Getting Out More from the Power Grid--
Mathematical Challenges of Energy Systems

Mihai Anitescu  SAMSI
2015

**M2ACS colleagues @ANL:** Emil Constantinescu, Victor Zavala, Cosmin Petra Shri Abhyankar, Francois Gilbert (ANL)

**M2ACS colleagues:** Henry Huang

**Other Contributors:** Miles Lubin (MIT), Dan Molzahn (Argonne)
1. The Problem
Complexity of the power grid

- $O(10^5)$ transmission lines,
- $O(10^4)$ generation nodes, and
- $O(10^4)$ demand nodes.

Distribution network: $O(10^8)$ components.

Reliability: – combinations of failures (2003 blackout)

time scale ranges: 1e-6 in power electronics to decades (1e 10s) in transmission planning.
Physical Constraints

**State Equations**
- $\sim 0.1$ second dynamics: swing equations; DAEs.
- Steady-State: Power Flow Equations, Nonlinear, Quadratic (ACPF)
- Their Linear Approximation: DC Power Flow.
- Sometimes, Power Flow Only.

\[ M\dot{x} = f(x, y, p) \]
\[ 0 = g(x, y, p) \]

**Decisions**
- Generation Levels
- Generation Schedules
- Asset Location
- System states (angles, inertia; for inverse problems)
Decision Making Structure and Optimization Tasks

Planning

Transmission/Generation Expansion: ISO, Yearly, MILP

Unit Commitment: ISO, Daily, DC Flow, MILP
Day-Ahead Bidding: GENCOs/Utilities, Daily, LP/QP
Economic Dispatch: ISO, 5 Minutes, DC Flow, LP/QP
Real-Time Bidding: GENCOs/Utilities, 5 Minutes, LP/QP

Markets

AC Power Flow: ISO, 1-2 Minutes, NLP
State Estimation: ISO, 1-2 Minutes, QP/NLP
Generation Control: GENCOs, Seconds, QP/NLP
Voltage and Dynamic Stability: ISO, MilliSeconds, No Optimization
Energy Management: Utilities/Consumers, Seconds/Minutes, LP/QP/NLP

Control
Transmission Planning -- Upper Level

- In principle encompasses all the complexity of the grid analysis.
- Upper level: the 20 year financial problem with hourly increments (100K+ time steps) Lower level reliability constraints

\[
\min_{\{x_i\}, \{y_i\}, \{u_i\}} \mathbb{E}_\omega \left[ \sum_{i=T_i}^{T_f} h(x_i, y_i, u_i, d_i, \omega_i) \right]
\]

\[
x_{i+1} = f_d(x_i, y_i, u_i, d_i, \omega_i)
\]

\[
0 = f_a(x_i, y_i, u_i, d_i, \omega_i)
\]

\[
g(x_i, y_i, u_i, d_i, \omega_i) \leq 0
\]

\[
R(x_i, y_i, u_i, d_i, \omega_i) = TRUE
\]
Lower Level: Reliability Constraints

- These are the famous N-1 or N-2 constraints, denoted by the contingency set, $C$.

\[
R(x_i, y_i, u_i, d_i, \omega_i) = \text{TRUE} \iff \\
\forall c \in C:\begin{cases}
\int_0^T \psi_j(x^c(t), y^c(t)) \leq a_j, \forall j \in J_c \\
\psi_i(x^c(t), y^c(t)) \leq a_i, \forall i \in I_c
\end{cases},
\]

\[
\begin{aligned}
\frac{dx^c}{dt} &= f_d(x^c, y^c, u_i, d_i, \omega_i) \\
0 &= f_a(x^c, y^c, u_i, d_i, \omega_i) \\
x^c(0) &= x_i
\end{aligned}
\]

- Direct transcription of the problem results in $10^{26}$ variables,

- And schedules may be variable (unit commitment) – integer variables.

- Current practice? Long horizon problem is run without networks, and a few constraints sampled and run under some scenarios, using engineering judgment.

