

Risk-Sensitive Optimization for Power Systems

Avinash N. Madavan

May 10, 2021

Expecting uncertainty in power systems

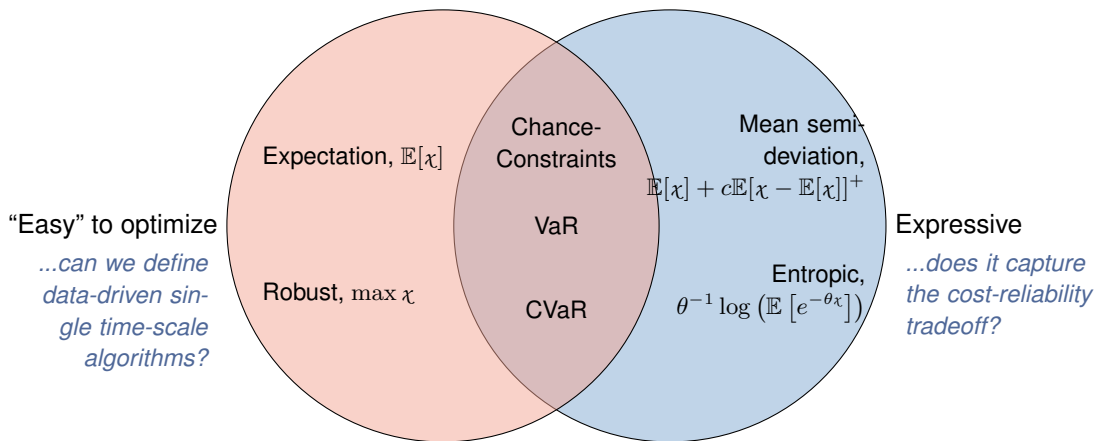
- ▶ Conventional power systems are composed of a grid connected to several large-scale dispatchable generators with predictable loads.
- ▶ This paradigm is shifting rapidly with
 - Increasing wind and solar plants that are intermittent, non-dispatchable, and uncertain.
 - Increasing adoption of distributed energy resources (rooftop solar or battery energy storage) and electric vehicles introducing further uncertainty and complexity.
 - Increasing frequency of extreme weather events posing greater risk of component availability and potential blackouts. For example, the 2021 Texas blackout.
- ▶ Modern models must more strongly account for uncertainty in component availability, available energy supply, and fluctuations in demand.

What is risk-sensitive optimization?

- For random functions, $f_\omega(x)$ and $g_\omega(x)$, we want to solve the (ill-formed) problem

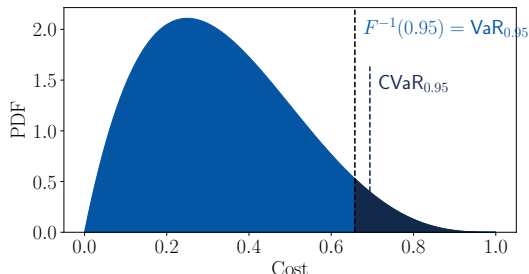
$$\underset{x}{\text{minimize}} \quad f_\omega(x), \quad \text{subject to} \quad g_\omega(x) \leq 0.$$

- We must extract some numerical property of these random functions to meaningfully optimize. These properties are referred to as *measures*.



Understanding the conditional value at risk (CVaR) measure

- It measures the average over the $(1 - \alpha)$ fraction of worst-case scenarios.



The CVaR measure has the properties:

- $\text{CVaR}_0[\chi] \equiv \mathbb{E}[\chi]$.
- $\lim_{\alpha \uparrow 1} \text{CVaR}_\alpha[\chi] = \text{ess sup } \chi$.
- $\text{CVaR}_\alpha[\chi] \geq \text{VaR}_\alpha[\chi]$.
- CVaR limits severity of exceedance.

