Gradient-Based Optimization of Rolling-Horizon Power System Models

Lilianna M. Gittoes
School of ORIE
Cornell University

Shane G. Henderson
School of ORIE
Cornell University
sgh9@cornell.edu

Jacob P. Mays
School of CEE
Cornell University

SIAM Conference on Optimization, 2026

Research supported, in part, by NSF grants OAC 2410950 and DMS 2230023

A Great Mentor and a Good Mate

Motivation, Goal and TL;DR

- What is the value of investments in renewable generation, battery storage at the grid level?
- Complex intertemporal dynamics: Upfront investment decisions, production costs unfold over years.
- Two-stage frameworks help, but struggle in capturing the sequential nature of uncertainty.
- We consider model predictive control (MPC) for production cost modeling. We parameterize MPC and optimize over the choice of parameters (terminal value of stored energy). (Centralized decision maker.)
- TL;DR: Focus is on automatic differentiation to get gradients and proof that this yields unbiased gradient estimators. Then we do gradient-based search over parameters.

Contributions

- Approach to getting gradient estimators when optimization is embedded within (stochastic) simulation replications.
- Conditions under which this provably yields unbiased estimators when using MPC.
- Framework giving guidance on when this will work more generally.
- A new approach to capacity expansion and resource adequacy studies.
- Relative to Ferris and Philpott (2023) more natural handling of "second-stage" dynamics.
- Relative to Khazaei and Powell (2018) don't need finite differences.
- Not the first to differentiate optimal solutions - long history; see especially "Layers" – Amos and Kolter (2017), Agrawal et al. (2019).

General System Structure

Stage 1	Stage 2			
	1	2	$\cdots$	T
	s_1 ξ_1 $\boxed{x_1(\theta, s_1)}$			
Choose θ		s_2 ξ_2 $\boxed{x_2(\theta, s_2)}$		
			$\ddots$	
				s_T ξ_T $\boxed{x_T(\theta, s_T)}$
$c_0(\theta)$	$C(x_1(\theta, s_1))$	$C(x_2(\theta, s_2))$	$\cdots$	$C(x_T(\theta, s_T))$

Objective Function and Derivative Estimation

$$\min_{\theta} f(\theta) = c_0(\theta) + \sum_{t=1}^{T} \mathbb{E} C(x_t(\theta, s_t, \xi_{1:t}))$$

$$\begin{aligned} \nabla f(\theta) &= \nabla c_0(\theta) + \sum_{t=1}^{T} \nabla \mathbb{E} C(x_t(\theta, s_t, \xi_{1:t})) \\ &= \nabla c_0(\theta) + \sum_{t=1}^{T} \mathbb{E} \nabla C(x_t(\theta, s_t, \xi_{1:t})) \text{ if interchange ok...} \end{aligned}$$

Automatic differentiation/infinitesimal perturbation analysis suggests using

$$\widehat{\nabla} f(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left[\nabla c_0(\theta) + \sum_{t=1}^{T} \nabla C(x_t(\theta, s_t(i), \xi_{1:t}(i))) \right]$$

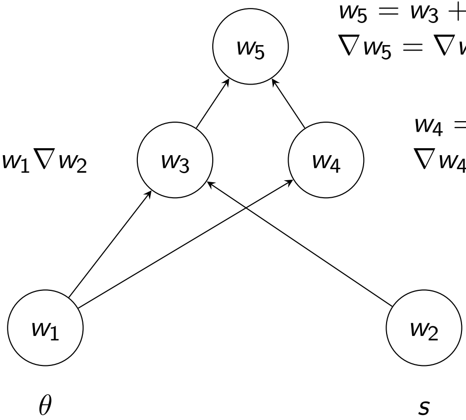
Computational Graph Warmup

$$
\begin{aligned}
z = g(\theta, s) &= \theta s + \sin\theta \\
w_5 = g(w_1, w_2) &= w_1 w_2 + \sin w_1 \\
&= w_3 + w_4
\end{aligned}
$$

$$
\begin{aligned}
w_5 &= w_3 + w_4 \\
\nabla w_5 &= \nabla w_3 + \nabla w_4
\end{aligned}
$$

