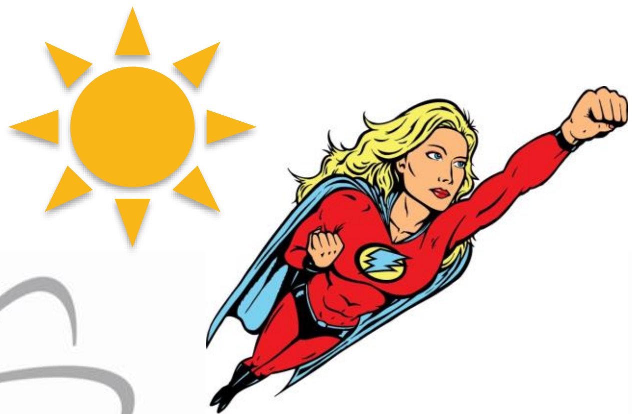
Statistical Learning for Optimization (and Control)

An Active Set Approach



Los Alamos NATIONAL LABORATORY EST. 1943

Seminar in Electric Power Systems

Line A. Roald (LANL)

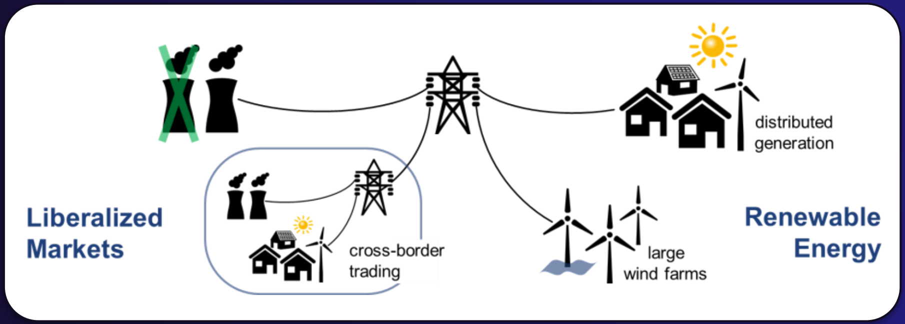
with **Sidhant Misra** (LANL)

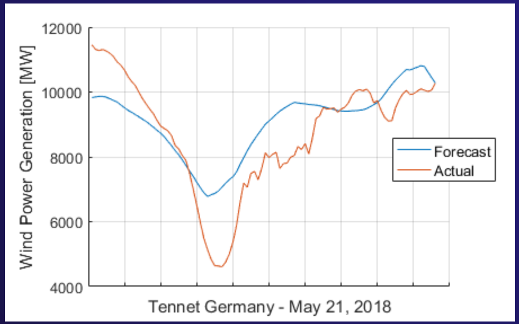
and **Yee Sian Ng** (MIT)

June 6, 2018

Every Day Decision Making Problems

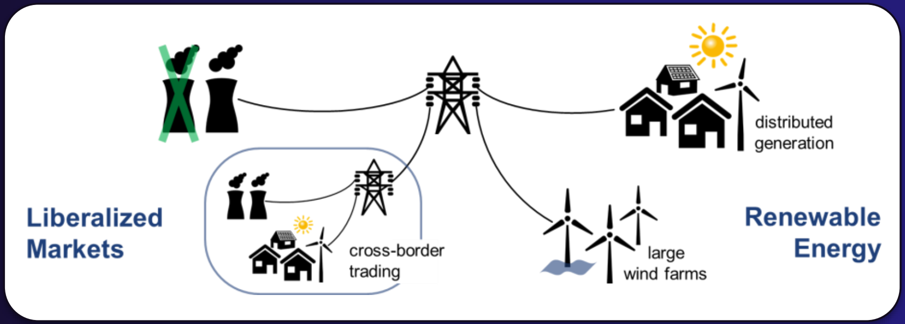
Energy System Optimization

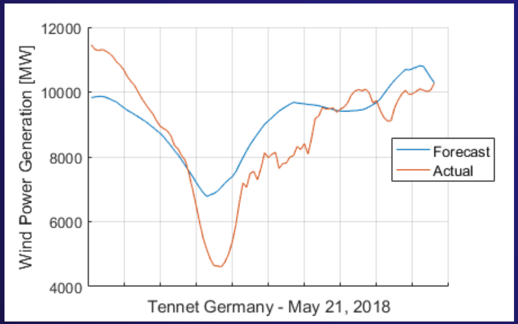




Model predictive/optimal control

Every Day Decision Making Problems





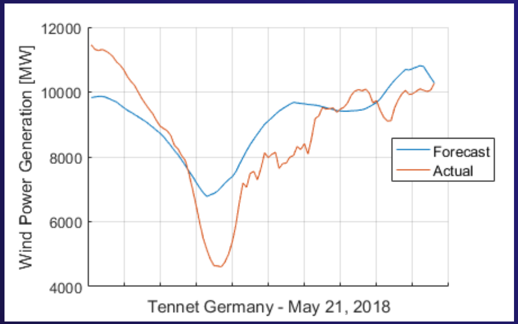
Relies increasingly on solving **optimization problems**

Having **fast** and **reliable** solution processes matters!

Every Day Decision Making Problems

Motivating example:

Optimal operation of electric grids with **variable renewable energy**



The Optimal Power Flow Problem

DC Optimal Power Flow

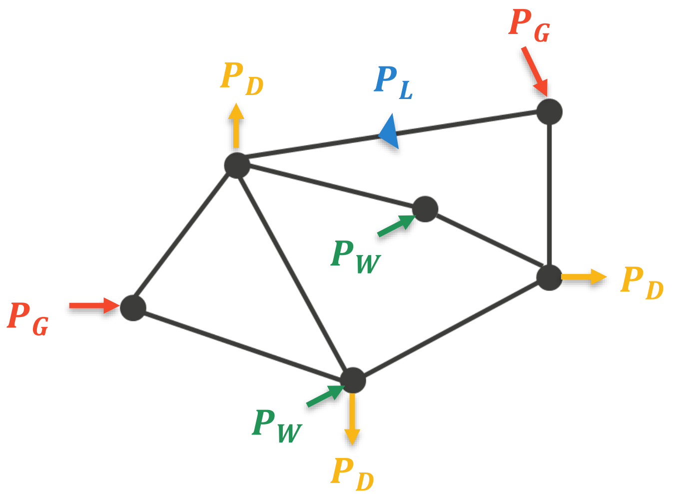
Goal: Low cost operation, while enforcing technical limits

P_G: **Controllable generation**

P_D: **System load**

P_W: **Renewable generation (uncontrollable, fluctuating)**

P_L: **Transmission line flow**



DC Optimal Power Flow

Goal: Low cost operation, while enforcing technical limits

$$\min_{P_G} \quad C_G^T P_G$$

minimize generation cost

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)} + P_{W(i)} - P_{D(i)}\right) = 0$$

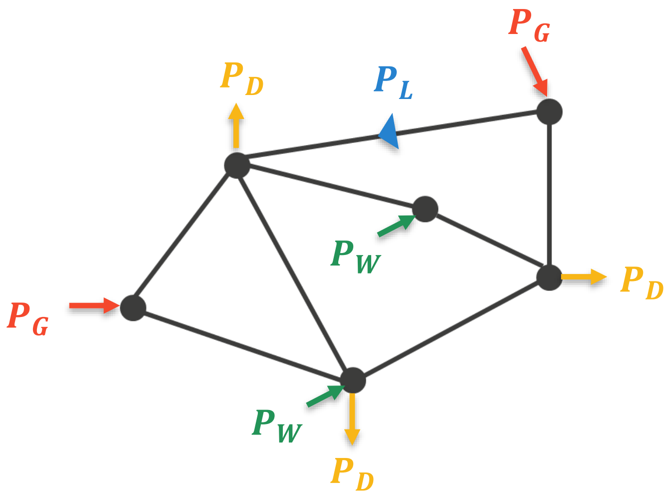
balanced operation

$$P_{G(g)} \leq P_{G(g)}^{max}\,,$$
$$P_{G(g)} \geq P_{G(g)}^{min}\,,$$
$$\forall\, g = 1, \dots, N_G$$

generator limits

$$A_{(l,\cdot)}(P_G + P_W - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G + P_W - P_D) \geq -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

transmission line limits



DC Optimal Power Flow

Renewable energy is unpredictable!