- **WG**: However, there is strong evidence that temporal decomposition would be very effective even with integer variables.
Current Features and State of the Analysis

- Problem: grid operates cheaply and safely, including following acts-of-God
- Same physical model, but historical separation between analyses.
- Direct transcription is beyond any compute power in 2050.
- Engineering judgment is used offline for reduction. Eg.
  - Scenarios to be used in transmission expansion planning.
  - Static constraints for post-contingencies states.
  - Screening of the possible contingencies in security-constrained problems.
- Recourses are generally implemented with operator-in-the-loop.
- While perhaps a bit dated, it served us well. US has one of the lowest electricity prices and higher reliabilities.
However, is it really a solved problem?

- The 2003 blackout affected 50 million people
  - Task Force Report: “we could not complete the simulation in the 6 months …”
- The 2013 Superbowl blackout (CNN)
  - “Manufacturer blames Super Bowl outage on incorrect setting”
  - “System operators essentially put the relay's trip setting too low,”
- The 2014 February Gas Shortage/Curtailment (Forbes) ... In TEXAS (and California, Mass)
- The 2015 White House Blackout ... (CBSDC)
  - “Explosion at Power Plant Responsible for D.C. Area Outages “ -- grid operations should ensure that no one event to an asset results in blackout.”
Why do these things not happen more often?

Hierarchy of Reserves

- Margins and off-market actions probably cost Bs of Dollars (5-20)
- And, the same reserve level may not hold for the NEW grid
Getting out more of the power grid: converging optimization and transient stability (RES, Abhyankar)

- You say: “give me the best energy portfolio--”dispatch”-- but make sure that I survive transient from any one contingency “.
- Computer sees:

$$\begin{array}{l}
\min \quad C(p) \\
\text{s.t.} \\
g_s(p) = 0 \\
h_s(p) \leq h^+ \\
p^- < p < p^+ \\
M\dot{x} = f(t, x, y, p), \quad x(t_0) = I_{x0}(p) \\
0 = g(t, x, y, p), \quad y(t_0) = I_{y0}(p) \\
h(x(t), y(t)) = 0, \quad \forall(t) \quad \text{path constraints}
\end{array}$$

If the dynamic tool cannot compute derivatives, this becomes dim(p) time harder.

S. Abhyankar, E. Constantinescu, and M.A.
A bundle method for optimization with transient constraints (RES, Gilbert)

- Integration of dynamics and optimization, given that direct transcription of dynamics is not possible.
- We use adjoint calculations/PETSc to compute the derivative of the path constraint violation with respect to the dispatch.
- We use a bundle method to compute it.
- (WG): Are there good “asynchronous” versions of it?
- The balance between computational effort at a new step versus local improvements
Some Results:

- Before and after:

- (WG) in what circumstances is the dynamic stability set convex? (for TSUC)
This is with the grid today, how about tomorrow?

- New usage drivers (PHEV, bidirectional, smart grid, smart buildings) are changing completely the response of the system.
Changing Demand Patterns

- Assuming 50% solar penetration.
- More dynamic content.
- The error would have to increase – needing larger reserves.

Solar Radiation winter 2008 in the COMED AREA

F. Gilbert and M.A.
Maybe things will turn out all right anyways? Doubtful -- Germany Energy Landscape
Increase Natural Gas Usage ➔ Very different issues

- Gas Pipelines Provide Storage Capacity to Mitigate Fluctuations (Ramps & Capacity)
- But, Gas Travels at 30-50 mph (Storage Has to be Built up Well in Advance)
- Aggressive Gas Withdrawals Cascade Upstream the Pipeline (Compromises Stability)
- Shortages of Natural Gas for Power Plants Not Uncommon
Complexity Drivers over next 15 years

- More pro-active consumer (“prosumer”)
- Significant increase of uncertainty in supply, demand, and inertia due to renewables expansion and prosumers.
- Increased Reliance on Distributed Generation and Natural Gas.
- Changes in Power Flow patterns
- Increased Dynamic Ranges, shorter analyses time scales and real times.
- Rapidly Increasing Data Streams.
- New Devices. (PHEV, PMUs, HVDC, mass and local storage)

- Larger state spaces, smaller time scales, more physics, more scenarios,
Total US R&D Investment