- For general distributions, CVaR is defined by the following variational form

$$\text{CVaR}_\alpha[\chi] := \underset{z}{\text{minimize}} \quad z + \frac{1}{1 - \alpha} \mathbb{E}[\chi - z]^+$$

Rockafellar and Uryasev 2000; 2002

Summary of presentation

- ▶ **Why** do we need to model uncertainty in power system operation?
- ▶ **What** is risk-sensitive optimization and why does it address this concern?
- ▶ **How** do we formulate, optimize, and price risk-sensitive models?
 - Discrete uncertainty in a risk-sensitive security-constrained economic dispatch problem.
 - Continuous uncertainty in a risk-sensitive economic dispatch problem with uncertain wind.
 - Either uncertainty for voltage regulation in distribution networks.

The economic dispatch problem

- ▶ System operators make dispatch decisions at various intervals before the real time by solving an economic dispatch (ED) problem such as

$$\begin{aligned} & \underset{\mathbf{g}}{\text{minimize}} && \mathbf{c}^\top \mathbf{g}, \\ & \text{subject to} && \mathbf{1}^\top (\mathbf{g} - \mathbf{d}) = 0, \\ & && \mathbf{H}(\mathbf{g} - \mathbf{d}) \leq \mathbf{f}, \\ & && \underline{\mathbf{g}} \leq \mathbf{g} \leq \bar{\mathbf{g}}. \end{aligned}$$

- ▶ For simplicity, we adopt the DC approximations that reduce the nonlinear power flow equations into a linear map from the net power injections.
- ▶ This deterministic problem fails to account for:
 1. Potential component failures such as line outages.
 2. Uncertainty in available supply introduced by renewable sources.

A risk-sensitive security-constrained ED (R-SCED) problem

- For potential line failures, enumerated $k = 1, \dots, K$, we augment the ED problem with additional constraints.

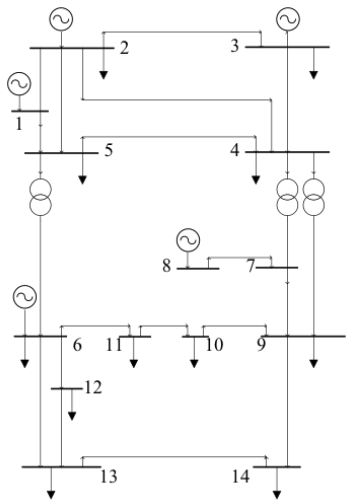
$$\begin{aligned} & \underset{g, \delta g, \delta d}{\text{minimize}} && \text{CVaR}_\alpha \left[c^\top g + C(\delta g, \delta d) \right], \\ & \text{subject to} && \mathbf{1}^\top (g - d) = 0, \quad \mathbf{H}^\top (g - d) \leq f, \quad g \in [\underline{g}, \bar{g}], \\ & && \mathbf{H}_k^\top (g - d) \leq f_k^{\text{DA}}, \\ & && \mathbf{H}_k^\top (g + \delta g_k - d + \delta d_k) \leq f_k^{\text{SE}}, \quad g + \delta g_k \in [\underline{g}, \bar{g}], \quad \delta g_k \in \Delta_g, \\ & && k = 1, \dots, K. \end{aligned}$$

where the random recourse cost $C(\delta g, \delta d)$ takes values

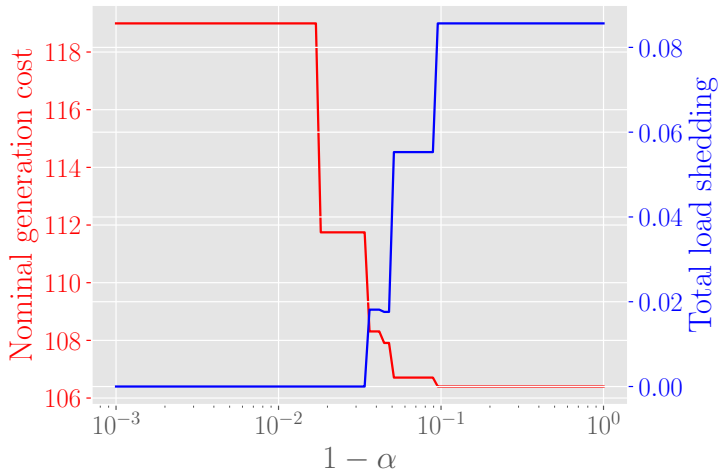
$$C_k(\delta g_k, \delta d_k) = c_+^\top [\delta g]^+ + c_-^\top [\delta g]^- + v^\top \delta d \quad \text{w.p. } p_k.$$

