization is Cool