The Global Innovation 1000: Comparison of R&D Spending by Regions and Industries

This graph allows you to compare R&D as a percentage of revenue (R&D "intensity") and total R&D spend by regions and industries as it changes from 2004-2011.
2. Unusual/unexpected mathematics
Classical OPF Problem

\[
\begin{align*}
\text{min} \quad & \sum_{k \in G} f_k(P_{Gk}) = \sum_{k \in G} c_{2k} P_{Gk}^2 + c_{1k} P_{Gk} + c_{0k} \\
\text{subject to} \quad & P_{Gk}^{\min} \leq P_{Gk} \leq P_{Gk}^{\max} \\
& Q_{Gk}^{\min} \leq Q_{Gk} \leq Q_{Gk}^{\max} \\
& \left(V_k^{\min}\right)^2 \leq V_{dk}^2 + V_{qk}^2 \leq \left(V_k^{\max}\right)^2 \\
& |S_{lm}| \leq S_{lm}^{\max}
\end{align*}
\]

\[
\begin{align*}
P_{Gk} - P_{Dk} &= V_{dk} \sum_{i=1}^n (G_{ik} V_{di} - B_{ik} V_{qi}) + V_{qk} \sum_{i=1}^n (B_{ik} V_{di} + G_{ik} V_{qi}) \\
Q_{Gk} - Q_{Dk} &= V_{dk} \sum_{i=1}^n (-B_{ik} V_{di} - G_{ik} V_{qi}) + V_{qk} \sum_{i=1}^n (G_{ik} V_{di} - B_{ik} V_{qi})
\end{align*}
\]

Rectangular voltage coordinates: \( \bar{V}_i = V_{di} + jV_{qi} \)
SDP Relaxations

- SDP relaxations were shown to be exact in many cases (Javaei and Low).
- Higher order moment relaxations of Laserre have been shown to significantly fill the gap (Molzahn, Lesieutre, Hiskens).
- However, practical estimates of the order of relaxations are lacking.
- We do not have sufficiently fast SDP implementations.

Source: D. Molzahn (ANL)
Random Graphs (RG)

- Essential if aim towards expert free planning (10s yrs horizons)
- Current RG models do not reproduce power networks well

<table>
<thead>
<tr>
<th>Graph type</th>
<th>Vertices, edges</th>
<th>Clust coeff</th>
<th>ASPL</th>
<th>Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western US</td>
<td>4941, 6594</td>
<td>0.0166</td>
<td>21.75</td>
<td></td>
</tr>
<tr>
<td>Random (of WECC)</td>
<td>4941, 6594</td>
<td><strong>0.00049</strong></td>
<td><strong>10</strong></td>
<td></td>
</tr>
<tr>
<td>Nordic</td>
<td>4789, 5571</td>
<td>0.0801</td>
<td>18.99</td>
<td></td>
</tr>
<tr>
<td>Random (of nordic)</td>
<td>4789, 5571</td>
<td><strong>0.00054</strong></td>
<td><strong>8.7</strong></td>
<td></td>
</tr>
<tr>
<td>Eastern inter</td>
<td>49597, 62985</td>
<td>0.071</td>
<td>35.8</td>
<td>96</td>
</tr>
<tr>
<td>Pref. attach (eastern)</td>
<td>49597, 62966</td>
<td><strong>0.0006</strong></td>
<td><strong>7.2</strong></td>
<td>18</td>
</tr>
<tr>
<td>Small world p=0.0882 (eastern)</td>
<td>49597, 62906</td>
<td><strong>0.27</strong></td>
<td>36.6</td>
<td>96</td>
</tr>
</tbody>
</table>

M Halappanavar, PNNL. (ASPL== average shortest path length)
Statistical Models of Sub-Second Noise

- Evidence of non-Gaussian, heavy-tailed, complex spectrum.
- Too Complex for AR models to cut it.
3. Stochasticity/Probabilistic Modeling
Sources of Uncertainty

- Reformulations: there appears to be a big push for probabilistic modeling and decision making in light of increased renewable penetration.
- See the FERC Technical Conferences.
- However, such techniques do not appear to be used in a major way operationally (at all?)
- Upfront cost increases, but so does reliability.
- To model:
  - Weather
  - Demand
  - Sub-second noise.
  - Fuel Cost
  - Reliability
  - ....
Stochastic unit commitment/economic dispatch w. wind

\[
\begin{align*}
\text{min} \quad \text{COST} = & \frac{1}{N_s} \sum_{s \in S} \left( \sum_{j \in N} \sum_{k \in T} c_{sjk}^p + c_{sjk}^u + c_{sjk}^d \right) \\
\text{s.t.} \quad \sum_{j \in N} p_{sjk} + \sum_{j \in N_{\text{wind}}} p_{sjk}^{\text{wind}} = D_k, s \in S, k \in T \\
\sum_{j \in N} \tilde{p}_{sjk} + \sum_{j \in N_{\text{wind}}} p_{sjk}^{\text{wind}} \geq D_k + R_k, s \in S, k \in T \\
\text{ramping constr., min. up/down constr.}
\end{align*}
\]