- Other security-constrained ED (SCED) formulations include
- Preventive-SCED: Alsac and Stott 1974; Capitanescu, Glavic, et al. 2007
 - Corrective-SCED: Capitanescu and Wehenkel 2007; 2008; Phan and Sun 2015 (without load shed), Francois Bouffard and Galiana 2008 (with load shed)

An illustrative example of R-SCED



IEEE 14-bus network



risk-averse ← → risk-neutral

The computational challenges of SCED at scale

- ▶ R-SCED and C-SCED are large linear programs which are *usually* considered easy.
- ▶ For the IEEE 2383-bus Polish network with 2,869 lines, 327 generators, and 1,817 loads

Formulation	Variables	Constraints
ED	327	6,447
C-SCED	947,319	28,949,071
R-SCED	6,212,249	28,954,865

“The resulting linear program is not solvable by traditional LP methods due to its large size”

-Liu et. al. 2015

- ▶ The following techniques have been used to make the problem more tractable:
 - Contingency filtering/pre-screening François Bouffard, Galiana, and Arroyo 2005; Capitanescu, Glavic, et al. 2007; Ejebe and Wollenberg 1979; Fliscounakis et al. 2013; Fu and A. Bose 1999
 - Exploiting decomposable structure Benders 2005; Gomez et al. 1991; Liu, Ferris, and Zhao 2015; Madavan, S. Bose, et al. 2019

A decomposition method for SCED problems

- The R-SCED (and C-SCED) formulation is an instance of the following with $\alpha' = (1 - \alpha)^{-1}$

$$\begin{aligned} & \underset{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_K}{\text{minimize}} && \mathbf{c}^\top \mathbf{x}_0 + \alpha' \sum_{k=1}^K p_k \mathbf{c}_k^\top \mathbf{x}_k, \\ & \text{subject to} && \mathbf{A}_0 \mathbf{x}_0 \leq \mathbf{b}_0, \\ & && \mathbf{E}_k \mathbf{x}_0 + \mathbf{A}_k \mathbf{x}_k \leq \mathbf{b}_k, \quad \text{for each } k = 1, \dots, K, \end{aligned}$$

- Making use of this structure, this can be written as

$$\underset{\mathbf{x}_0 \in \mathbb{X}_0}{\text{minimize}} \quad \mathbf{c}_0^\top \mathbf{x}_0 + \alpha' \sum_{k=1}^K J_k^*(\mathbf{x}_0),$$

where,

$$\begin{aligned} J_k^*(\mathbf{x}_0) &:= \underset{\mathbf{x}_k}{\text{minimize}} && p_k \mathbf{c}_k^\top \mathbf{x}_k, \\ & \text{subject to} && \mathbf{E}_k \mathbf{x}_0 + \mathbf{A}_k \mathbf{x}_k \leq \mathbf{b}_k. \end{aligned}$$

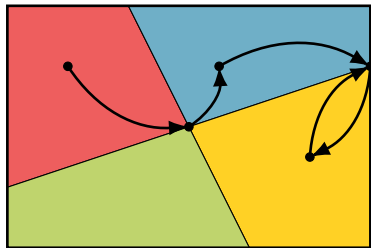
The critical region exploration (CRE) algorithm

- The decomposed form of R-SCED and C-SCED is given by

$$\begin{aligned} \underset{\mathbf{x}_0 \in \mathbb{X}_0}{\text{minimize}} \quad & \mathbf{c}_0^\top \mathbf{x}_0 + \alpha' \sum_{k=1}^K J_k^*(\mathbf{x}_0), \quad J_k^*(\mathbf{x}_0) := \underset{\mathbf{x}_k}{\text{minimize}} \quad p_k \mathbf{c}_k^\top \mathbf{x}_k, \\ & \text{subject to} \quad \mathbf{E}_k \mathbf{x}_0 + \mathbf{A}_k \mathbf{x}_k \leq \mathbf{b}_k. \end{aligned}$$

- The critical region exploration algorithm uses the following:

- J_k^* is a multi-parametric linear program (MPLP) linearly parameterized by \mathbf{x}_0 .
- J_k^* is piecewise affine and convex in \mathbf{x}_0 .
- The sets over which J_k^* is affine describe a polyhedral partition of \mathbb{X}_0 .