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

minimize generation cost

s.t. $\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$

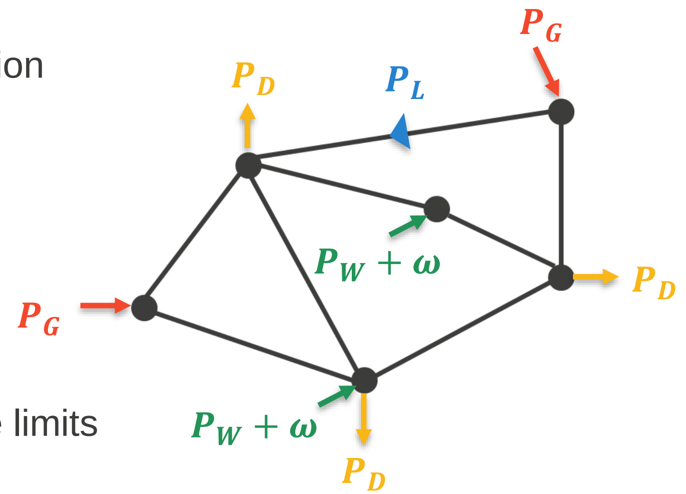
balanced operation

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min}, \quad \forall\, g = 1, \dots, N_G$$

generator limits

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max}, \quad \forall\, l = 1, \dots, N_L$$

transmission line limits



DC Optimal Power Flow

Renewable energy is unpredictable! → Resolve problem for each new $\boldsymbol{\omega}$ to obtain $P_G^*(\boldsymbol{\omega})$ **every 5-15 min**

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

minimize generation cost

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

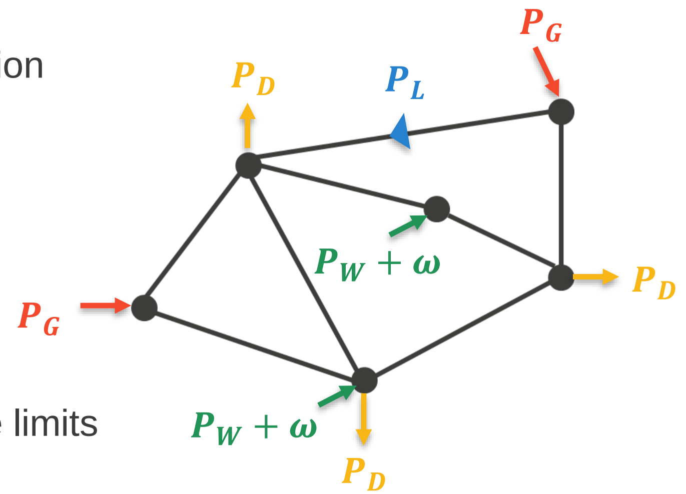
balanced operation

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min}, \quad \forall\, g = 1, \dots, N_G$$

generator limits

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max}, \quad \forall\, l = 1, \dots, N_L$$

transmission line limits



Repeated Solution Process

OPF at $T_1, \boldsymbol{\omega_1}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

s.t. $\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max} \,,$$
$$P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min} \,,$$
$$\forall \, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max},$$
$$\forall \, l = 1, \dots, N_L$$

Repeated Solution Process

OPF at $T_1, \boldsymbol{\omega}_1$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

OPF at $T_2, \boldsymbol{\omega}_2$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

Repeated Solution Process

OPF at $T_1, \boldsymbol{\omega_1}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

OPF at $T_2, \boldsymbol{\omega_2}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

OPF at $T_3, \boldsymbol{\omega_3}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

Repeated Solution Process

OPF at $T_1, \boldsymbol{\omega_1}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

s.t. $\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$

$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max}$,
$P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min}$,
$\forall\ g = 1, \dots, N_G$

$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max}$,
$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max}$,
$\forall\ l = 1, \dots, N_L$

OPF at $T_2, \boldsymbol{\omega_2}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

s.t. $\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$

$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max}$,
$P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min}$,
$\forall\ g = 1, \dots, N_G$

$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max}$,
$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max}$,
$\forall\ l = 1, \dots, N_L$

OPF at $T_3, \boldsymbol{\omega_3}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

s.t. $\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$

$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max}$,
$P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min}$,
$\forall\ g = 1, \dots, N_G$

$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max}$,
$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max}$,
$\forall\ l = 1, \dots, N_L$

....

Repeated Solution Process

OPF at $T_1, \boldsymbol{\omega_1}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)}\right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min}, \quad \forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max}, \quad \forall\, l = 1, \dots, N_L$$

OPF at $T_2, \boldsymbol{\omega_2}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)}\right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min}, \quad \forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max}, \quad \forall\, l = 1, \dots, N_L$$

OPF at $T_3, \boldsymbol{\omega_3}$

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)}\right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min}, \quad \forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max}, \quad \forall\, l = 1, \dots, N_L$$

....

Can we use learning to speed up the solution process

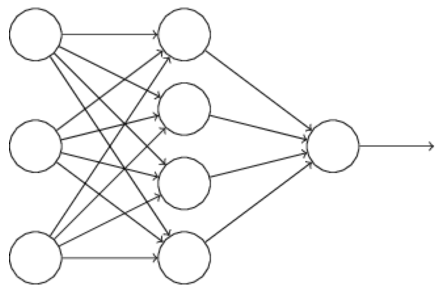
by using information from past solutions $\left(\boldsymbol{\omega_i}, P(\boldsymbol{\omega_i})\right)$?

First attempt:

Train a neural net

First Attempt – Train a Neural Net

ω_i

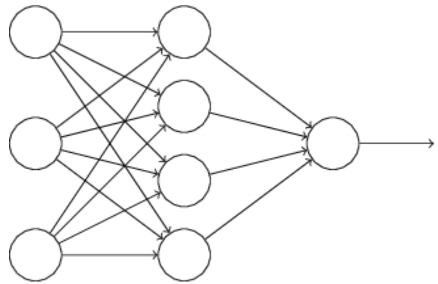


$P_G(\omega_i)$

- **This didn't work well...**
  - Hard to enforce safety constraints
  - Projection of back into feasible domain → sub-optimal solution
  - Sample complexity for the learning step is high

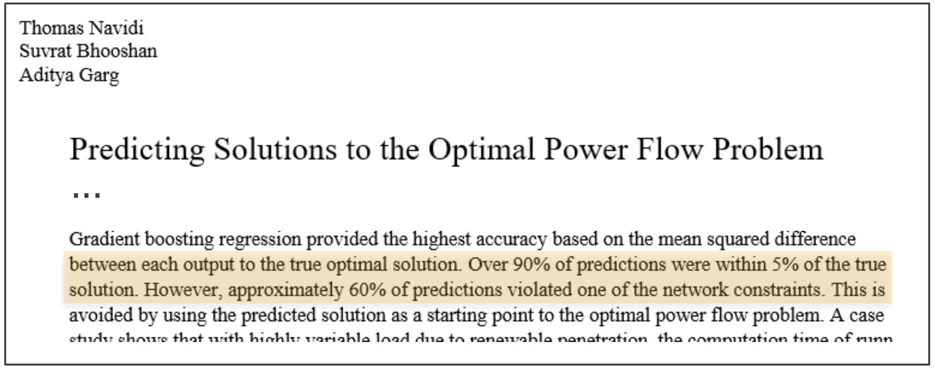
First Attempt – Train a Neural Net

ω_i



$P_G(\omega_i)$

- **General problems when applying machine learning!**
  - Issues with constraint violation



Thomas Navidi
Suvrat Bhooshan
Aditya Garg

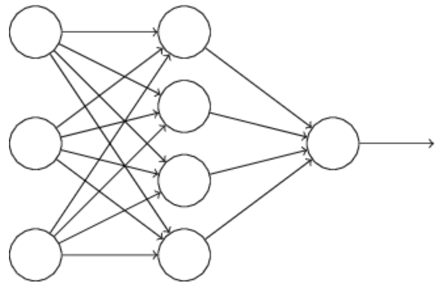
Predicting Solutions to the Optimal Power Flow Problem

...