- Most data approaches lose coherence after 2-3 hrs.
- Numerical weather prediction (NWP) expensive
- NWP was never meant to simulate < 1 km resolution.

Wind Forecast – WRF(Weather Research and Forecasting) Model
- Real-time grid-nested 24h simulation
- 30 samples require 1h on 500 CPUs (Jazz@Argonne)

Images courtesy of V. Zavala & E. Constantinescu @ANL
Hybrid NWP/Data Models

- joint distribution of weather simulations (NWP) and observations (Obs)

\[
\begin{pmatrix}
Y_{\text{Obs}} \\
Y_{\text{NWP}}
\end{pmatrix} \sim \mathcal{N}
\begin{pmatrix}
\mu_{\text{Obs}} \\
\mu_{\text{NWP}}
\end{pmatrix},
\begin{pmatrix}
\Sigma_{\text{Obs}} & \Sigma_{\text{Obs,NWP}} \\
\Sigma_{\text{Obs,NWP}}^T & \Sigma_{\text{NWP}}
\end{pmatrix}
\]

- wind speed prediction fusing NWP and Data

Julie Bessac, Emil Constantinescu, and Mihai Anitescu; *Stochastic simulation of predictive space-time scenarios of wind speed using observations and physical models*, 2015.
Optimal Bandwidth Selection for Multitapers (RES), w C. Haley

- Multitapers are a fascinating spectral estimation technique, and the only one that can get noise estimates.
- However, it has a bandwidth parameter $W$ that needs to be selected, the standard is $NW=4$.
- What is the optimal bandwidth selection?
- We show that the mean square error is (asymptotically)
  \[ \hat{\Lambda}(N,W) = \sum_{m=0}^{M-1} \left[ \frac{W^3 (\tilde{S}^H(f_m))^2}{3} \right] \hat{\sigma}^2(f_m); \]
- Then we minimize over $W$ to obtain the optimal bandwidth.
- Results for Lorentzian and real PMU data

Charlotte Haley and Mihai Anitescu; Optimal Bandwidth for Multitaper Spectrum Estimation (2016)
Optimal versus Suboptimal Bandwidth Selection, PMU data

- NW=4 has much more variability and can result in false positives for “peaks”
- Optimal choice NW=16 has a much smoother appearance.
- Cis are a lot better (and valid)
Gas Networks

Transport Equations for link $\ell \in \mathcal{L} := \mathcal{L}_p \cup \mathcal{L}_a$

\[
\frac{\partial p_\ell}{\partial t} + \frac{1}{A_\ell \rho_\ell} \frac{\partial f_\ell}{\partial x} = 0
\]
\[
- \frac{1}{A_\ell} \frac{\partial f_\ell}{\partial t} + \frac{\partial p_\ell}{\partial x} + \frac{8 \lambda_\ell}{\pi^2 D_\ell^5} \frac{f_\ell | f_\ell |}{\rho_\ell} = 0
\]
\[
f_\ell |_{x=0} = f_\ell^{in}
\]
\[
f_\ell |_{x=L_\ell} = f_\ell^{out}
\]
\[
p_\ell |_{x=L_\ell} = \theta_{rec(\ell)}
\]
\[
p_\ell |_{x=0} = \theta_{snd(\ell)}, \ \ell \in \mathcal{L}_p
\]
\[
p_\ell |_{x=0} = \theta_{snd(\ell)} + \Delta \theta_\ell, \ \ell \in \mathcal{L}_a
\]

Conservation at node $n \in \mathcal{N}$

\[
\sum_{\ell: rec(\ell)=n} f_\ell^{out} + \sum_{i: sup(i)=n} s_i - \sum_{\ell: snd(\ell)=n} f_\ell^{in} - \sum_{j: dem(j)=n} d_j = 0
\]