Theorem. CRE converges in finite iterations.

Preliminary results

- ▶ CRE and Benders' was implemented in C++ leveraging Eigen and SuiteSparse libraries. Available at: <https://github.com/amadavan/Stuka>.
- ▶ CRE is competitive with Benders' for small- and medium-sized problems but suffers from complexity issues for large-scale problems.
- ▶ Critical region defined by very large set of constraints, making main problem very large.
- ▶ Benders' requires several iterations to estimate global cost even if the initial point is close.
- ▶ Benders' solves all subproblems at each iteration while CRE solves a small subset.

Future work

- ▶ Design a composite algorithm making use of advantages of both approaches.
- ▶ Evaluate pricing schemes under the R-SCED formulation.

Avinash N. Madavan, Subhonmesh Bose, et al. (2019). "Risk-Sensitive Security-Constrained Economic Dispatch via Critical Region Exploration". In: *2019 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5

Modeling uncertainty in wind in the ED problem

- Consider the ED problem augmented with some forecasted renewable supply

$$\begin{aligned}\mathcal{P}_{\text{det}} : \quad & \underset{\mathbf{g} \in [\underline{\mathbf{g}}, \bar{\mathbf{g}}]}{\text{minimize}} && \mathbf{c}^\top \mathbf{g}, \\ & \text{subject to} && \mathbf{1}^\top (\mathbf{g} - \mathbf{d} + \boldsymbol{\omega}) = 0, \\ & && \mathbf{H}^\top (\mathbf{g} - \mathbf{d} + \boldsymbol{\omega}) \leq \mathbf{f}.\end{aligned}$$

- When the forecast is incorrect, the dispatch will change accordingly, i.e. $\mathbf{g} = \mathbf{g}(\boldsymbol{\omega})$. For simplicity, assume affine recourse policies, with $\Delta\boldsymbol{\omega} = \boldsymbol{\omega} - \mathbb{E}[\boldsymbol{\omega}]$,

$$\mathbf{g}(\boldsymbol{\omega}) = \hat{\mathbf{g}} + \mathbf{G}\Delta\boldsymbol{\omega}.$$

- The risk-sensitive problem is given by

$$\begin{aligned}\mathcal{P}_{\text{risk}} : \quad & \underset{\hat{\mathbf{g}}, \mathbf{G}}{\text{minimize}} && \text{CVaR}_\alpha \left[\mathbf{c}^\top (\hat{\mathbf{g}} + \mathbf{G}\Delta\boldsymbol{\omega}) \right], \\ & \text{subject to} && \mathbf{1}^\top (\hat{\mathbf{g}} + \mathbf{G}\Delta\boldsymbol{\omega} - \mathbf{d} + \boldsymbol{\omega}) = 0 \quad \text{a.s.}, \\ & && \text{CVaR}_{\beta\ell} \left[\mathbf{H}^\top (\hat{\mathbf{g}} + \mathbf{G}\Delta\boldsymbol{\omega} - \mathbf{d} + \boldsymbol{\omega}) - \mathbf{f} \right] \leq 0, \\ & && \text{CVaR}_{\beta g} [\hat{\mathbf{g}} + \mathbf{G}\Delta\boldsymbol{\omega} - \bar{\mathbf{g}}] \leq 0, \quad \text{CVaR}_{\beta g} [\underline{\mathbf{g}} - \hat{\mathbf{g}} - \mathbf{G}\Delta\boldsymbol{\omega}] \leq 0.\end{aligned}$$

