Learning Congestion Patterns in Optimal Power Flow Problems

Line Roald (University of Wisconsin - Madison)

IEEE BDA Tutorial Series:
Big Data & Analytics for Power Systems

Feb 12, 2020

Learning Congestion Patterns in Optimal Power Flow Problems

Line Roald (University of Wisconsin - Madison)

with Sidhant Misra (LANL), Yeesian Ng (MIT), Sean Simpson (UW Madison), and Dan Molzahn (Georgia Tech)

Power System Patterns

https://electrical-engineering-portal.com/wp-content/uploads/2018/02/power-transmission-control-center.jpg

Power System Patterns

https://electrical-engineering-portal.com/wp-content/uploads/2018/02/power-transmission-control-center.jpg

$$\min \quad \sum_{i \in \mathcal{G}} \left(c_{2,i} p_{G,i}^2 + c_{1,i}\, p_{G,i} \right)$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j \in \delta i} b_{ij} (\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \leq p_{G,g} \leq p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \leq b_{ij} \left(\theta_i - \theta_j \right) \leq f_{ij}^{max}, j \in \mathcal{L}$$

"Power system operators solve power flow in their heads"

Informal discussion with Thomas Trøtscher (Head of Data Science, Statnett, Norway)

Power System Patterns

https://electrical-engineering-portal.com/wp-content/uploads/2018/02/power-transmission-control-center.jpg

$$\min \quad \sum_{i \in \mathcal{G}} \left(c_{2,i} p_{G,i}^2 + c_{1,i}\, p_{G,i} \right)$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j \in \delta i} b_{ij} (\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

IDEA: Can we use *machine learning* to identify the optimal dispatch?

Optimal Power Flow

Goal: Low cost operation, while enforcing technical limits

$$\min \quad \sum_{i \in \mathcal{G}} \left(c_{2,i} p_{G,i}^2 + c_{1,i}\, p_{G,i} \right)$$ minimize generation cost

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j \in \delta i} b_{ij} (\theta_i - \theta_j)\,,$$ power flow balance

$$p_{G,g}^{min} \leq p_{G,g} \leq p_{G,g}^{max},\ \ g \in \mathcal{G}$$ generation constraints

$$-f_{ij}^{max} \leq b_{ij} \left(\theta_i - \theta_j \right) \leq f_{ij}^{max}, j \in \mathcal{L}$$ transmission constraints

Optimal Power Flow

Goal: Low cost operation, while enforcing technical limits

$$\min \quad \sum_{i \in \mathcal{G}} \left(c_{2,i} p_{G,i}^2 + c_{1,i}\, p_{G,i} \right)$$ minimize generation cost

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j \in \delta i} b_{ij} (\theta_i - \theta_j)\,,$$ power flow balance

$$p_{G,g}^{min} \leq p_{G,g} \leq p_{G,g}^{max}, \;\; g \in \mathcal{G}$$ generation constraints

$$-f_{ij}^{max} \leq b_{ij} \left(\theta_i - \theta_j \right) \leq f_{ij}^{max}, j \in \mathcal{L}$$ transmission constraints

Could include much more details (security constraints, generator variables on/off variables, non-linear AC power flow, …)

Optimal Power Flow

System conditions (load, renewables) are constantly changing

$$\min \quad \sum_{i \in \mathcal{G}} \left(c_{2,i} p_{G,i}^2 + c_{1,i}\, p_{G,i} \right)$$ minimize generation cost

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j \in \delta i} b_{ij} (\theta_i - \theta_j)\,,$$ power flow balance

$$p_{G,g}^{min} \leq p_{G,g} \leq p_{G,g}^{max},\ \ g \in \mathcal{G}$$ generation constraints

$$-f_{ij}^{max} \leq b_{ij} \left(\theta_i - \theta_j \right) \leq f_{ij}^{max}, j \in \mathcal{L}$$ transmission constraints

Could include much more details (security constraints, generator variables on/off variables, non-linear AC power flow, …)

Optimal Power Flow

System conditions (load, renewables) are constantly changing

$$\min \quad \sum_{i \in \mathcal{G}} \left(c_{2,i} p_{G,i}^2 + c_{1,i}\, p_{G,i} \right)$$ minimize generation cost

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j \in \delta i} b_{ij} (\theta_i - \theta_j)\,,$$ power flow balance

$$p_{G,g}^{min} \leq p_{G,g} \leq p_{G,g}^{max}, \;\; g \in \mathcal{G}$$ generation constraints

$$-f_{ij}^{max} \leq b_{ij} \left(\theta_i - \theta_j \right) \leq f_{ij}^{max}, j \in \mathcal{L}$$ transmission constraints

Resolve this problem **every 5-15 min** for varying load profile $\boldsymbol{p_D}$