Compression Power for link $\ell \in \mathcal{L}_A$

\[
P_\ell = f_\ell^{in} c_p T \left( \left( \frac{\theta_{snd(\ell)} + \Delta \theta_\ell}{\theta_{snd(\ell)}} \right)^{\frac{\gamma - 1}{\gamma}} - 1 \right)
\]
Stochastic Versus Deterministic Gas Pipeline Operations for Electricity

Nominal Demand | High Demand
---|---

Nominal Demand | High Demand
---|---

V. Zavala
Converging Infrastructure leads to increased weather sensitivity

- (Forbes article on Feb 2014) events.
- **California**’s electric grid operator asked generators to reduce their gas usage.
  - In **Texas**, which produces more natural gas than any other state, the Electric Reliability Council of Texas (ERCOT) called a gas-scarcity state of emergency.
  - In **New England** the price of natural gas to soar **20-fold** to more than $100 per thousand cubic feet. Coal, Jet Fuel, Oil plants were emergently started.
- This may have been foreseeable with probabilistic analysis tools.
Picture and Challenges for Stochastic Modeling

- Space-time, non-stationary, models of uncertainty.
- High quality PDFs of weather.
- New physics: PDEs. DAEs.
- (addition cascades, sub-second noise)

- The benefits are increased reliability and more transparent pricing, but not yet operationally widespread.
- Challenge: Elicit Risk Functionals out of Decision Makers.
4 Computing/Supercomputing
State of Computing

- Computing is central on all these analyses.
- Parallel computing is starting to make inroads, but the use of supercomputing is rare.
- However, it is quite probable it will cost less than using humans for scenario screening and be much more accurate.
- E.g N-1 scenario dispatch over an RTO area, with stochasticity in demand and supply.
  - 10000 x 100 scenarios x 10000 degrees of freedom.
- It is certainly worth asking to what extent our approaches for certain problems scale to that level.
Stochastic programming - computational patterns

- Computational pattern

1\textsuperscript{st} stage computations

Optimization

- Deterministic model
- Effects of uncertainty

2\textsuperscript{nd} stage computations

Optimization
- Recourse problem outcome 1
Optimization
- Recourse problem outcome 2
Optimization
- Recourse problem outcome N

- Security-Constrained Problems.
- Two Stage Stochastic Programming
- Model Calibration.
Scalable optimization solver - PIPS-IPM

- **Schur-complement scenario decomposition**

\[
\begin{bmatrix}
K_1 & B_1 \\
\vdots & \vdots \\
B^T_1 & K_N & B_N \\
\end{bmatrix}
\begin{bmatrix}
\Delta z_1 \\
\vdots \\
\Delta z_N & \Delta z_0 \\
\end{bmatrix} =
\begin{bmatrix}
r_1 \\
\vdots \\
r_N & r_0 \\
\end{bmatrix}
\]

\[
\left(K_0 - \sum_{i=1}^{N} B^T_i K_i^{-1} B_i \right) \Delta z_0 = r_0 - \sum_{i=1}^{N} B^T_i K_i^{-1} r_i
\]

- **Incomplete augmented factorization w. perturbation**
  - Multithreading and full sparsity exploiting
  - BiCGStab for maintaining numerical stability
  - >10 x speed-up

- **Technical idea: halt factorization before 2x2 block**

\[
\begin{bmatrix}
K_i & B^T_i \\
B_i & 0
\end{bmatrix} =
\begin{bmatrix}
L_{11} & 0 \\
L_{21} & L_{22}
\end{bmatrix}
\begin{bmatrix}
U_{11} & U_{12} \\
0 & U_{22}
\end{bmatrix}
\]

\[
L_{22}U_{22} = -B^T_i K_i^{-1} B_i
\]

---

Full HPC evaluation of PIPS-IPM on power grid models

Relaxation of 24-hour unit commitment

- 4 billion decisions variables
- 4.1 billion constraints
  (32,768 scenarios)
- Solution time: < 40 mins
- FLOPS peak performance 12%
- Parallel efficiency > 90%