Physically enforced

A stochastic approximation algorithm

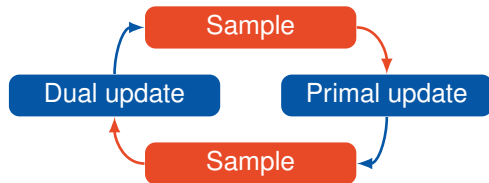
- The $\mathcal{P}_{\text{risk}}$ problem can be written more generally as

$$\begin{aligned} \mathcal{P}^{\text{CVaR}} : \quad & \underset{\mathbf{x}}{\text{minimize}} \quad \text{CVaR}_{\alpha} [f_{\omega}(\mathbf{x})], \\ & \text{subject to} \quad \text{CVaR}_{\beta} [g_{\omega}^i(\mathbf{x})] \leq 0, \quad i = 1, \dots, m. \end{aligned}$$

- With the variational form of CVaR and appropriately modified \mathbf{x} , $f_{\omega}(\mathbf{x})$, $\mathbf{g}_{\omega}(\mathbf{x})$, this can be

$$\begin{aligned} \mathcal{P}^{\text{E}} : \quad & \underset{\mathbf{x}}{\text{minimize}} \quad \mathbb{E} [f_{\omega}(\mathbf{x})], \\ & \text{subject to} \quad \mathbb{E} [g_{\omega}^i(\mathbf{x})] \leq 0, \quad i = 1, \dots, m. \end{aligned}$$

- We consider a primal-dual stochastic subgradient (PDSS) algorithm for the problem



$$\begin{aligned} \mathcal{L}_{\omega}(\mathbf{x}, \mathbf{z}) &= f_{\omega}(\mathbf{x}) + \mathbf{z}^{\top} \mathbf{g}_{\omega}(\mathbf{x}), \\ \mathbf{x}_{k+1} &= \mathbf{x}_k - \gamma \nabla_{\mathbf{x}} \mathcal{L}_{\omega_k}(\mathbf{x}_k, \mathbf{z}_k), \\ \mathbf{z}_{k+1} &= \left[\mathbf{z}_k + \gamma \nabla_{\mathbf{z}} \mathcal{L}_{\omega_{k+1/2}}(\mathbf{x}_{k+1}, \mathbf{z}_k) \right]^+. \end{aligned}$$

Converges without knowing dual bound!

Theoretical guarantees of the PDSS algorithm

Theorem (Asymptotic convergence). Under certain assumptions and $\{\gamma_k\}_{k=1}^{\infty}$ non-summable square-summable nonnegative sequence, i.e., $\sum_{k=1}^{\infty} \gamma_k = \infty, \sum_{k=1}^{\infty} \gamma_k^2 < \infty$. Then, $(\mathbf{x}_k, \mathbf{z}_k)$ from PDSS remains bounded and $\lim_{k \rightarrow \infty} \mathcal{L}(\mathbf{x}_k, \mathbf{z}_k) - \mathcal{L}(\mathbf{x}_\star, \mathbf{z}_k) = 0$

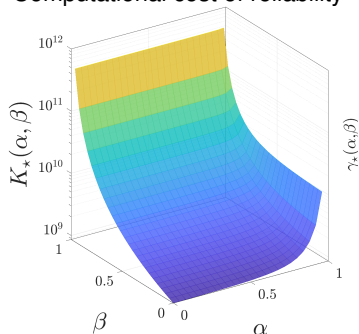
Theorem (Finite-time performance). Under certain assumptions, specified constants P_1, P_2, P_3 , and positive sequence $\{\gamma_k\}_{k=1}^K$, with $P_3 \sum_{k=1}^K \gamma_k^2 < 1$, then the iterates satisfy

$$\mathbb{E}[F(\bar{\mathbf{x}}_{K+1})] - p_\star^{\mathbb{E}} \leq \frac{1}{4 \sum_{k=1}^K \gamma_k} \left(\frac{P_1 + P_2 \sum_{k=1}^K \gamma_k^2}{1 - P_3 \sum_{k=1}^K \gamma_k^2} \right),$$
$$\mathbb{E}[G^i(\bar{\mathbf{x}}_{K+1})] \leq \frac{1}{4 \sum_{k=1}^K \gamma_k} \left(\frac{P_1 + P_2 \sum_{k=1}^K \gamma_k^2}{1 - P_3 \sum_{k=1}^K \gamma_k^2} \right)$$