Could include much more details (security constraints, generator variables on/off variables, non-linear AC power flow, …)

Repeated solution process

OPF at T_1 with load $\boldsymbol{p_D^{(1)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}\left(\theta_i - \theta_j\right) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(1)}\right)$

Repeated solution process

OPF at T_1 with load $\boldsymbol{p_D^{(1)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}\left(\theta_i - \theta_j\right) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(1)}\right)$

OPF at T_2 with load $\boldsymbol{p_D^{(2)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}\left(\theta_i - \theta_j\right) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(2)}\right)$

Repeated solution process

OPF at T_1 with load $\boldsymbol{p_D^{(1)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(1)}\right)$

OPF at T_2 with load $\boldsymbol{p_D^{(2)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(2)}\right)$

OPF at T_3 with load $\boldsymbol{p_D^{(3)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(3)}\right)$

Repeated solution process

OPF at T_1 with load $\boldsymbol{p_D^{(1)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(1)}\right)$

OPF at T_2 with load $\boldsymbol{p_D^{(2)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(2)}\right)$

OPF at T_3 with load $\boldsymbol{p_D^{(3)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ \ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(3)}\right)$

……

Repeated solution process

OPF at T_1 with load $\boldsymbol{p_D^{(1)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(1)}\right)$

OPF at T_2 with load $\boldsymbol{p_D^{(2)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(2)}\right)$

OPF at T_3 with load $\boldsymbol{p_D^{(3)}}$

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\,p_{G,i}\right)$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Optimal solution $p_G^*\left(p_D^{(3)}\right)$

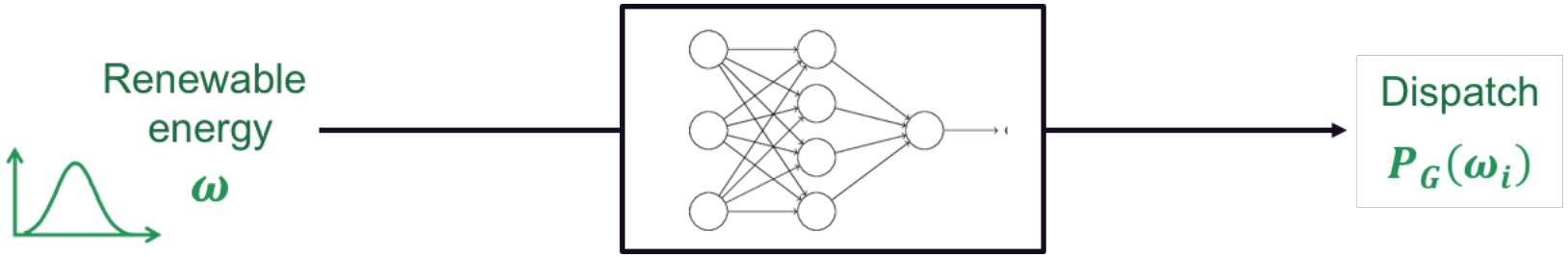
……

Can we use solutions for previous load profiles $\boldsymbol{p_D}$ to learn the new solution $\boldsymbol{p_G^*}$?

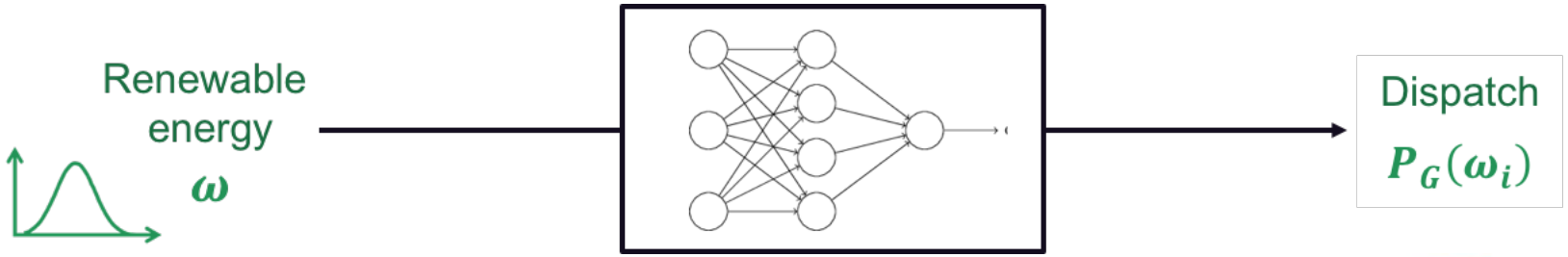
First attempt:

Train a neural net!

