Control of Uncertain Power Systems via Convex Optimization

Eilyan Bitar

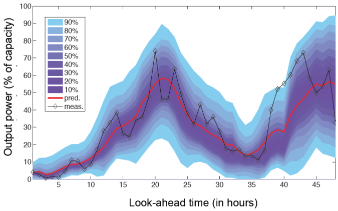
School of Electrical and Computer Engineering

Cornell University

CERTS Annual Review Meeting, Ithaca, Aug. 4, 2015

The Problem of Variability in Power Systems

Existing operational tools and electricity market designs not equipped to efficiently accommodate variability in renewable supply at scale.



1. **Predominant approach to economic dispatch** makes inefficient use of information

2. **Existing market designs** reflect this inefficiency of operation and rely on ad hoc schemes to allocate the cost of variability (e.g., A/S costs) to market participants

H. Madsen et al. (2011). *Forecasting Wind and Solar Power Production*.

Broad Project Objectives

(A) Develop **computational tools to manage uncertainty** in power system operations

- data driven
- computational scalable
- provable performance guarantees

(B) Design **novel market systems** that:

- provide a competitive medium through which variable power producers can sell their supply on equal footing with conventional power producers.
- efficiently reflect and *allocate the cost of uncertainty* owing to variable renewable generation.

Talk Outline

1. Power system model
2. Economic Dispatch as Stochastic Control
3. Policy approximation
4. Constraint approximation
5. Numerical Experiment

The Power System Model

The power system is modeled as a **partially-observed linear stochastic system**:

$$x_{t+1} = Ax_t + Bu_t + Gw_t$$
$$y_t = Cx_t + Hw_t$$

for all $t = 0, \ldots, T$.

- $x_t \in \mathbf{R}^{\mathsf{n_x}}$ (state) generator operating points, energy storage levels, etc.
- $u_t \in \mathbf{R}^{\mathsf{n_u}}$ (control) generator and storage power injections
- $y_t \in \mathbf{R}^{\mathsf{n_y}}$ (measurements) partial noisy observation of system state
- $w_t \in \mathbf{R}^{\mathsf{n_w}}$ (exogenous random process) wind, solar, demand, etc.

Notation

- $x^t = (x_0, \ldots, x_t) \in \mathbf{R}^{\mathsf{n_x(t+1)}}$ – history of process until time t
- $\mathbf{x} = (x_0, \ldots, x_{T-1}) \in \mathbf{R}^{\mathsf{n_x T}}$ – history of a enitre process

The Power System Model

The power system is subjected to **linear constraints**

$$F_x x_t + F_u u_t + F_w w_t \leq g,$$

which we **enforce probabilistically** as

$$\textbf{Prob}\left\{F_x x_t + F_u u_t + F_w w_t \leq g\right\} \;\geq\; 1 - \epsilon$$

for all $t = 0, \ldots, T$, where $\epsilon \in (0, 1]$.

Model Features

- **Uncertainty** in renewable supply, demand, network line outages
- **Transmission network constraints** subject to linearized power flow
- **Resource constraints** – storage, generator ramping, flexible demand
- **Doesn't capture** AC power flow or unit commitment decisions

The Stochastic Economic Dispatch (SED) Problem

Compute a causal output-feedback dispatch policy $\pi = (\mu_0, \ldots, \mu_{T-1})$

$$u_t = \mu_t(y_0, \ldots, y_t)$$

that solves **chance constrained stochastic control problem:**

$$
\begin{aligned}
(\mathcal{P}) \quad & \underset{\pi}{\text{minimize}} \quad && \mathbf{E}^{\pi}\left[x_T' Q x_T \;+\; \sum_{t=0}^{T-1} x_t' Q x_t \;+\; u_t' R u_t \right] \\
& \text{subject to} && \mathbf{Prob}\left\{ F_x x_t + F_u u_t + F_w w_t \le g \right\} \;\ge\; 1 - \epsilon \\
& && x_{t+1} = A x_t + B u_t + G w_t \\
& && y_t = C x_t + H w_t \\
& && u_t = \mu_t(y^t), \qquad t = 0, \ldots, T.
\end{aligned}
$$

We assume convex quadratic costs, $Q, R \succeq 0$.