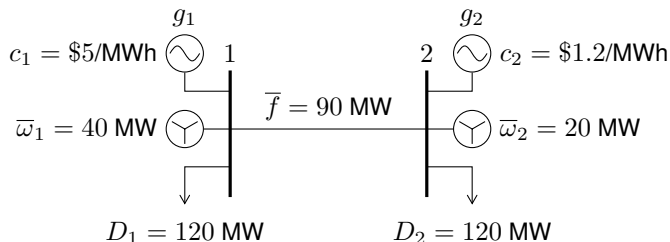
for each $i = 1, \dots, m$, where $\bar{\mathbf{x}}_{K+1} := \frac{\sum_{k=1}^K \gamma_k \mathbf{x}_{k+1}}{\sum_{k=1}^K \gamma_k}$.

Illustrating PDSS with examples

Computational cost of reliability



A two-bus network example with wind sample pairs (ω_1, ω_2) taken from NREL's synthetic dataset over 5 min intervals between 2008-2011.



Aversion	α	β	γ	Highest Cost (\$)	Worst-Case Line (MW)	Constraint Violation Generation (MW)	Iterations ($\times 10^9$)	Runtime (s)
None	0	0	0	400	25	24	2.9	3767
Line capacity	0	0.6	0.2	402	0	50	9.2	21400
Dispatch cost	0.6	0.2	0.2	352	13	0	4.6	9660

Pricing for the risk-sensitive ED problem

- Assume $\alpha = 0$, and replace the almost sure constraint using

$$\mathbf{1}^\top (\hat{\mathbf{g}} + \mathbf{G}\Delta\boldsymbol{\omega} - \mathbf{d} + \boldsymbol{\omega}) = 0 \text{ a.s.} \iff \mathbf{1}^\top (\hat{\mathbf{g}} - \mathbf{d} + \mathbb{E}[\boldsymbol{\omega}]) = 0, \quad \mathbf{G}^\top \mathbf{1} = \mathbf{1}.$$

- Applying the variational form, the risk-sensitive ED problem can be written as

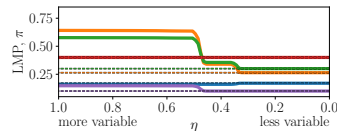
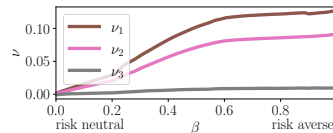
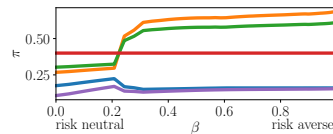
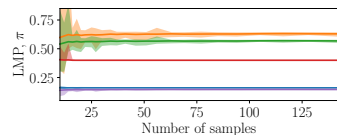
$$\begin{aligned} & \underset{\mathbf{g}_0, \mathbf{G}, \mathbf{u}, \underline{\mathbf{v}}, \bar{\mathbf{v}}}{\text{minimize}} && \mathbf{c}^\top \mathbf{g}_0, \\ & \text{subject to} && \mathbf{1}^\top (\mathbf{g}_0 + \mathbb{E}[\boldsymbol{\xi}] - \mathbf{d}) = 0, \quad \mathbf{G}^\top \mathbf{1} = \mathbf{1} && : \lambda, \boldsymbol{\nu} \\ & && \mathbf{u} + \frac{1}{\beta} \mathbb{E} [(\mathbf{H}(\mathbf{g}_0 - \mathbf{G}\Delta\boldsymbol{\xi} + \boldsymbol{\xi} - \mathbf{d}) - \mathbf{f} - \mathbf{u})^+] \leq 0, && : \boldsymbol{\mu} \\ & && \underline{\mathbf{v}} + \frac{1}{\gamma} \mathbb{E} [(-\mathbf{g}_0 + \mathbf{G}\Delta\boldsymbol{\xi} - \underline{\mathbf{v}})^+] \leq 0, \\ & && \bar{\mathbf{v}} + \frac{1}{\gamma} \mathbb{E} [(\mathbf{g}_0 - \mathbf{G}\Delta\boldsymbol{\xi} - \bar{\mathbf{g}} - \bar{\mathbf{v}})^+] \leq 0. \end{aligned}$$

The risk-sensitive locational marginal prices (risk-LMPs): $\boldsymbol{\pi} = \lambda \mathbf{1} - \mathbf{H}^\top \boldsymbol{\mu}$.