First attempt: Train a neural net



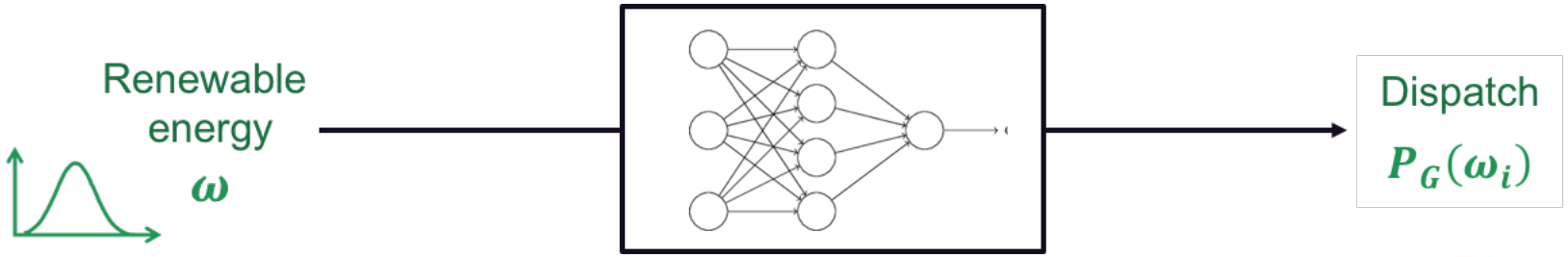
First attempt: Train a neural net



- This didn’t work well...

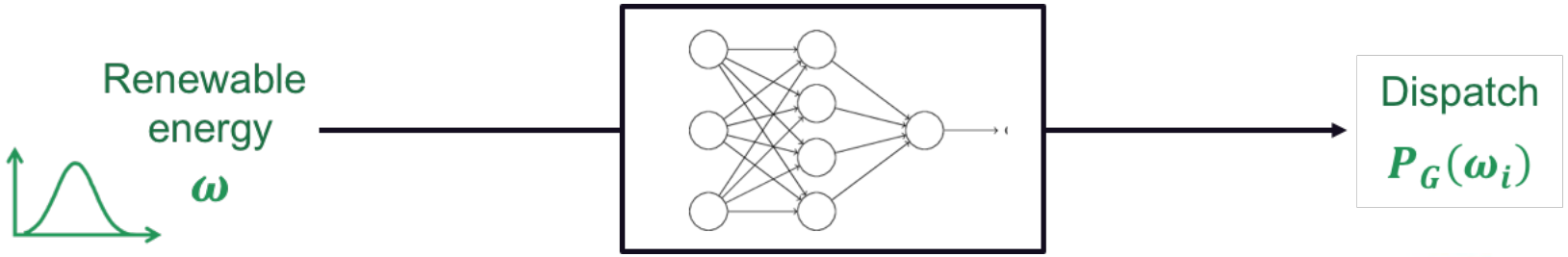
(though I will admit that we gave up quite fast)

First attempt: Train a neural net



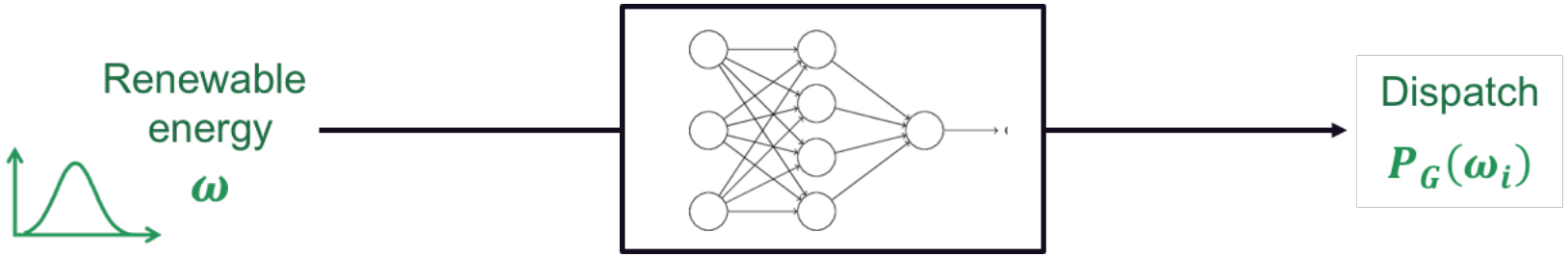
- This didn't work well...
  - Hard to satisfy safety constraints!