Why is Problem $\mathcal{P}$ Difficult to Solve?

In general, problem $\mathcal{P}$ is difficult to solve...

- **Infinite-dimensional** in its optimization variables
- **Nonconvex** in its feasible region
- **Requires specification of a prior** probability distribution

We reformulate problem $\mathcal{P}$ as a **finite-dimensional convex program** that:

- is **computationally scalable**,
- requires **minimal distributional information**,
- admits **computable performance guarantees**,
- and enables the systematic trade off between **computational burden** and **performance**.

Why is Problem $\mathcal{P}$ Difficult to Solve?

In general, problem $\mathcal{P}$ is difficult to solve...

- **Infinite-dimensional** in its optimization variables
- **Nonconvex** in its feasible region
- **Requires specification of a prior** probability distribution

We reformulate problem $\mathcal{P}$ as a **finite-dimensional convex program** that:

- is **computationally scalable**,
- requires **minimal distributional information**,
- admits **computable performance guarantees**,
- and enables the systematic trade off between **computational burden** and **performance**.

Policy Approximation

An Output Transformation

Define the **purified observation process** $z_t = y_t - \widehat{y}_t$, where

$$\widehat{x}_{t+1} = A\widehat{x}_t + Bu_t \quad \text{and} \quad \widehat{y}_t = C\widehat{x}_t$$

The purified output z_t can be thought of as an output prediction error.

Lemma (Kailath '68, Ben-Tal et al. '09, Hadjiyiannis et al. '10)

1. $\{z_t\}$ **generates the same amount of information** as $\{y_t\}$

$$\sigma(z_0, \dots, z_t) = \sigma(y_0, \dots, y_t).$$

2. $\{z_t\}$ **is independent of the control process** $\{u_t\}$ and satisfies

$$z_t = L_t w^t,$$

where the matrix L_t is easily constructed from problem data. We write the entire purified output vector as

$$\mathbf{z} = L\mathbf{w}.$$

An Output Transformation

Define the **purified observation process** $z_t = y_t - \widehat{y}_t$, where

$$\widehat{x}_{t+1} = A\widehat{x}_t + Bu_t \qquad \text{and} \qquad \widehat{y}_t = C\widehat{x}_t$$

The purified output z_t can be thought of as an output prediction error.

Lemma (Kailath '68, Ben-Tal et al. '09, Hadjiyiannis et al. '10)

1. $\{z_t\}$ **generates the same amount of information** as $\{y_t\}$

$$\sigma(z_0, \ldots, z_t) = \sigma(y_0, \ldots, y_t).$$

2. $\{z_t\}$ **is independent of the control process** $\{u_t\}$ and satisfies

$$z_t = L_t \mathbf{w}^t,$$

where the matrix L_t is easily constructed from problem data. We write the entire purified output vector as

$$\mathbf{z} = L\mathbf{w}.$$

Purified Output-Feedback

Reparameterize the causal feedback policies in the purified observation $\{z_t\}$:

$$
\begin{aligned}
(\widehat{\mathcal{P}}) \qquad & \underset{\pi}{\text{minimize}} && \mathbf{E}^{\pi}\left[x_T' Q x_T \;+\; \sum_{t=0}^{T-1} x_t' Q x_t \;+\; u_t' R u_t \right] \\
& \text{subject to} && \mathbf{Prob}\left\{ F_x x_t + F_u u_t + F_w w_t \le g \right\} \;\ge\; 1 - \epsilon \\
& && x_{t+1} = A x_t + B u_t + G w_t \\
& && z_t = L_t w^t \\
& && u_t = \mu_t(z^t), \qquad t = 0, \dots, T.
\end{aligned}
$$

- Eliminates dependency of observations on control inputs
- This is **without loss of optimality**, i.e. problem $\widehat{\mathcal{P}}$ equivalent to $\mathcal{P}$
- Problem $\widehat{\mathcal{P}}$ still intractable however....