Gradient boosting regression provided the highest accuracy based on the mean squared difference between each output to the true optimal solution. Over 90% of predictions were within 5% of the true solution. However, approximately 60% of predictions violated one of the network constraints. This is avoided by using the predicted solution as a starting point to the optimal power flow problem. A case study shows that with highly variable load due to renewable penetration, the computation time of runn

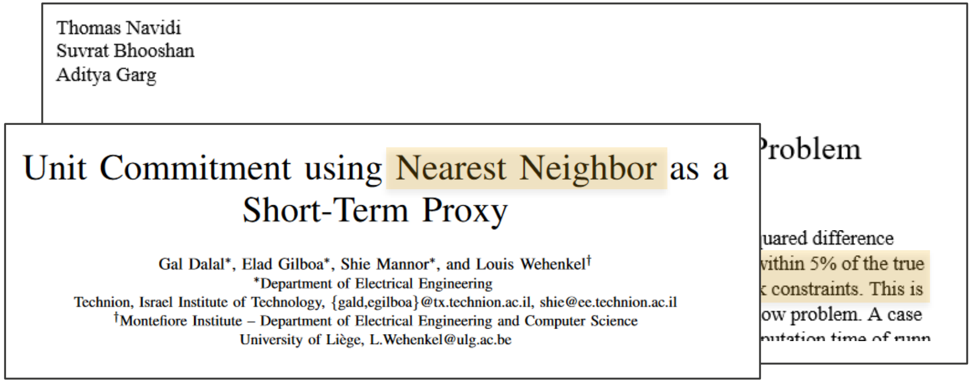
First Attempt – Train a Neural Net

ω_i



$P_G(\omega_i)$

- **General problems when applying machine learning!**
  - Issues with constraint violation
  - Simplified (suboptimal) solutions



Thomas Navidi
Suvrat Bhooshan
Aditya Garg

Problem

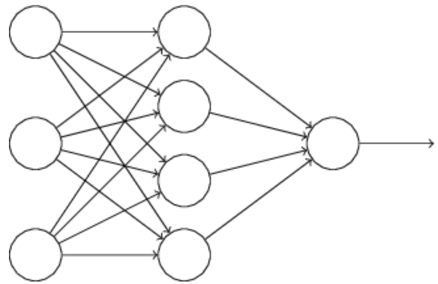
uared difference
within 5% of the true
constraints. This is
ow problem. A case

Unit Commitment using Nearest Neighbor as a Short-Term Proxy

Gal Dalal*, Elad Gilboa*, Shie Mannor*, and Louis Wehenkel†
*Department of Electrical Engineering
Technion, Israel Institute of Technology, {gald,egilboa}@tx.technion.ac.il, shie@ee.technion.ac.il
†Montefiore Institute – Department of Electrical Engineering and Computer Science
University of Liège, L.Wehenkel@ulg.ac.be

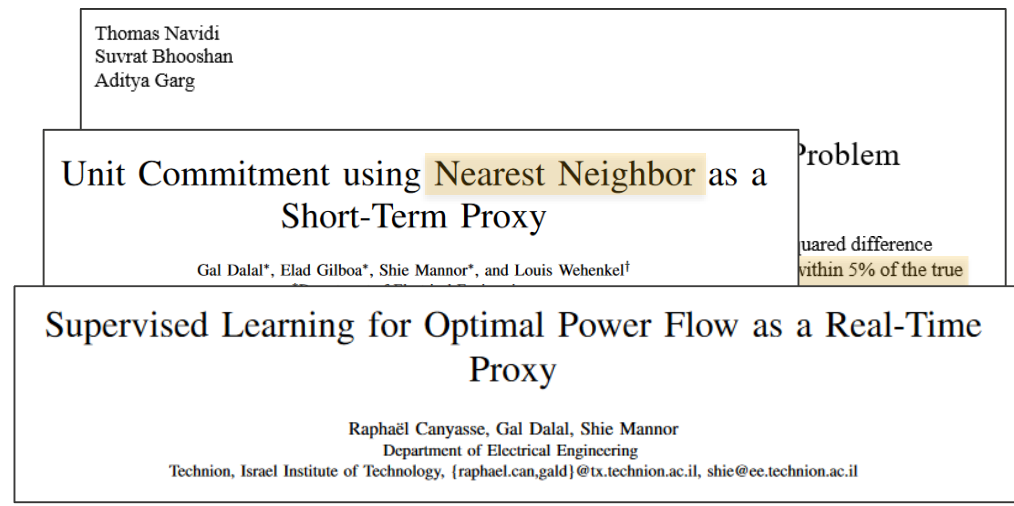
First Attempt – Train a Neural Net

ω_i



$P_G(\omega_i)$

- **General problems when applying machine learning!**
  - Issues with constraint violation
  - Simplified (suboptimal) solutions
  - Predict only infeasibility/ objective value



Thomas Navidi
Suvrat Bhooshan
Aditya Garg

Problem

uared difference
within 5% of the true

Unit Commitment using Nearest Neighbor as a Short-Term Proxy

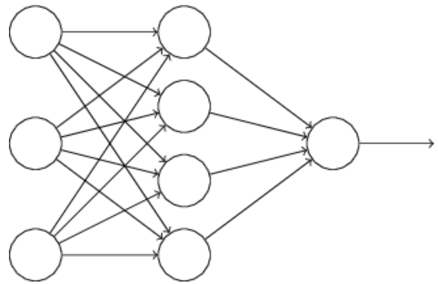
Gal Dalal*, Elad Gilboa*, Shie Mannor*, and Louis Wehenkel†

Supervised Learning for Optimal Power Flow as a Real-Time Proxy

Raphaël Canyasse, Gal Dalal, Shie Mannor
Department of Electrical Engineering
Technion, Israel Institute of Technology, {raphael.can,gald}@tx.technion.ac.il, shie@ee.technion.ac.il

First Attempt – Train a Neural Net

ω_i



$P_G(\omega_i)$

We have a **parametric optimization problem**

– can we use **more information** about the **problem structure**?

Think again:

How can we leverage pre-existing knowledge about the solution?

What do we already know?

Full mathematical model is known!

– Solution must satisfy physical laws

– Solution must respect safety constraints

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

s.t. $\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)}\right) = 0$ physical laws

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

safety constraints

What do we already know?

Full mathematical model is known!

– Solution must satisfy physical laws

– Solution must respect safety constraints

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)}\right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

New Idea:
Learn Active Set!

- Inherently encodes problem constraints
- Finite number
- Meaningful physical interpretation

Recovering Feasible Solution from Active Set

- **DC Optimal Power Flow Problem**
  - Linear program with n variables → there is a *basic feasible* optimal solution
  - Active contraints: Power balance and $(n-1)$ inequalities

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max},$$
$$P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min},$$
$$\forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max},$$
$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max},$$
$$\forall\, l = 1, \dots, N_L$$

Recovering Feasible Solution from Active Set

- **DC Optimal Power Flow Problem**
  - Linear program with n variables → there is a *basic feasible* optimal solution
  - Active contraints: Power balance and $(n-1)$ inequalities

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

s.t. $\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$

$$P_{G(g)}(\boldsymbol{\omega}) \leq P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) \geq P_{G(g)}^{min}, \quad \forall\, g = 1, \dots, N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \leq P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \geq -P_{L(l)}^{max}, \quad \forall\, l = 1, \dots, N_L$$

for $p^*(\omega)$ →

$$\sum_{i=1}^{N_B} \left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)} \right) = 0$$

$$\left.\begin{aligned} P_{G(g)}(\boldsymbol{\omega}) &= P_{G(g)}^{max}, \\ P_{G(g)}(\boldsymbol{\omega}) &= P_{G(g)}^{min}, \\ A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) &= P_{L(l)}^{max}, \\ A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) &= -P_{L(l)}^{max}, \end{aligned}\right\} (n-1) \text{ active constraints}$$

Recovering Feasible Solution from Active Set

- **DC Optimal Power Flow Problem**
  - Linear program with n variables → there is a *basic feasible* optimal solution
  - Active contraints: Power balance and $(n-1)$ inequalities

$$\min_{P_G(\boldsymbol{\omega})} \quad C_G^T P_G(\boldsymbol{\omega})$$

$$\text{s.t.} \quad \sum_{i=1}^{N_B}\left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)}\right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) \le P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) \ge P_{G(g)}^{min}, \quad \forall\, g = 1,\dots,N_G$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \le P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) \ge -P_{L(l)}^{max}, \quad \forall\, l = 1,\dots,N_L$$

for $p^*(\omega)$ →

$$\sum_{i=1}^{N_B}\left(P_{G(i)}(\boldsymbol{\omega}) + P_{W(i)} + \boldsymbol{\omega}_{(i)} - P_{D(i)}\right) = 0$$

$$P_{G(g)}(\boldsymbol{\omega}) = P_{G(g)}^{max}, \quad P_{G(g)}(\boldsymbol{\omega}) = P_{G(g)}^{min},$$

$$A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) = P_{L(l)}^{max}, \quad A_{(l,\cdot)}(P_G(\boldsymbol{\omega}) + P_W + \boldsymbol{\omega} - P_D) = -P_{L(l)}^{max},$$

$(n-1)$ active constraints

$$\boldsymbol{p}^*(\boldsymbol{\omega}) = A_{act}^{-1}(b_{act} + C_{act}\boldsymbol{\omega})$$

Solve set of linear equations!