First attempt: Train a neural net



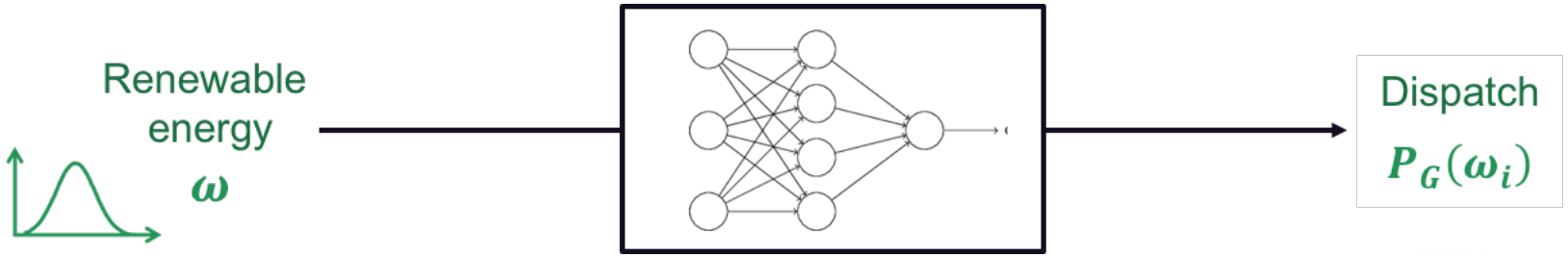
- This didn't work well...
  - Hard to satisfy safety constraints!
  - Projection back onto feasible space cause suboptimality…

First attempt: Train a neural net



- This didn't work well...
  - Hard to satisfy safety constraints!
  - Projection back onto feasible space cause suboptimality…
  - Challenging: High-dimensional input → High dimensional output

First attempt: Train a neural net



- This didn't work well...
  - Hard to satisfy safety constraints!
  - Projection back onto feasible space cause suboptimality…
  - Challenging: High-dimensional input → High dimensional output

- This can work well under some circumstances

Wide enough and deep enough
[Karg and Lucia, 2018]

Power System Patterns

https://electrical-engineering-portal.com/wp-content/uploads/2018/02/power-transmission-control-center.jpg

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

"Power system operators solve power flow in their heads"

Informal discussion with Thomas Trøtscher (Head of Data Science, Statnett, Norway)

Power System Patterns

How is this possible?

Limited number of ***operational patterns***

- Set of congested lines
- Set of generators at their limits

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

"Power system operators solve power flow in their heads"

Informal discussion with Thomas Trøtscher
(Head of Data Science, Statnett, Norway)

Power System Patterns

How is this possible?

Limited number of ***operational patterns***

- Set of congested lines
- Set of generators at their limits

This corresponds to the active constraints of an OPF problem!

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

"Power system operators solve power flow in their heads"

Informal discussion with Thomas Trøtscher
(Head of Data Science, Statnett, Norway)

Power System Patterns

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

How is this possible?

Limited number of ***operational patterns***

- Set of congested lines
- Set of generators at their limits

This corresponds to the active constraints of an OPF problem!

Operational pattern = optimal active set $\mathcal{A}^*$

"Power system operators solve power flow in their heads"

Informal discussion with Thomas Trøtscher (Head of Data Science, Statnett, Norway)

Power System Patterns

How is this possible?

Limited number of ***operational patterns***

- Set of congested lines
- Set of generators at their limits

This corresponds to the active constraints of an OPF problem!

Operational pattern = optimal active set $\mathcal{A}^*$

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Shorthand version of the OPF:

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\, p_{G,i}\right)$$

$$\text{s.t.} \quad Ap_G \le b + C\boldsymbol{p_D}$$

Power System Patterns

How is this possible?

Limited number of ***operational patterns***

- Set of congested lines
- Set of generators at their limits

This corresponds to the active constraints of an OPF problem!

Operational pattern = optimal active set $\mathcal{A}^*$

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Shorthand version of the OPF:

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\, p_{G,i}\right)$$

$$\text{s.t.} \quad Ap_G \le b + C\boldsymbol{p_D}$$

Active constaints are satisfied with equality $A_{act}p_G = b_{act} + C_{act}p_D$

Power System Patterns

How is this possible?

Limited number of ***operational patterns***

- Set of congested lines
- Set of generators at their limits

Each optimal operational pattern $\mathcal{A}^*$ has a corresponding ***optimal control policy***

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

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Shorthand version of the OPF:

$$\min \quad \sum_{i\in\mathcal{G}}\left(c_{2,i}p_{G,i}^2 + c_{1,i}\, p_{G,i}\right)$$

$$\text{s.t.} \quad Ap_G \le b + C\boldsymbol{p}_D$$

Active constaints are satisfied with equality $A_{act}p_G = b_{act} + C_{act}p_D$

Power System Patterns

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

How is this possible?