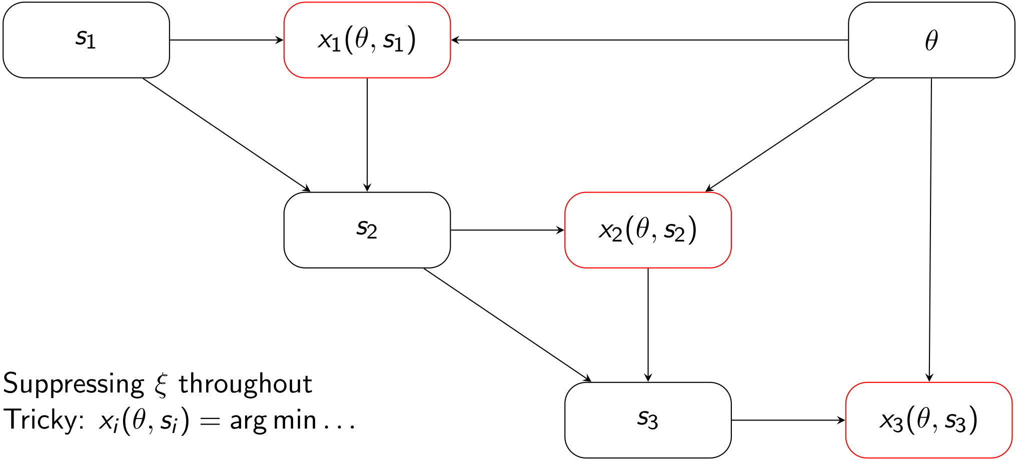
$$
\begin{aligned}
w_3 &= w_1 w_2 \\
\nabla w_3 &= w_2 \nabla w_1 + w_1 \nabla w_2
\end{aligned}
$$

$$
\begin{aligned}
w_4 &= \sin(w_1) \\
\nabla w_4 &= \cos w_1 \cdot \nabla w_1
\end{aligned}
$$



$$\nabla w_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \nabla w_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Computational Graph



Suppressing ξ throughout
Tricky: $x_i(\theta, s_i) = \arg\min \ldots$

What is "∇ arg min..."?

Basically need strictly convex minimization problem at each step. To be more precise (and to guarantee unbiasedness later) we assume

MPC to choose x_t is a QP

$$\begin{aligned} \min_x \quad & \frac{1}{2} x^\top Q x + p(\theta, s_t, \xi_{1:t})^\top x \\ \text{s.t.} \quad & a_j^\top x \geq b_j(\theta, s_t, \xi_{1:t}), \; j \in \mathcal{I} \\ & c_j^\top x = d_j(\theta, s_t, \xi_{1:t}), \; j \in \mathcal{E} \end{aligned}$$

where

- Q symmetric, positive definite (think regularization),
- Functions p, b_j, d_j affine in $\theta, s_t, \xi_{1:t}$,
- Problem is feasible for all $(\theta, s_t, \xi_{1:t})$ in certain polyhedral domains.

Getting Derivatives

Derivatives come from duality/implicit function theorem.
KKT system and its derivative:

$$\begin{pmatrix} Q & -A_{\mathcal{A}}^\top & -C^\top \\ A_{\mathcal{A}} & 0 & 0 \\ C & 0 & 0 \end{pmatrix} \begin{pmatrix} x^* \\ \lambda^* \\ \nu^* \end{pmatrix} = \begin{pmatrix} -p \\ b_{\mathcal{A}} \\ d \end{pmatrix}$$

$$\begin{pmatrix} Q & -A_{\mathcal{A}}^\top & -C^\top \\ A_{\mathcal{A}} & 0 & 0 \\ C & 0 & 0 \end{pmatrix} \begin{pmatrix} \nabla x^* \\ \nabla \lambda^* \\ \nabla \nu^* \end{pmatrix} = \begin{pmatrix} -\nabla p \\ \nabla b_{\mathcal{A}} \\ \nabla d \end{pmatrix}$$

Linear independence constraint qualification ensures nonsingular system.