Similar Results for Quadratic Programs

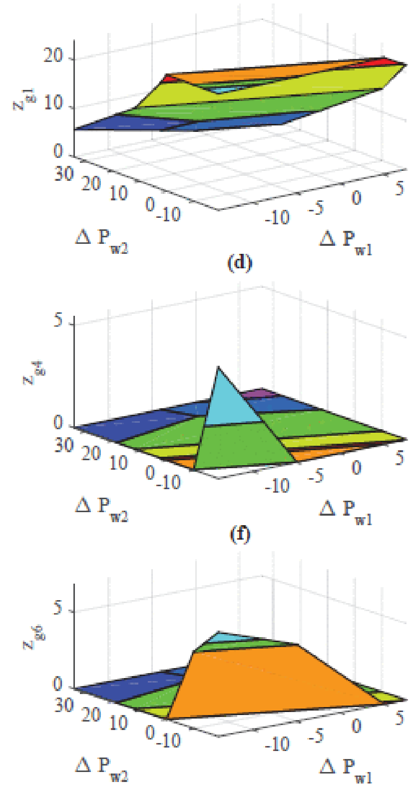
$$\min_{\boldsymbol{p}} \frac{1}{2}\boldsymbol{p}^T Q \boldsymbol{p} + (F\boldsymbol{\omega} + f)^T \boldsymbol{p}$$

$$s.t. \quad A\boldsymbol{p} \leq b + C\boldsymbol{\omega}$$

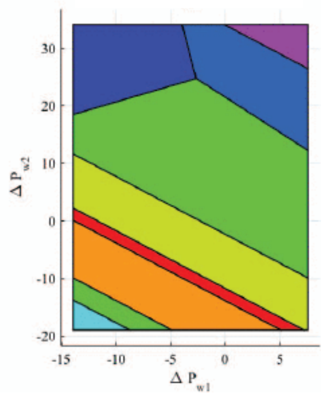
The optimal solution $\boldsymbol{p}^*(\boldsymbol{\omega})$ is a *piecewise affine* function of the uncertain parameters.

Each piece corresponds to a particular active set.

Piecewise affine functions



Regions in parameter space



M. Vrakopoulou and I. Hiskens, "Optimal policy-based control of generation and HVDC lines in power systems under uncertainty," IEEE PowerTech, Manchester, 2017

General Convex Programs

- **Active set encodes all information about the optimal solution**
  - «Minimal» information
  - Number of active contraints $|\mathcal{A}|$ is typically significantly smaller than the full number of constraints m

$$|\mathcal{A}| \ll m$$

- **Solve a relaxed optimization problem with only active constraints**

New approach:

Use Active Set

as a stepping stone

to Learn Optimal Solution!

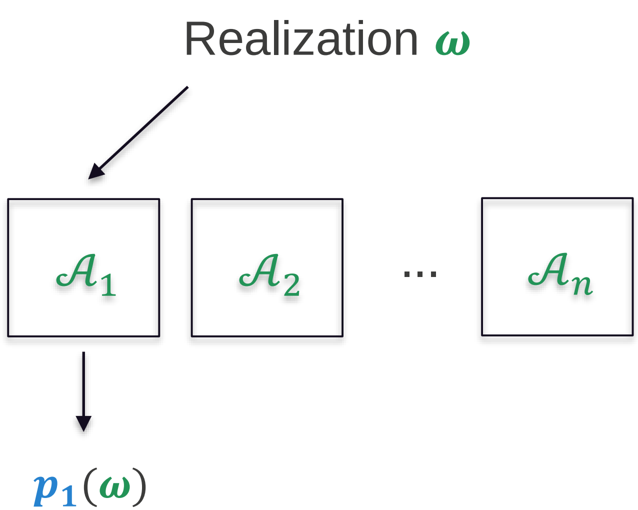
Ensemble Policy

Realization $\boldsymbol{\omega}$

$\mathcal{A}_1$	$\mathcal{A}_2$	...	$\mathcal{A}_n$

Candidates for optimal active set

Ensemble Policy

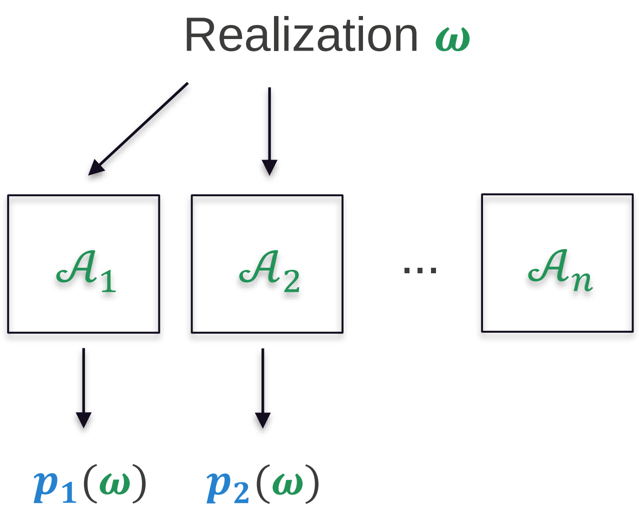


Candidates for
optimal active set

Obtain candidate solution $\tilde{\boldsymbol{p}}_{\boldsymbol{i}}(\boldsymbol{\omega})$ for each $\mathcal{A}_i$:

- LP: Solve set of linear equations
- General convex: Solve relaxed problem

Ensemble Policy

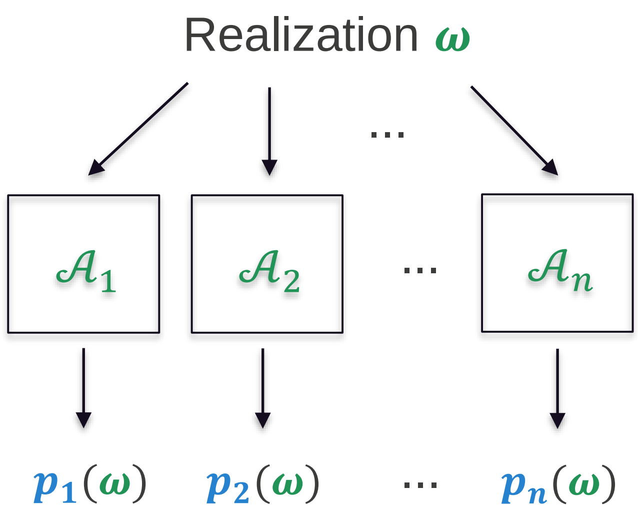


Candidates for
optimal active set

Obtain candidate solution $\tilde{\boldsymbol{p}}_{\boldsymbol{i}}(\boldsymbol{\omega})$ for each $\mathcal{A}_i$:

- LP: Solve set of linear equations
- General convex: Solve relaxed problem

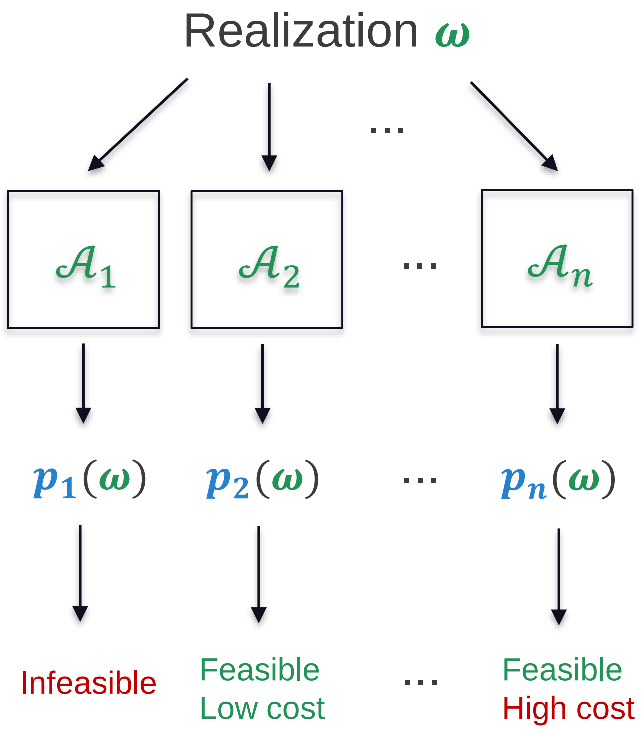
Ensemble Policy



Candidates for
optimal active set

Obtain candidate solution $\tilde{\boldsymbol{p}}_i(\boldsymbol{\omega})$ for each $\mathcal{A}_i$:
- LP: Solve set of linear equations
- General convex: Solve relaxed problem

Ensemble Policy



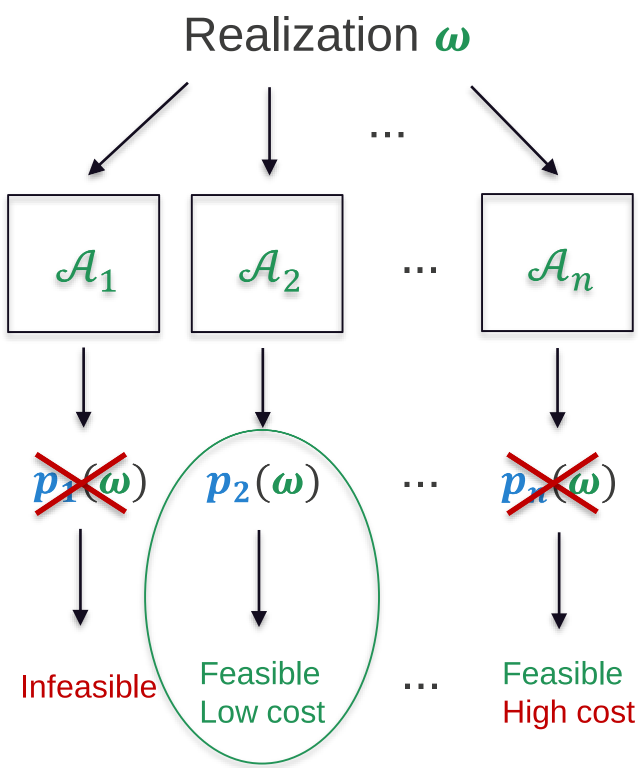
Candidates for optimal active set

Obtain candidate solution $\tilde{p}_i(\omega)$ for each $\mathcal{A}_i$:

- LP: Solve set of linear equations
- General convex: Solve relaxed problem

Evaluate cost and feasibility

Ensemble Policy



Candidates for optimal active set

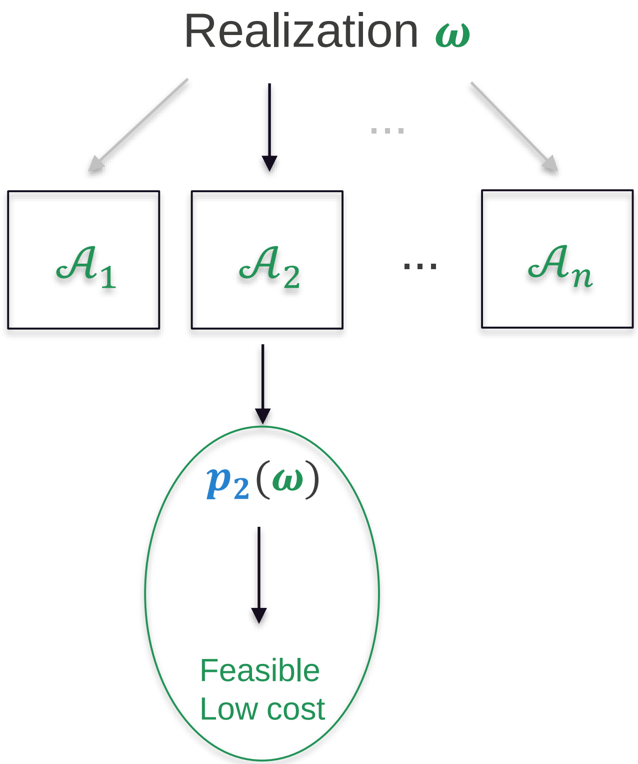
Obtain candidate solution $\tilde{p}_i(\omega)$ for each $\mathcal{A}_i$:

- LP: Solve set of linear equations
- General convex: Solve relaxed problem

Evaluate cost and feasibility

Pick **best** solution

Using Machine Learning to Simplify Further

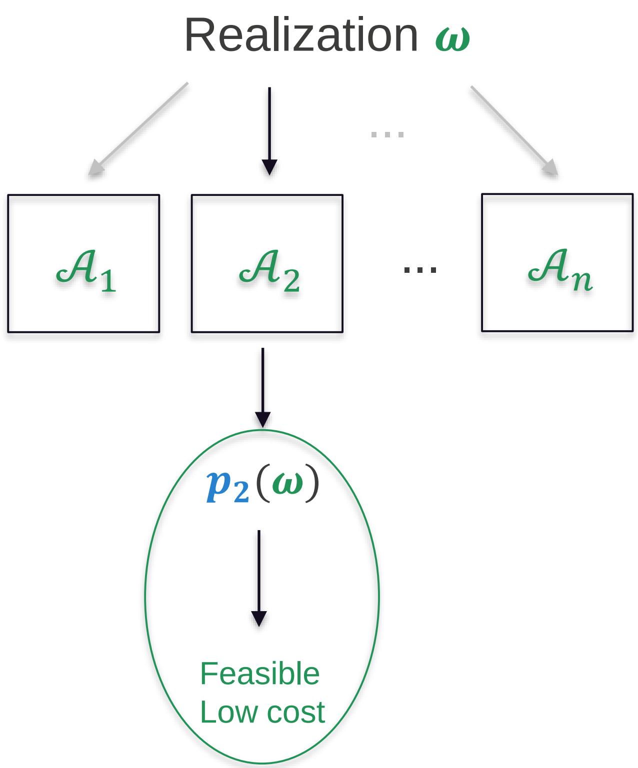


Classification:
Finite number of possible classes

Related work on learning regions with constant lagrangian multipliers

X.Geng, L. Xie, "Learning the LMP-Load Coupling From Data: A Support Vector Machine Based Approach", IEEE TPWRS, 2016

Using Machine Learning to Simplify Further



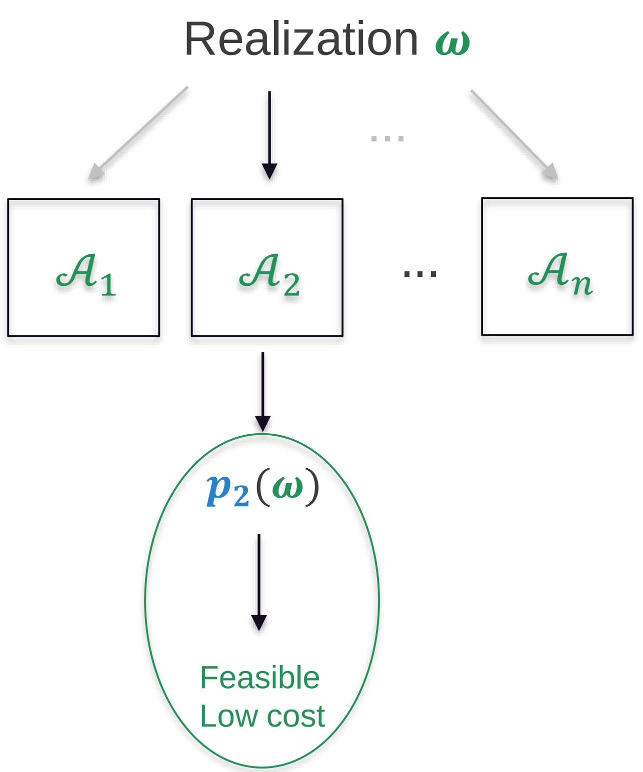
Classification:
Finite number of possible classes