Limited number of ***operational patterns***

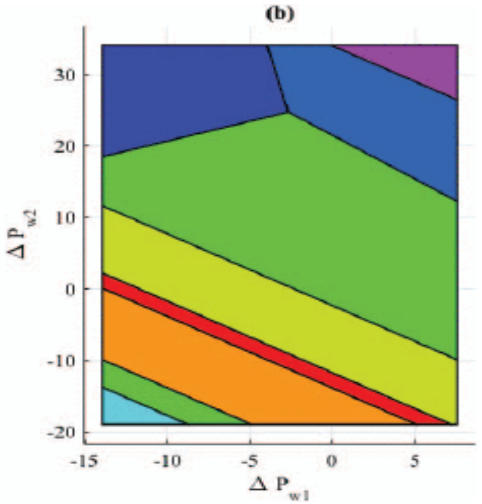
- Set of congested lines
- Set of generators at their limits

Each optimal operational pattern $\mathcal{A}^*$ has a corresponding ***optimal control policy***

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

Optimal dispatch is a ***linear function*** of the load profile $\boldsymbol{p_D}$ for *small** changes!

*Small changes = same active constraints at optimal solution



[Vrakopoulou and Hiskens, 2017]

Power System Patterns

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

How is this possible?

Limited number of ***operational patterns***

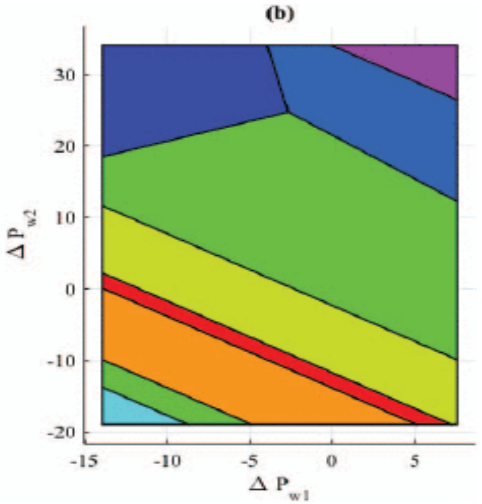
- Set of congested lines
- Set of generators at their limits

Each optimal operational pattern $\mathcal{A}^*$ has a corresponding ***optimal control policy***

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

Multiparametric programming, explicit MPC…

[Ji, Thomas and Tong '16], [Geng and Xie '17]



[Vrakopoulou and Hiskens, 2017]

Power System Patterns

Multi-parametric programming is inconvenient and computationally expensive for large systems ☹

The number of possible active sets is exponential in problem size ☹

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

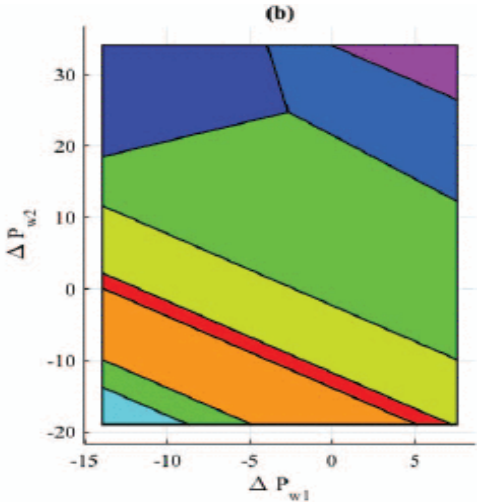
$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Multiparametric programming, explicit MPC…

[Ji, Thomas and Tong ‘16], [Geng and Xie ‘17]



[Vrakopoulou and Hiskens, 2017]

Power System Patterns

Multi-parametric programming is inconvenient and computationally expensive for large systems ☹

The number of possible active sets is exponential in problem size ☹

… but maybe only a few matters in practice? ☺

Can we learn the practically relevant active sets?

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

Think again!

https://electrical-engineering-portal.com/wp-content/uploads/2018/02/power-transmission-control-center.jpg

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - p_{D(i)} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},\ g \in \mathcal{G}$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max}, j \in \mathcal{L}$$

IDEA: Can we use *statistical learning* to identify optimal active sets and corresponding optimal policies?

Learning the Patterns

Learning optimal active sets through sampling

load profile 1 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 1**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}\,, ij \in \mathcal{L}$$

Learning the Patterns

Learning optimal active sets through sampling

load profile 1 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 1**

load profile 2 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 2**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}\,, ij \in \mathcal{L}$$

Learning the Patterns

Learning optimal active sets through sampling

load profile 1 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 1**

load profile 2 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 2**

load profile 3 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 3**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}\,, ij \in \mathcal{L}$$

Learning the Patterns

Learning optimal active sets through sampling

load profile 1 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 1**

load profile 2 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 2**

load profile 3 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set 3**

…

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}\,, ij \in \mathcal{L}$$

Learning the Patterns

Learning optimal active sets through sampling

load profile 1 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**, $\mathcal{A}^*$ **Optimal active set 1**

load profile 2 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**, $\mathcal{A}^*$ **Optimal active set 2**

load profile 3 $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**, $\mathcal{A}^*$ **Optimal active set 3**

…

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

Benefit:
We only look for active sets that occur in practice!