Properties of the risk-sensitive ED prices

- ▶ The risk-sensitive locational marginal prices (risk-LMPs), $\pi = \lambda \mathbf{1} - \mathbf{H}^\top \boldsymbol{\mu}$, satisfy
 - Equal to marginal sensitivity of optimal cost to nodal demands.
 - Varies with congestion through $\boldsymbol{\mu}$.
 - Uniform under no congestion.
 - Equal to classical LMP for deterministic wind.
- ▶ The payment scheme for players at bus i is
 - Demander pays $\pi_i d_i$.
 - Wind producer paid $\pi_i \mathbb{E}[\omega_i] - \nu_i$.
 - Dispatchable generator paid $\pi_i \bar{g}_i + \sum_{k=1}^n G_{ik} \nu_k$.
- ▶ Merchandising surplus $\geq \boldsymbol{\mu}^\top \mathbf{f} + \text{risk-dependent term}$.

Risk-LMPs generalize traditional LMPs to risk-sensitive case.



Preliminary Results

- ▶ Created a performant, open-source C++ implementation of PDSS algorithm. Available at: <https://github.com/amadavan/Stuka>.
- ▶ Sample-average approximation approach can solve risk-sensitive problem with good accuracy while PDSS is slow but has strong theoretical guarantees.

Future work

- ▶ Leverage power systems problem structure to improve runtime of the PDSS algorithm.
- ▶ Evaluate pricing under model with uncertain wind and unit commitment.
- ▶ Consider a distributed variant and extend the i.i.d. sampling model of PDSS to the Markov case.

Avinash N. Madavan and Subhonmesh Bose (2021). *A Stochastic Primal-Dual Method for Optimization with Conditional Value at Risk Constraints*. arXiv: 1908.01086 [math.OC]

Avinash N. Madavan and Subhonmesh Bose (2019). “Risk-Sensitive Energy Procurement with Uncertain Wind”. In: *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1–5

Mariola Ndrio, Avinash N. Madavan, and Subhonmesh Bose (2021). “Conditional-Value-at-Risk-Sensitive Locational Marginal Pricing for Electricity Markets”. In: *2021 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE

Dispatching for distribution networks

- ▶ Volt-VAR control (VVC) seeks to determine optimal control actions for voltage regulating and VAR control devices.
- ▶ Traditionally, VVC uses models of the network. However, network models are notoriously unreliable for distribution networks due to size and poor record-keeping.
- ▶ Model-free methods, such as reinforcement learning (RL), offer a solution for this problem.
- ▶ There are two important considerations when applying RL to power systems decision-making:
 - The applied policy must satisfy the constraints of the grid, such as line flow capacity limits, and resource limits.
 - The policy will be trained on a faulty or simulated model and must be able to provide strong out-of-sample guarantees.

Applying a simulator-trained model to the real world

- ▶ Safe RL studies methods to avoid risky rewards or satisfy constraints.
 - Risky policy avoidance: Gaskett 2003; Heger 1994 (worst-case), Basu, Bhattacharyya, and Borkar 2008; Borkar 2001; 2002; Howard and Matheson 1972; Kashima 2007 (risk-sensitive)
 - Constraint satisfaction: Di Castro, Tamar, and Mannor 2012; Kadota, Kurano, and Yasuda 2006; Moldovan and Abbeel 2012
- ▶ Evaluation of the optimal policy typically requires exploration over adverse states, the effect of which can be mitigated through a simulator.
- ▶ Robust RL seeks to establish strong out-of-sample guarantees when the transition kernel is not reflective of the true system (ex. from a simulator). Iyengar 2005; Nilim and El Ghaoui 2005; Wiesemann, Kuhn, and Rustem 2013
- ▶ Robust RL, distributionally robust optimization (DRO) and CVaR are closely related.