Obtain solution only
for most relevant active set(s):
Develop good **local** approximation

Misra, Dvijotham, Molzahn, "Optimal Adaptive Linearizations of the AC Power Flow Equations", PSCC 2018

Hohmann, Warrington, Lygeros, «Optimal Linearizations of Power Systems with Uncertain Supply and Demand", May 2018

Limits of the Approach



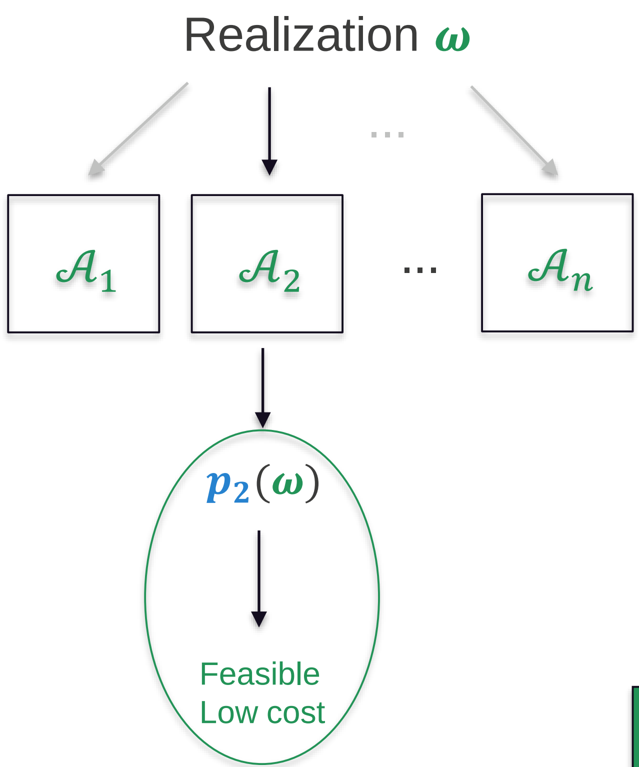
Works well if the number of optimal active sets is **small**!

Total number of possible active sets is exponential ☹

Many applications typically only have few optimal active sets! ☺

«Typical operation patterns»

Limits of the Approach



Works well if the number of optimal active sets is **small**!

Total number of possible active sets is exponential ☹

Many applications typically only have few optimal active sets! ☺

«Typical operation patterns»

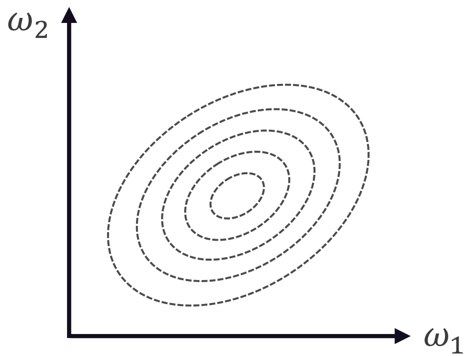
How do we know that we are dealing with a **nice** system?

How do we find the **important** active sets?

Using Sampling to Learn Important Active Sets

Important active sets

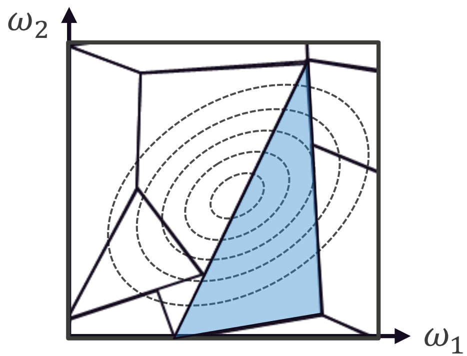
Probability distribution
over input parameters



---- level sets of probability distribution

Important active sets

Probability distribution
over input parameters



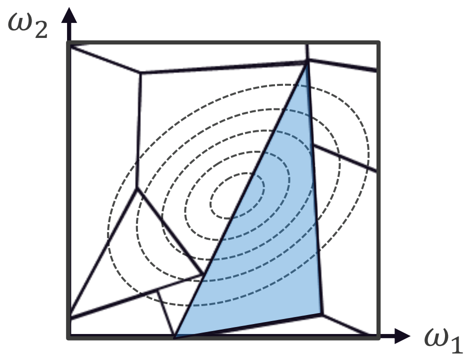
---- level sets of probability distribution

unknown
regions with
different
optimal
active sets

(in general not polyhedral)

Important active sets

Probability distribution
over input parameters



---- level sets of probability distribution

unknown regions with different optimal active sets

Importance of active set:

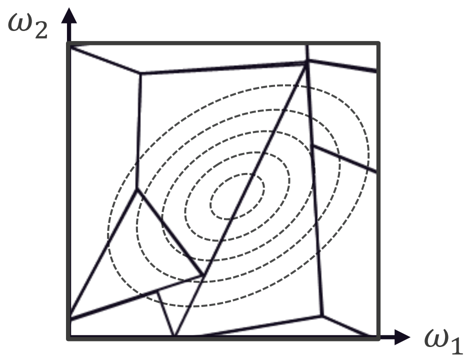
Probability that the active set is optimal

$$\mathbb{P}_\omega(\mathcal{A} = \mathcal{A}^*(\omega))$$

Main goal:

Discover a **collection of active sets** that together have high probabiliy of being optimal!

Finding Important Active Sets



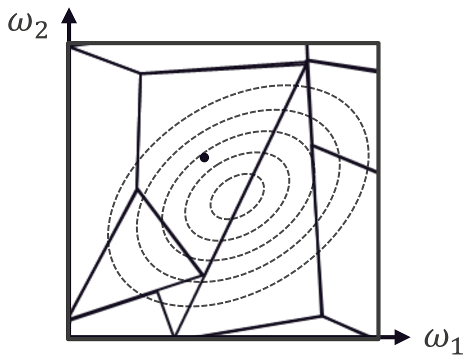
---- level sets of probability distribution

unobserved optimal active sets

discovered optimal active sets

samples of parameters

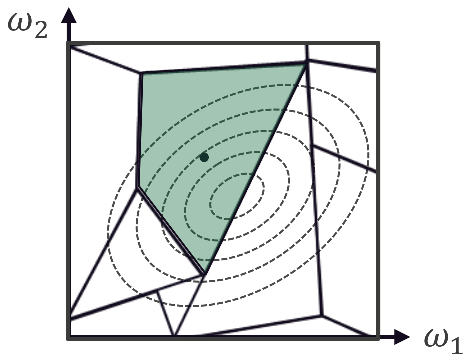
Finding Important Active Sets



---- level sets of probability distribution

- unobserved optimal active sets
- discovered optimal active sets
- samples of parameters

Finding Important Active Sets



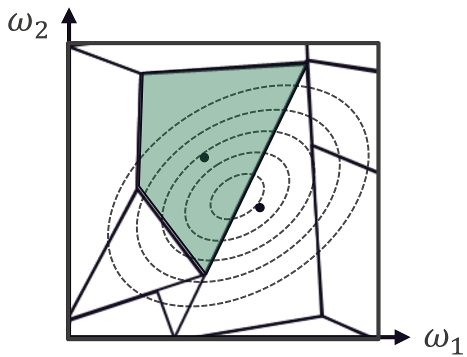
---- level sets of probability distribution

unobserved optimal active sets

discovered optimal active sets

samples of parameters

Finding Important Active Sets



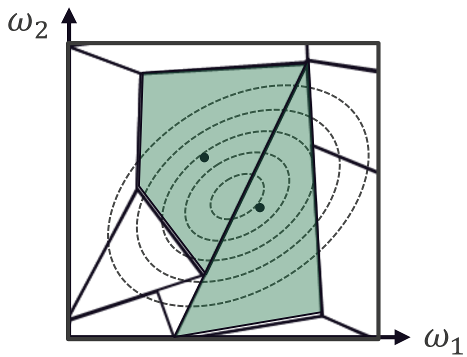
---- level sets of probability distribution

unobserved optimal active sets

discovered optimal active sets

samples of parameters

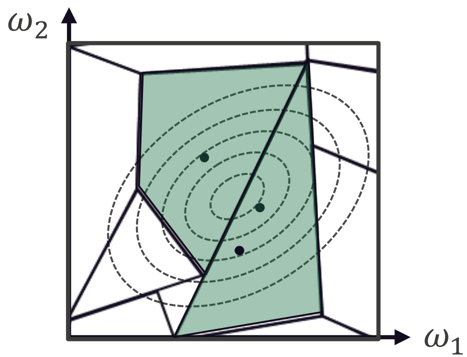
Finding Important Active Sets



---- level sets of probability distribution

- unobserved optimal active sets
- discovered optimal active sets
- samples of parameters