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}\,, ij \in \mathcal{L}$$

When to stop? Combined **probability** of observed patterns $\mathcal{A}^* = \{\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots\}$ is high.

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

When to stop? Combined **probability** of observed patterns $\mathcal{A}^* = \{\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots\}$ is high.

Combined probability of observed patterns $\mathcal{A}^*$ = the probability that $\mathcal{A}^*$ contains the optimal policy and give us an optimal solution for a new sample p_D

Load $\boldsymbol{p_D}$

$\mathcal{A}^{(1)} \rightarrow \boldsymbol{p_G^*}$

$\mathcal{A}^{(2)} \rightarrow p_G^{(2)}$

$\dots \quad \dots$

$\mathcal{A}^{(n)} \rightarrow p_G^{(n)}$

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ ⟶ **Optimal Power Flow** ⟶ p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ ⟶ **Optimal Power Flow** ⟶ p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

Algorithm based on a "stream" of samples: [Misra, Roald and Ng '18]

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

Algorithm based on a “stream” of samples: [Misra, Roald and Ng ‘18]

1. Use first samples to identify optimal active sets $\mathcal{A}^*$

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

Algorithm based on a "stream" of samples: [Misra, Roald and Ng '18]

1. Use first samples to identify optimal active sets $\mathcal{A}^*$
2. Use next samples to assess the "rate of discovery" (i.e., empirical probability of observing previously unseen $\mathcal{A}^*$)

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

Algorithm based on a "stream" of samples: [Misra, Roald and Ng '18]

1. Use first samples to identify optimal active sets $\mathcal{A}^*$
2. Use next samples to assess the "rate of discovery" (i.e., empirical probability of observing previously unseen $\mathcal{A}^*$)
3. Use empirically observed probability to derive bounds on the true probability

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

Algorithm based on a "stream" of samples: [Misra, Roald and Ng '18]

1. Use first samples to identify optimal active sets $\mathcal{A}^*$
2. Use next samples to assess the "rate of discovery" (i.e., empirical probability of observing previously unseen $\mathcal{A}^*$)
3. Use empirically observed probability to derive bounds on the true probability
4. If true probability is too high: Add additional samples and move to step 1

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ → **Optimal Power Flow** → p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

Algorithm based on a “stream” of samples: [Misra, Roald and Ng ‘18]

- Guaranteed to converge
- Converges fast if there are few active sets
- No assumptions on problem structure/probability distribution required!

Learning the Patterns

Learning optimal active sets through sampling

load profile $\boldsymbol{i}$ $\boldsymbol{p_D}$ ⟶ **Optimal Power Flow** ⟶ p_G^* **Optimal dispatch**

$\mathcal{A}^*$ **Optimal active set $\boldsymbol{i}$**

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j)\,,$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

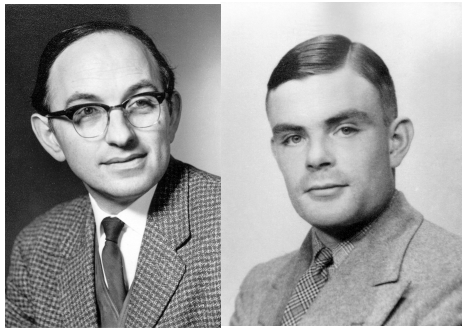
When to stop? Combined **probability** of observed patterns $\mathcal{A}^*$ is high.

Requires estimating this probability!

Algorithm based on a “stream” of samples: [Misra, Roald and Ng ‘18]

Missing mass problem and Good-Turing estimator!

[Bertsimas and Stellato ‘18]



Using the Patterns

Operating the system based on learned policies

Load $\boldsymbol{p_D}$

$\mathcal{A}^{(1)} \rightarrow \boldsymbol{p_G^*}$

$\mathcal{A}^{(2)} \rightarrow \boldsymbol{p_G^{(2)}}$

$\ldots \quad \ldots$

$\mathcal{A}^{(n)} \rightarrow \boldsymbol{p_G^{(n)}}$

$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j),$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

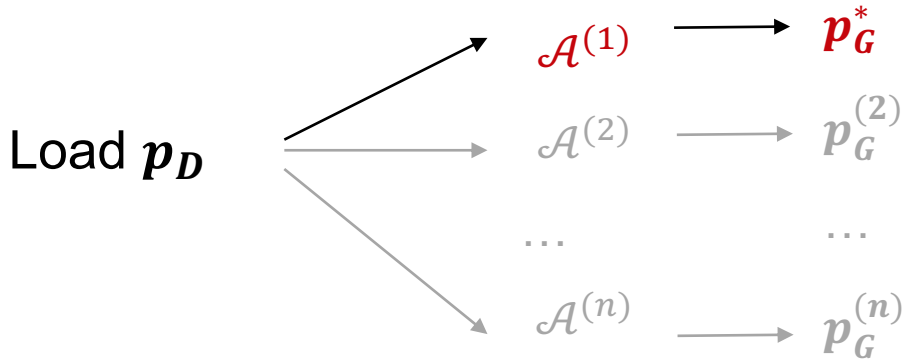
$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

Operator experience!