Recap of Computing Gradients

For each of N replications:

- Layout computational tree as seen earlier.
- Use chain rule to percolate gradients through.
- Gradients of MPC at each stage computed by solving a linear system obtained from KKT conditions.
- Don't need to assume affine functions, but doing so simplifies things. Need strict convexity of MPC, continuity in θ, s_i

But are the gradient estimators so obtained unbiased?

Proving Unbiasedness I: Piecewise Affine x_t

Assume

- x_t obtained by solving previous quadratic program $\forall t$
- The matrix Q is fixed, symmetric positive definite.
- The functions p, b_j, d_j are affine in (θ, s, ξ).
- The space of θs is polyhedral, as is the space containing the states and random variables.
- The quadratic program is always feasible.

Can now partition domain into *critical regions* with constant active sets of constraints. Adapting Bemporad et al. (2002), Tøndel et al. (2003a, 2003b) we get

Proposition

x_t is piecewise affine, continuous in θ, s_t, ξ_t over a finite polyhedral partition.

Proving Unbiasedness II: Sample-Path Obj Function

$$f(\theta, s_1, \xi) = c_0(\theta) + \sum_{t=1}^{T} C(x_t(\theta, s_t, \xi_{1:t}))$$

$$s_{t+1} = g(s_t, x_t)$$

Assume, in addition, state transition function g is affine.

Proposition

Sample path function $f(\cdot, \cdot, \cdot)$ is piecewise convex quadratic, continuous.

Proving Unbiasedness III

Assume, also, that ξ has a continuous distribution.

Proposition

For almost all s_1, and for almost all θ in the interior of its feasible region, the sample path function $f(\cdot, s_1, \xi)$ is differentiable at θ with probability 1.

Assume, also, that $\mathbb{E}\|\xi\|^2 < \infty$.

Proposition

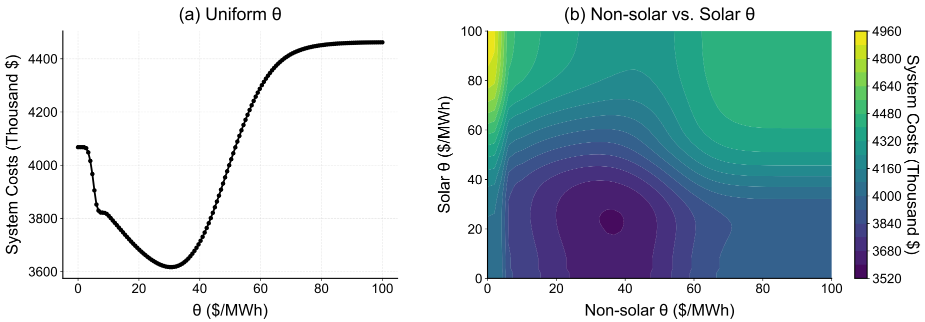
For almost all s_1, and for almost all θ in the interior of its feasible region, the cost function $\mathbb{E}f(\cdot, s_1, \xi)$ is differentiable at θ and