Finding Important Active Sets



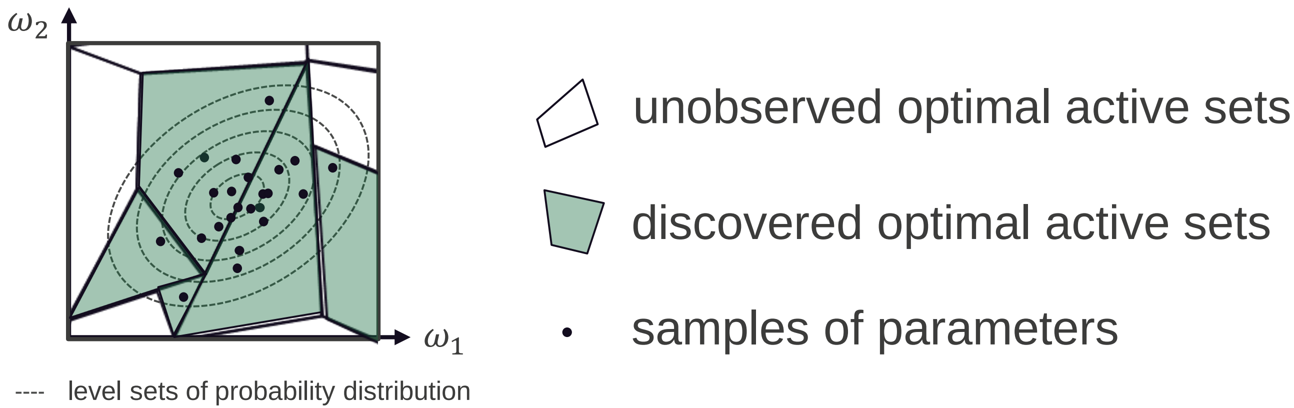
---- level sets of probability distribution

unobserved optimal active sets

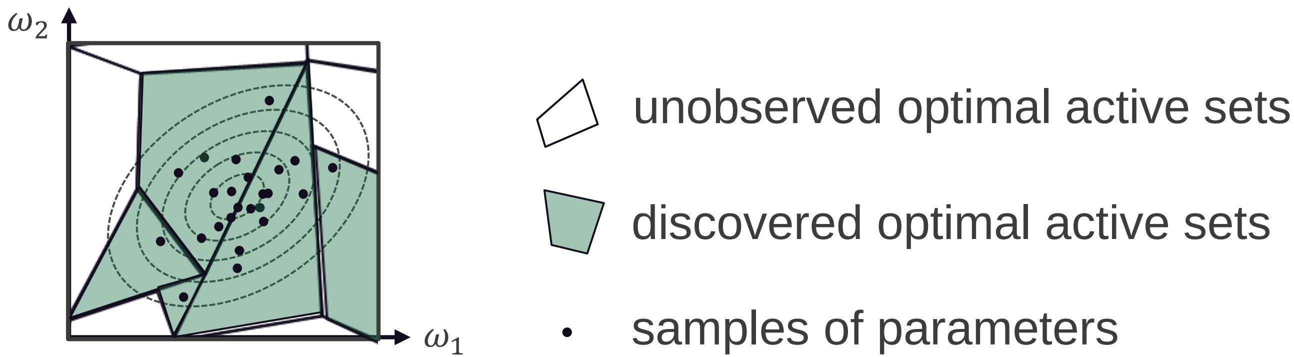
discovered optimal active sets

samples of parameters

Finding Important Active Sets



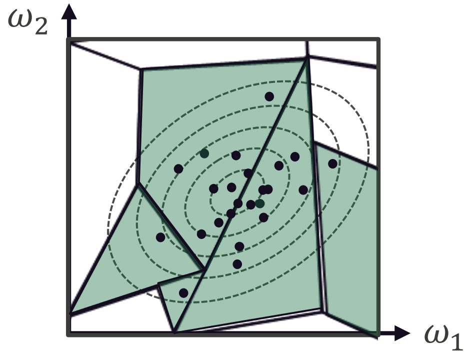
Finding Important Active Sets



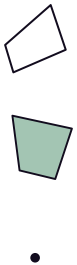
---- level sets of probability distribution

Draw samples until «sufficient» optimal active sets have been discovered

Finding Important Active Sets



---- level sets of probability distribution



unobserved optimal active sets

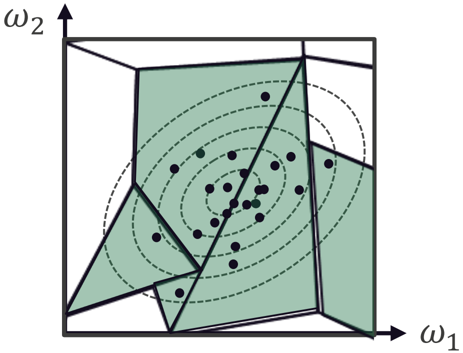
discovered optimal active sets

samples of parameters

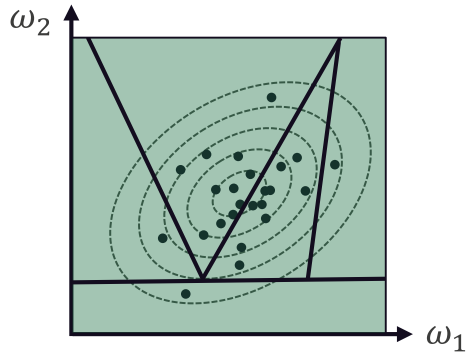
Draw samples until «sufficient» optimal active sets have been discovered

How do we know when it is safe to stop?

Finding Important Active Sets



---- level sets of probability distribution

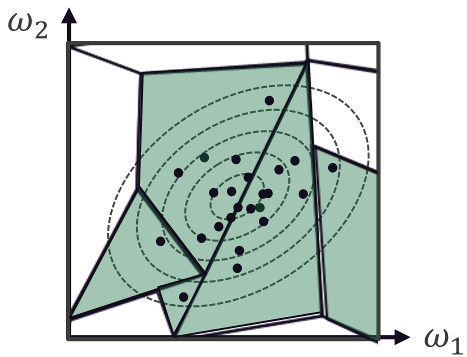


---- level sets of probability distribution

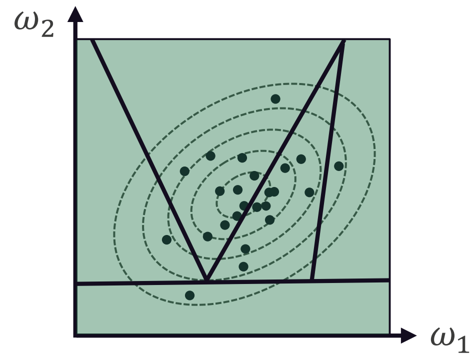
Draw samples until «sufficient» optimal active sets have been discovered

How do we know when it is safe to stop?

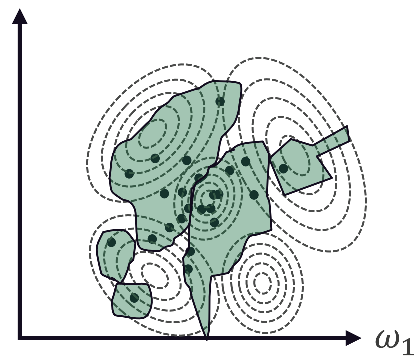
Finding Important Active Sets



---- level sets of probability distribution



---- level sets of probability distribution



---- level sets of probability distribution

Draw samples until «sufficient» optimal active sets have been discovered

How do we know when it is safe to stop?

Streaming Algorithm

with Probabilistic Learning Guarantees

Streaming Algorithm with Probabilistic Guarantees

Considering M i.i.d. samples, we obtain a set of unique active sets $\mathcal{A}_i$.

Streaming Algorithm with Probabilistic Guarantees

Considering M i.i.d. samples, we obtain a set of unique active sets $\mathcal{A}_i$.