Some Challenges

- The usual suspects: nonconvexity and integrality.
- Scenario Clustering. (Contingency x Load x Demand x Weather)
- Convergence of Planning and Operations ➔ Convergence of optimization and Dynamics (including MINLP + Dynamics)
- Very Long Time Horizons in Planning.
- Very large scale nonlinear programming, -- using decomposition.
- Rolling horizon: Warm/hot starting.
5. Decision-Oriented Algebraic Modeling for Complex Systems
State of Modeling in Optimization

- Optimization shines in the simplicity and power of its abstraction.
- Compare with partial differential equations.
- Optimization is one of the few areas where algebraic modeling is common (e.g. AMPL, GAMS, PyOMO etc).
- On the other hand, for power grids we need to incorporate new patterns fast, and support parallelism, dynamics, PDEs.
- We may also be a very small market for AML providers.
- For huge problems, the pricing model per core does not suit us.
DIY Algebraic Modeling using Julia (Lubin et al)

- Julia: a high-level, high-performance, open-source dynamic language for technical computing
- Keeps productivity of dynamic languages without giving up speed (2x of C/C++/Fortran)
- Macros instead of operator overloading (known to have poor performance).
- Ideal for creating and extending AMLs without giving up speed JuMP (Lubin et al.)

![Model example]

```
m = Model(:Max)
@defVar(m, 0 <= x[j=1:N] <= 1)
@setObjective(m, sum{profit[j] * x[j], j=1:N})
@addConstraint(m, sum{weight[j] * x[j], j = 1:N} <= C)
```

---

**Table: Linear-quadratic control benchmark results. N=M is the grid size. Total time (in seconds) to process the model definition and produce the output file in LP and MPS formats (as available).**

<table>
<thead>
<tr>
<th>N</th>
<th>JuMP/Julia</th>
<th>AMPL</th>
<th>Gurobi/C++</th>
<th>Pulp/PyPy</th>
<th>Pyomo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LP MPS</td>
<td>MPS</td>
<td>LP MPS</td>
<td>LP MPS</td>
<td>LP</td>
</tr>
<tr>
<td>250</td>
<td>0.5 0.9</td>
<td>0.8</td>
<td>1.2 1.1</td>
<td>8.3 7.2</td>
<td>13.3</td>
</tr>
<tr>
<td>500</td>
<td>2.0 3.6</td>
<td>3.0</td>
<td>4.5 4.4</td>
<td>27.6 24.4</td>
<td>53.4</td>
</tr>
<tr>
<td>750</td>
<td>5.0 8.4</td>
<td>6.7</td>
<td>10.2 10.1</td>
<td>61.0 54.5</td>
<td>121.0</td>
</tr>
<tr>
<td>1000</td>
<td>9.2 15.5</td>
<td>11.6</td>
<td>17.6 17.3</td>
<td>108.2 97.5</td>
<td>214.7</td>
</tr>
</tbody>
</table>

- Allows for different levels of user expertise within same language.
- We started using it last year, 3 of us spend > 50%, one 100%.
Example DIY: Parallel SP extension StochJuMP

- Extension of Julia JuMP for stochastic LP/QP/MILP/
- Interfaced with PIPS, runs efficiently on “Blues” LCRC cluster (MPI)
- AMPL took 3 days to instantiate and 1 hour to run on BG.

Some Challenges in Efficient Modeling

- If we are to break the barriers between analyses, a good algebraic modeling environment is necessary. Some challenges:
  - PDEs -- Natural Gas is Hyperbolic.
  - Dynamical/differential algebraic systems.
  - Determining well posedness of composite models (easier in dynamics than in optimization – see SIMULINK or Modelica)
  - Some abstraction of hierarchical modeling
    - interconnect-distribution-building
    - yearly-daily-hourly-second
Summary

- We manage the grid very well today, but it is not a solved problem.
- The US grid will change dramatically over the next 15 years.
- This will challenge our analytical abilities at multiple levels.
- Mathematical Foundations.
- Probabilistic Modeling
- New Algorithms
- Problem Representations
Unsolved and unapproached math conceptual challenges

- Global Optimization with nonconvex constraints.
- Equilibrium formulations with integer variables
- Probability density for cascade events.
- Random graphs
- Multiscale Math for Networks

- Spatio-temporal, nonstationary, data modeling.
- Rare event simulations
- Stochastic numerical simulation
- Multistage stochastic programming
- Error models for engineered physical processes
- Non-Gaussian noise
- Hybrid, multisource data analysis.