Robust RL \longrightarrow DRO \longleftarrow CVaR-optimization
- ▶ Based on the literature behind safe RL and robust RL, we expect that these notions can be combined using primal-dual analysis and DRO.

Summary of presentation

- ▶ **Why** do we need to model uncertainty in power system operation?
- ▶ **What** is risk-sensitive optimization and why does it address this concern?
- ▶ **How** do we formulate, optimize, and price risk-sensitive models?
 - Discrete uncertainty in a risk-sensitive security-constrained economic dispatch problem.
 - Continuous uncertainty in a risk-sensitive economic dispatch problem with uncertain wind.
 - Either uncertainty for voltage regulation in distribution networks.

My collaborators



Subhonmesh Bose
UIUC



Mariola Ndrino
UIUC



Lang Tong
Cornell



Ye Guo
*Tsinghua-Berkeley Shenzhen
Institute*



Hassan Hijazi
Los Alamos National Labs

Thank you!

`amadavan.github.io`



Appendix: CRE algorithm stated in full

Algorithm 1: A critical region exploration algorithm for R-SCED.

Initialization: Choose $x_0 \in \mathbb{X}_0$, $J^* = \infty$, $\mathbb{D} = \emptyset$ and a small positive ϵ .

```
1 while  $v^* \neq 0$  do
2   Given  $x^0$ , compute  $\rho_k, \eta_k, \mathbb{C}_k$  for  $k = 1, \dots, K$ .
3   Solve main problem over current critical region.
4    $[x^0]^{\text{opt}} \leftarrow$  lexicographically smallest minimizer of step 3
5    $J^{\text{opt}} \leftarrow$  optimal cost of step 3
6   if  $J^{\text{opt}} < J^*$  then
7      $[x^0]^* \leftarrow [x^0]^{\text{opt}}, J^* \leftarrow J^{\text{opt}}, \mathbb{D} \leftarrow \{c\}$ 
8   else
9      $\mathbb{D} \leftarrow \mathbb{D} \cup \{c^0 + \alpha' \sum_{k=1}^K \rho_k\}$ 
10     $v^* \leftarrow \operatorname{argmin}_{v \in \operatorname{conv}(\mathbb{D}) + N_{\mathbb{X}_0}([x^0]^*)} \|v\|^2$ 
11     $x^0 \leftarrow [x^0]^{\text{opt}} - \epsilon v^*$ 
```

Appendix: PDSS algorithm stated in full

Algorithm 2: Primal-dual stochastic subgradient method for \mathcal{P}^E .

Initialization: Choose $\mathbf{x}_1 \in \mathbb{X}$, $\mathbf{z}_1 = 0$, and a positive sequence γ .

1 **for** $k \geq 1$ **do**

2 Sample $\omega_k \in \Omega$. Update \mathbf{x} as

$$\mathbf{x}_{k+1} \leftarrow \operatorname{argmin}_{\mathbf{x} \in \mathbb{X}} \left\langle \nabla f_{\omega_k}(\mathbf{x}_k) + \sum_{i=1}^m z_k^i \nabla g_{\omega_k}^i(\mathbf{x}_k), \mathbf{x} - \mathbf{x}_k \right\rangle + \frac{1}{2\gamma_k} \|\mathbf{x} - \mathbf{x}_k\|^2. \quad (1)$$

3 Sample $\omega_{k+1/2} \in \Omega$. Update \mathbf{z} as

$$\mathbf{z}_{k+1} \leftarrow \operatorname{argmax}_{\mathbf{z} \in \mathbb{R}_+^m} \left\langle \mathbf{g}_{\omega_{k+1/2}}(\mathbf{x}_{k+1}), \mathbf{z} - \mathbf{z}_k \right\rangle - \frac{1}{2\gamma_k} \|\mathbf{z} - \mathbf{z}_k\|^2. \quad (2)$$