For the next $W(M)$ i.i.d. samples, we calculate the «rate of discovery» $R_{M,W}$ of previously unobserved active sets: $R_{M,W} = \frac{N_{unobserved}}{N}$

Streaming Algorithm with Probabilistic Guarantees

Considering M i.i.d. samples, we obtain a set of unique active sets $\mathcal{A}_i$.

For the next $W(M)$ i.i.d. samples, we calculate the «rate of discovery» $R_{M,W}$ of previously unobserved active sets: $R_{M,W} = \frac{N_{unobserved}}{N}$

Theorem 1 and 2 [Misra, Roald, Ng, 2018]**:**

If the window size $W(M)$ is defined as

$$W = \frac{2\gamma}{\epsilon^2} \max\{\log M, \log \underline{M}\} \text{ with } \underline{M} = 1 + \left(\frac{\gamma}{\delta(\gamma-1)}\right)^{\frac{1}{\gamma-1}}$$

Then $\mathbb{P}\big(\pi(U_M) - R_{M,W} \leq \epsilon \;\; \forall\, M > 1\big) \leq 1 - \delta$

- ϵ difference between true and empirical probability of unobserved active sets
- δ confidence level
- γ hyperparameter > 1

Streaming Algorithm with Probabilistic Guarantees

Considering M i.i.d. samples, we obtain a set of unique active sets $\mathcal{A}_i$.

For the next $W(M)$ i.i.d. samples, we calculate the «rate of discovery» $R_{M,W}$ of previously unobserved active sets: $R_{M,W} = \frac{N_{unobserved}}{N}$

If rate of discovery is small enough $\left(R_{M,W} \leq \alpha - \epsilon\right)$: **Terminate**

α max. probability of unobserved active set

ϵ max. difference between true and empirical probability of unobserved active sets

Streaming Algorithm with Probabilistic Guarantees

Considering M i.i.d. samples, we obtain a set of unique active sets $\mathcal{A}_i$.

For the next $W(M)$ i.i.d. samples, we calculate the «rate of discovery» $R_{M,W}$ of previously unobserved active sets: $R_{M,W} = \frac{N_{unobserved}}{N}$

If rate of discovery is small enough $\left(R_{M,W} \leq \alpha - \epsilon\right)$: **Terminate**

If we have too many undiscovered active sets $\left(R_{M,W} > \alpha - \epsilon\right)$: $M = M + 1$

Streaming Algorithm with Probabilistic Guarantees

1. **Guaranteed to terminate,**
   no need to decide on the number of M samples apriori

2. **Guaranteed to terminate fast for *benign* systems**

If a (small) number of relevant active sets K_0 that contains more than $1-\alpha_0$ probability mass, then with probability at least $1-\delta-\delta_0$ the algorithm terminates in less iterations than

$$M = \frac{1}{\alpha - \alpha_0}\left(K_0 \log 2 + \log \frac{1}{\delta_0}\right)$$

Results for the

DC Optimal Power Flow Problem

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇨ Termination: $R_{M,W} \leq 0.01$

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇒ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇨ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇒ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Terminates fast.

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇒ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Terminates fast.

Some outliers:

Large number of active sets

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ → Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Terminates fast.

Some outliers:

Large number of active sets

Rate of discovery ≥ 0.01

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ → Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Terminates fast.

Some outliers:

Large number of active sets

Rate of discovery ≥ 0.01

Degeneracy?

Many small generators...

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇨ Termination: $R_{M,W} \leq 0.01$

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ → Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd										
case5_pjm										
case14_ieee										
case30_ieee										
case39_epri										
case118_ieee										
case57_ieee										
case1888_rte										
case1951_rte										
case162_ieee_dtc										
case24_ieee_rts										

System sizes ranging from 3 to 1951 nodes

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ → Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd										
case5_pjm										
case14_ieee										
case30_ieee										
case39_epri										
case118_ieee										
case57_ieee										
case1888_rte										
case1951_rte										
case162_ieee_dtc										
case24_ieee_rts										

Normal distribution
$\omega \sim \mathcal{N}(0, \sigma = 0.03d)$

Uniform distribution
with support
$\omega \in [-0.09d, 0.09d]$

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇨ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇨ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941

Few active sets!

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇨ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941

Few active sets!

Terminates fast.

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ➡ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Terminates fast.

Some outliers:

Large number of active sets

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ → Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Terminates fast.

Some outliers:

Large number of active sets

Rate of discovery ≥ 0.01

Streaming Algorithm Results – PGLib-OPF v 17.08

Max. undiscovered: $\alpha = 0.05$, Max. difference: $\epsilon = 0.04$ ⇨ Termination: $R_{M,W} \leq 0.01$

	Normal distribution					Uniform distribution				
	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	K_M	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$
Low-Complexity										
case3_lmbd	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case5_pjm	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case14_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case30_ieee	1	1	13'259	0.0	1.0	1	1	13'259	0.0	0.0
case39_epri	2	2	13'259	0.0	1.0	2	2	13'259	0.0008	0.9998
case118_ieee	2	33	13'259	0.0	1.0	2	4	13'259	0.0019	0.9984
case57_ieee	2	2	13'259	0.0003	0.9997	3	46	13'259	0.0	0.0
case1888_rte	3	6	13'259	0.0	0.0	3	10	13'259	0.0	0.0
case1951_rte	5	47	13'259	0.0069	0.9943	11	63	13'259	0.0084	0.9901
case162_ieee_dtc	7	91	13'259	0.0054	0.9925	17	192	13'259	0.0085	0.9926
case24_ieee_rts	10	1456	18'209	0.0	0.0	11	64	13'259	0.0047	0.9941
High-Complexity										
case73_ieee_rts	19	1258	17'844	0.0087	0.9931	130	22'000	24'977	0.0136	-
case300_ieee	24	1257	17'842	0.0073	0.9919	293	9095	22'789	0.0099	0.9897
case200_pserc	174	4649	21'112	0.0099	0.9909	236	6741	22'040	0.0099	0.9901
case240_pserc	2993	22'000	24'997	0.0795	-	2993	22'000	24'997	0.0795	-

Few active sets!

Terminates fast.

Some outliers:

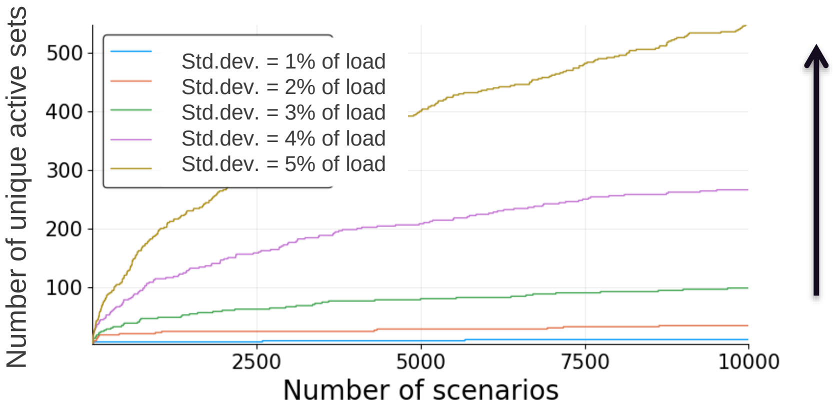
Large number of active sets

Rate of discovery ≥ 0.01

Degeneracy?

Many small generators...

IEEE 300 bus system with normally distributed load



Increasing parameter uncertainty =
Increasing number of important active sets

«Power systems operation becomes more unpreditable and complex with increasing uncertainty»

General perception among system operators

Summary

- **Active sets are useful for learning solutions to optimization problems**
  - Active sets encode the problem structure, including relevant constraints
  - Solution efficiently recoverable from simpler optimization problem
  - Finite number of relevant active sets («typically» few)

- **Three-step machine learning procedure:**

1. Learn active sets (offline learning) – establish learnability of the task!
2. Find the right one (classification)
3. Recover optimal solution (simpler problem/approximation)

} **FUTURE WORK** (steps 2–3)

THANK YOU!

Line Roald, roald@lanl.gov

Yee Sian Ng, Sidhant Misra, Line Roald and Scott Backhaus, «Statistical Learning for DC Optimal Power Flow», accepted to Power System Computation Conference (PSCC), 2018

Sidhant Misra, Line Roald and Yee Sian Ng, «Learning for Convex Optimization», submitted to NIPS 2018