$$\nabla \mathbb{E} f(\cdot, s_1, \xi) = \mathbb{E} \nabla f(\cdot, s_1, \xi).$$

Production Cost Model: Proof of Concept

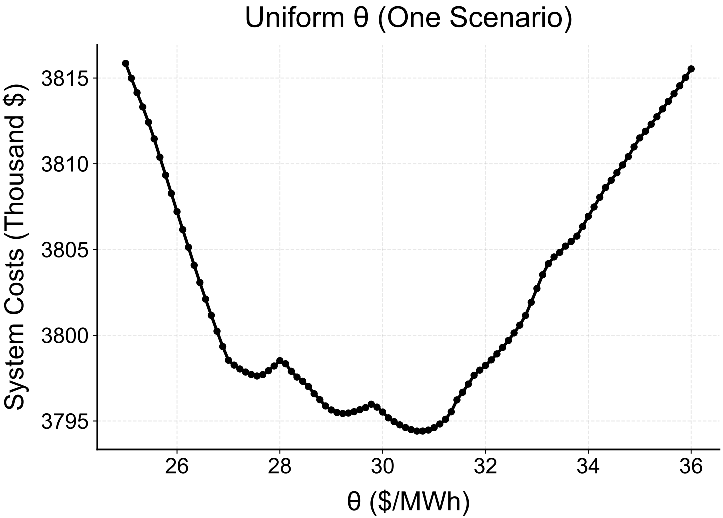
- One thermal generator, one solar generator, one battery.
- Strictly convex quadratic costs for each resource.
- Decision variables: for each hour, thermal generation, renewable generation, battery charging, battery discharging, shed load.
- State: battery charge $s_{t+1} = s_t + (1 - \ell)$charge - discharge.
- ξ_t is just demand. (Assuming deterministic solar.)
- System operator uses H-hour deterministic lookahead (MPC) to select x_t.
- Encourage charge/discharge using $\theta_{t+H} s_{t+H}$

In-sample Cost Surfaces



- Plot is cost surface in θ for a single θ for all times (left), and one θ for day, one θ for night (right)
- Used hourly data from NYISO Central zone from May-Sept 2025. Took $N = 160$ sample paths, used a certain AR(1) process for demand, centered on hourly averages.

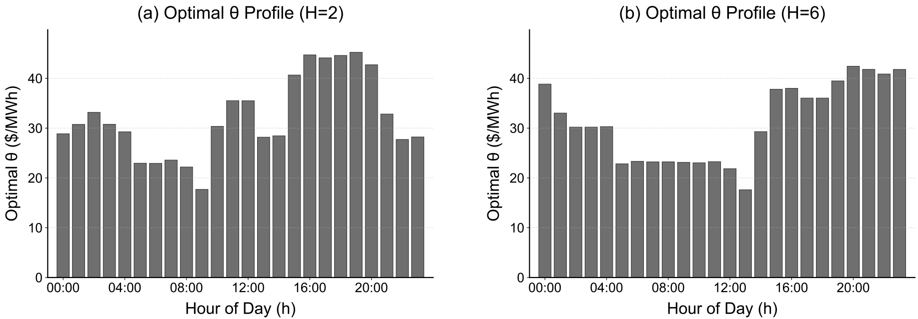
Zooming In


Uniform θ (One Scenario)
System Costs (Thousand $)
3815
3810
3805
3800
3795
26
28
30
32
34
36
θ ($/MWh)

Out-of-Sample Cost Comparison

Policy	$H = 2$	$H = 6$	$H = 10$	$H = 24$
No terminal value ($\theta = 0$)	\$4,769	\$4,100	\$3,722	\$3,604
Single terminal value	\$3,704	\$3,659	\$3,613	\$3,604
Optimal θ^*	*32.26*	*30.68*	*31.30*	–
Two-regime terminal value	\$3,689	\$3,636	\$3,605	\$3,604
Optimal θ^* *(Non-solar, Solar)*	*(34.94, 27.12)*	*(35.96, 24.11)*	*(36.93, 25.98)*	–

Optimized θ Profiles



Now we allow θ to depend on the hour of the day, so there are 24 values.

Conclusion

- IPA (automatic differentiation) can work with MPC.
- Gradients come from KKT systems - care needed, especially at boundaries of critical regions.
- IPA gradients are unbiased and obtained efficiently (low overhead compared to finite differences).
- Parameterized MPC with optimized terminal energy values mitigates horizon boundary effects.
- On stylized high-storage systems, short horizons with tuned terminal values approximate long-horizon performance.
- Framework is extensible to capacity expansion and other MPC contexts.