A Finite-Dimensional Approximation of the Policy Space

Restrict the space of admissible control policies to those representable as finite linear combinations of a **preselected basis functions** $\phi_t = (\phi_t^1, \ldots, \phi_t^{n_t})'$,

$$u_t = K_t \phi_t(z^t) = \begin{bmatrix} | \\ K_t^1 \\ | \end{bmatrix} \phi_t^1(z^t) + \cdots + \begin{bmatrix} | \\ K_t^{n_t} \\ | \end{bmatrix} \phi_t^{n_t}(z^t)$$

where $K_t \in \mathbf{R}^{\mathsf{n_u} \times \mathsf{n_t}}$ is matrix of weighting coefficients.

We write the entire input vector as $\mathbf{u} = K\phi(\mathbf{z})$, where

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{T-1} \end{bmatrix} = \begin{bmatrix} K_0 & 0 & \cdots & 0 \\ 0 & K_1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & K_{T-1} \end{bmatrix} \begin{bmatrix} \phi_0(z^0) \\ \phi_1(z^1) \\ \vdots \\ \phi_{T-1}(z^{T-1}) \end{bmatrix}$$

A Finite-Dimensional Approximation of the Policy Space

Restrict the space of admissible control policies to those representable as finite linear combinations of a **preselected basis functions** $\phi_t = (\phi_t^1, \ldots, \phi_t^{n_t})'$,

$$u_t \;=\; K_t \phi_t(z^t) \;=\; \begin{bmatrix} | \\ K_t^1 \\ | \end{bmatrix} \phi_t^1(z^t) \;+\; \cdots \;+\; \begin{bmatrix} | \\ K_t^{n_t} \\ | \end{bmatrix} \phi_t^{n_t}(z^t)$$

where $K_t \in \mathbf{R}^{\mathsf{n_u} \times \mathsf{n_t}}$ is matrix of weighting coefficients.

We write the entire input vector as $\mathbf{u} = K\phi(\mathbf{z})$, where

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{T-1} \end{bmatrix} = \begin{bmatrix} K_0 & 0 & \cdots & 0 \\ 0 & K_1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & K_{T-1} \end{bmatrix} \begin{bmatrix} \phi_0(z^0) \\ \phi_1(z^1) \\ \vdots \\ \phi_{T-1}(z^{T-1}) \end{bmatrix}$$

Reduction of $\mathcal{P}$ to a Chance Constrained Program

Proposition

Given a restriction to dispatch policies of the form

$$\mathbf{u} = K\phi(\mathbf{z}), \quad \text{where} \quad \mathbf{z} = L\mathbf{w},$$

problem $\mathcal{P}$ reduces to a **finite-dimensional chance constrained program**

$$
\begin{aligned}
(\mathcal{CCP}) \qquad & \underset{K}{\text{minimize}} && \mathbf{tr}(M_\phi K' V K) \; + \; \mathbf{tr}(N_\phi U K) \\
& \text{subject to} && \mathbf{Prob}\left\{E_t K\phi(\mathbf{z}) \; + \; F_t \mathbf{w} \le g_t\right\} \; \ge \; 1 - \epsilon \\
& && t = 0, \ldots, T \\
& && K = \text{diag}(K_0, \ldots, K_{T-1})
\end{aligned}
$$

Moment matrices given by $M_\phi := \mathbf{E}\left[\phi(\mathbf{z})\phi(\mathbf{z})'\right]$ and $N_\phi := \mathbf{E}\left[\phi(\mathbf{z})\mathbf{w}'\right]$.

Matrices $V \succeq 0$, U, and $\{E_t, F_t, g_t\}$ are computed from the primitive data.

Reduction to a Chance Constrained Program

$$
\begin{array}{lll}
(\mathcal{CCP}) & \underset{K}{\text{minimize}} & \mathbf{tr}(M_\phi K' V K) \; + \; \mathbf{tr}(N_\phi U K) \\
& \text{subject to} & \mathbf{Prob}\left\{ E_t K \phi(\mathbf{z}) \; + \; F_t \mathbf{w} \le g_t \right\} \; \ge \; 1 - \epsilon \\
& & t = 0, \ldots, T \\
& & K = \text{diag}(K_0, \ldots, K_{T-1})
\end{array}
$$

- $\mathcal{CCP}$ is a **inner approximation** of the original problem $\mathcal{P}$
- Objective is convex
- Feasible region is nonconvex, in general

There is a rich literature[1] on convex approximations of chance constraints

- convex inner approximations
- randomized convex approximations

1. Nemirovski et al. (2006). *Convex approximations of chance constrained programs*

Chance Constraint Approximation

A Scenario Constrained Program

Definition

Define the **Scenario Constrained Program** induced by $\mathcal{CCP}_\epsilon$ as

$$
\begin{array}{lll}
(\mathcal{SCP}_N) & \underset{K}{\text{minimize}} & \mathbf{tr}(M_\phi K' V K) \; + \; \mathbf{tr}(N_\phi U K) \\
& \text{subject to} & E_t K \phi(L\mathbf{w}^i) \; + \; F_t \mathbf{w}^i \leq g_t \\
& & t = 0, \ldots, T \\
& & i = 1, \ldots, N
\end{array}
$$

where $(\mathbf{w}^1, \ldots, \mathbf{w}^N)$ are N i.i.d. samples of the random vector $\mathbf{w}$.

$\mathcal{SCP}_N$ is a random convex **quadratic program (QP)**

- However, solutions to $\mathcal{SCP}_N$ are random and may not be feasible for $\mathcal{CCP}_\epsilon$
- How large does N have to be?

A Scenario Constrained Program

Definition

Define the **Scenario Constrained Program** induced by $\mathcal{CCP}_\epsilon$ as

$$
\begin{aligned}
(\mathcal{SCP}_N) \qquad & \underset{K}{\text{minimize}} \quad && \mathbf{tr}(M_\phi K' V K) \;+\; \mathbf{tr}(N_\phi U K) \\
& \text{subject to} && E_t K \phi(L\mathbf{w}^i) \;+\; F_t \mathbf{w}^i \le g_t \\
& && t = 0, \ldots, T \\
& && i = 1, \ldots, N
\end{aligned}
$$

where $(\mathbf{w}^1, \ldots, \mathbf{w}^N)$ are N i.i.d. samples of the random vector $\mathbf{w}$.

$\mathcal{SCP}_N$ is a random convex **quadratic program (QP)**

- However, solutions to $\mathcal{SCP}_N$ are random and may not be feasible for $\mathcal{CCP}_\epsilon$
- How large does N have to be?

Bound on the Number of Scenarios

The following result is an immediate consequence of Campi et al. (2008).*

Theorem

Fix a choice of basis ϕ and a probability level $\epsilon \in (0,1)$. If

$$N \;\geq\; \frac{1}{\epsilon}\left(\ln\frac{1}{\beta} \;+\; 1 \;+\; n_u \cdot \mathsf{card}(\phi)\right),$$

then an optimal solution to $\mathcal{SCP}_N$ will be feasible for $\mathcal{CCP}_\epsilon$ with probability greater than or equal to $1-\beta$.

*M.C. Campi et al. (2008). *The exact feasibility of randomized solutions of uncertain convex programs.*

Discussion of $\mathcal{SCP}_N$

How does $\mathcal{SCP}_N$ fair as a surrogate the original stochastic control problem $\mathcal{P}$?

Minimal Distributional Assumptions

- Results hold for any distribution on **w**
- Ability to procure independent samples of **w**

Polynomial-Time Complexity

- $\mathcal{SCP}_N$ is a convex QP
- It has dimension $=$ poly(n_y, n_u, T), if card$(\phi) \leq$ poly(n_y, T)
- Not exponential in the horizon T....

Fidelity

- Optimal solution to $\mathcal{SCP}_N$ is feasible for $\mathcal{P}$ with probability $1 - \beta$
- Richness of basis ϕ controls fidelity of approximation
- Similar techniques can be applied to obtain dual lower bounds for $\mathcal{P}$

Discussion of $\mathcal{SCP}_N$

How does $\mathcal{SCP}_N$ fair as a surrogate the original stochastic control problem $\mathcal{P}$?

Minimal Distributional Assumptions

- Results hold for any distribution on **w**
- Ability to procure independent samples of **w**

Polynomial-Time Complexity

- $\mathcal{SCP}_N$ is a convex QP
- It has dimension $=$ poly(n_y, n_u, T), if card$(\phi) \leq$ poly(n_y, T)
- Not exponential in the horizon T....

Fidelity

- Optimal solution to $\mathcal{SCP}_N$ is feasible for $\mathcal{P}$ with probability $1-\beta$
- Richness of basis ϕ controls fidelity of approximation
- Similar techniques can be applied to obtain dual lower bounds for $\mathcal{P}$

Discussion of $\mathcal{SCP}_N$

How does $\mathcal{SCP}_N$ fair as a surrogate the original stochastic control problem $\mathcal{P}$?

Minimal Distributional Assumptions

- Results hold for any distribution on **w**
- Ability to procure independent samples of **w**

Polynomial-Time Complexity

- $\mathcal{SCP}_N$ is a convex QP
- It has dimension $=$ poly(n_y, n_u, T), if card$(\phi) \leq$ poly(n_y, T)
- Not exponential in the horizon T....

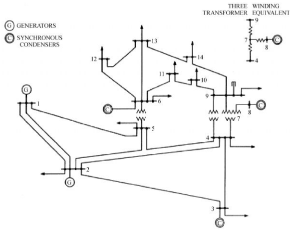
Fidelity

- Optimal solution to $\mathcal{SCP}_N$ is feasible for $\mathcal{P}$ with probability $1-\beta$
- Richness of basis ϕ controls fidelity of approximation
- Similar techniques can be applied to obtain dual lower bounds for $\mathcal{P}$

Numerical Experiment

Test Case Description

We consider a modified IEEE 14-bus power system dispatched over $T = 12$ hours.



System Features

- Constrained DC power flow
- Dispatchable thermal generation
- Energy storage
- Wind generation (30% penetration)

Wind power data acquired from NREL Eastern Wind Integration and Transmission Study (EWITS).

Experiment Description

Parameters

- Constraint probability, $\epsilon = 0.1$
- Approximation confidence, $\beta = 0.001$
- Basis, $\phi = \{$set of all d-degree monomials$\}$ in $\mathbf{z}$
    - $d = 1$ (affine control policies)
    - $d = 2$ (quadratic control policies)

Procedure

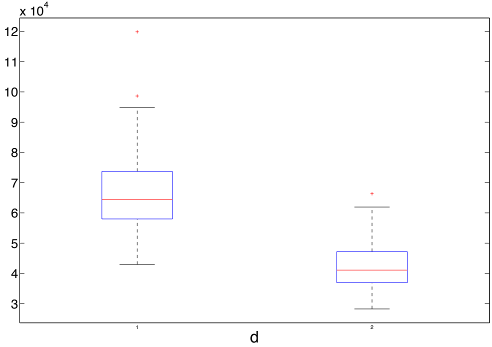
1. Fix the degree d
2. Set the sample size $N = N(\epsilon, \beta, \phi)$, where

$$N(\epsilon, \beta, \phi) = \frac{1}{\epsilon}\left(\ln\frac{1}{\beta} \ + \ 1 \ + \ n_u \binom{T \cdot n_y}{d}\right)$$

3. Solve 100 instances of $\mathcal{SCP}_N$

Empirical Results

Optimal cost distribution of $\mathcal{SCP}_N$ vs. d



There is a 36% reduction in average cost in moving from affine $(d = 1)$ to quadratic $(d = 2)$ policies.