Classification [Deka and Misra, 2019]

Using the Patterns

Operating the system based on learned policies



$$\min_{p_G,\theta} \quad \sum_{i\in\mathcal{G}} c_{1,i}\, p_{G,i}$$

s.t.

$$p_{G(i)} - \boldsymbol{p_{D(i)}} = \sum_{j\in\delta i} b_{ij}(\theta_i - \theta_j),$$

$$p_{G,g}^{min} \le p_{G,g} \le p_{G,g}^{max},$$

$$-f_{ij}^{max} \le b_{ij}(\theta_i - \theta_j) \le f_{ij}^{max},$$

$$i \in N, g \in \mathcal{G}, ij \in \mathcal{L}$$

Operator experience / classification [Deka and Misra, 2019]

- Works well if the number of optimal active sets is small
- Dangerous if a correct active set has not yet been observed

Is it effective?

Stopping criterion: Rate of discovery (Empirical probability) < 0.01

- Works well if the number of optimal active sets is small
- Dangerous if a correct active set has not yet been observed

Is it effective?

Stopping criterion: Rate of discovery (Empirical probability) < 0.01

Test case

	Normal distribution					Uniform distribution				
	K_M	M	W_M	R_W	$\mathbb{P}(p^*)$	K_M	M	W_M	R_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	-

For many DC OPF test cases* it works well:

Few active sets required to achieve a high probability

- Works well if the number of optimal active sets is small
- Dangerous if a correct active set has not yet been observed

*from pglib_opf

Is it effective?

of active sets

# of active sets

Test case

	Normal distribution					Uniform distribution				
	$\mathbb{M}$	M	W_M	$R_{M,W}$	$\mathbb{P}(p^*)$	$\mathbb{M}$	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	-

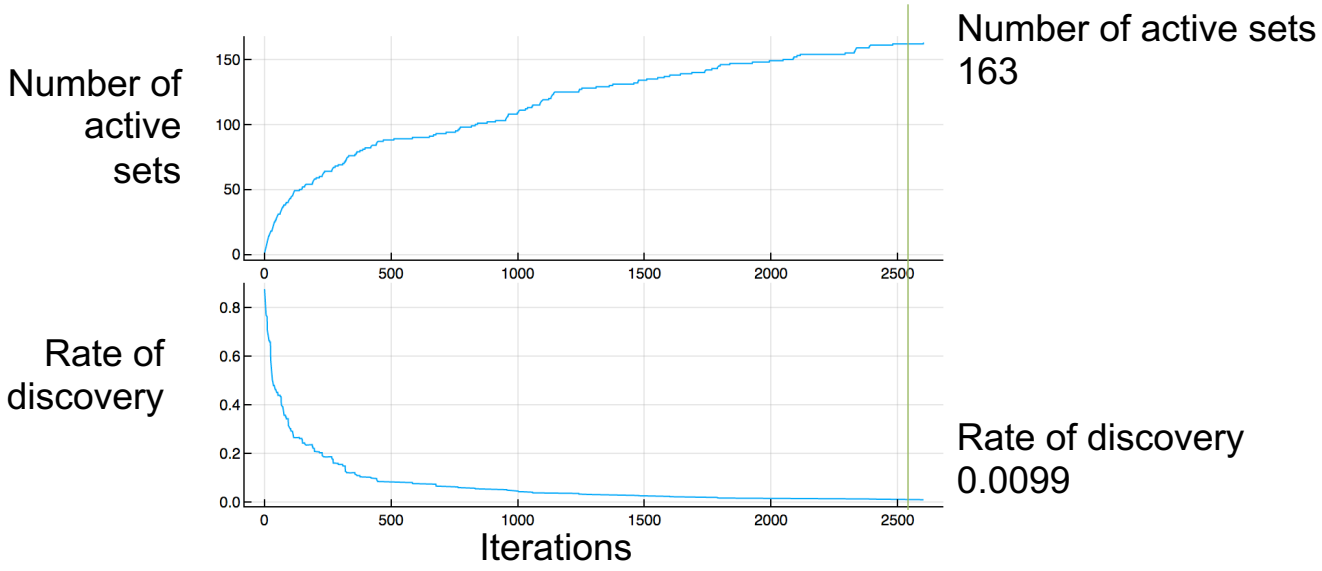
For many DC OPF test cases*
it works well:

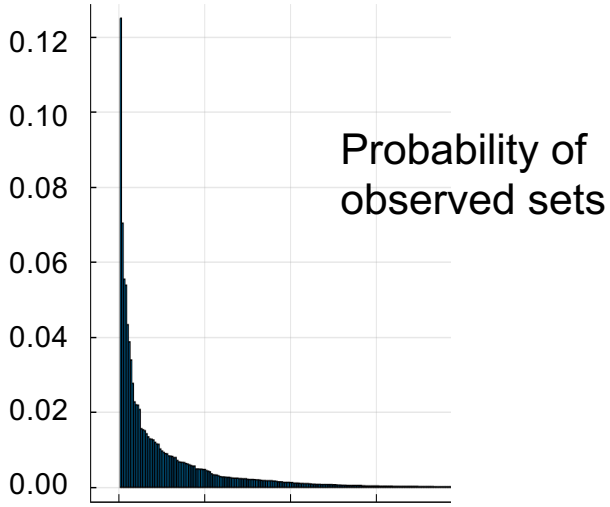
Few active sets required to achieve a high probability

- Works well if the number of optimal active sets is small
- Dangerous if a correct active set has not yet been observed

*from pglib_opf

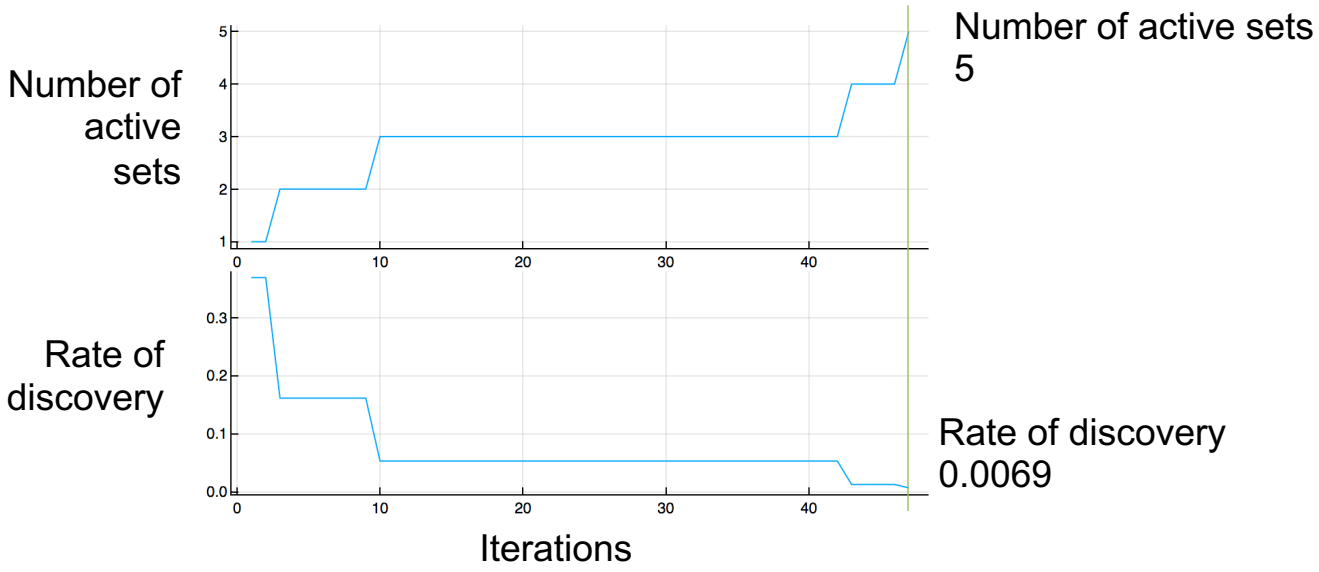
PSERC 200 bus test case

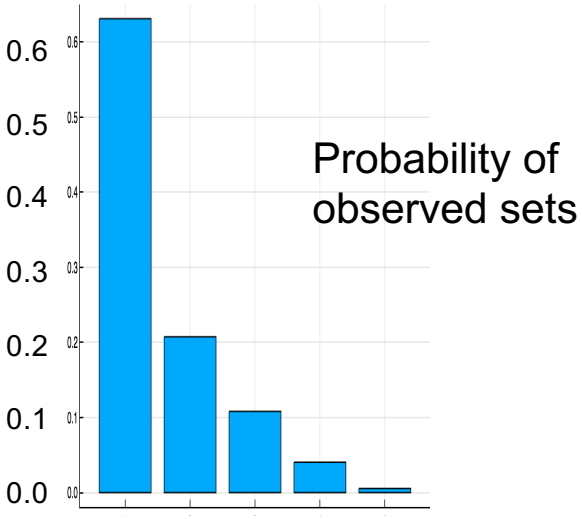




Many optimal active sets with non-negligible probability = the algorithm terminates slowly!

RTE 1951 bus test case

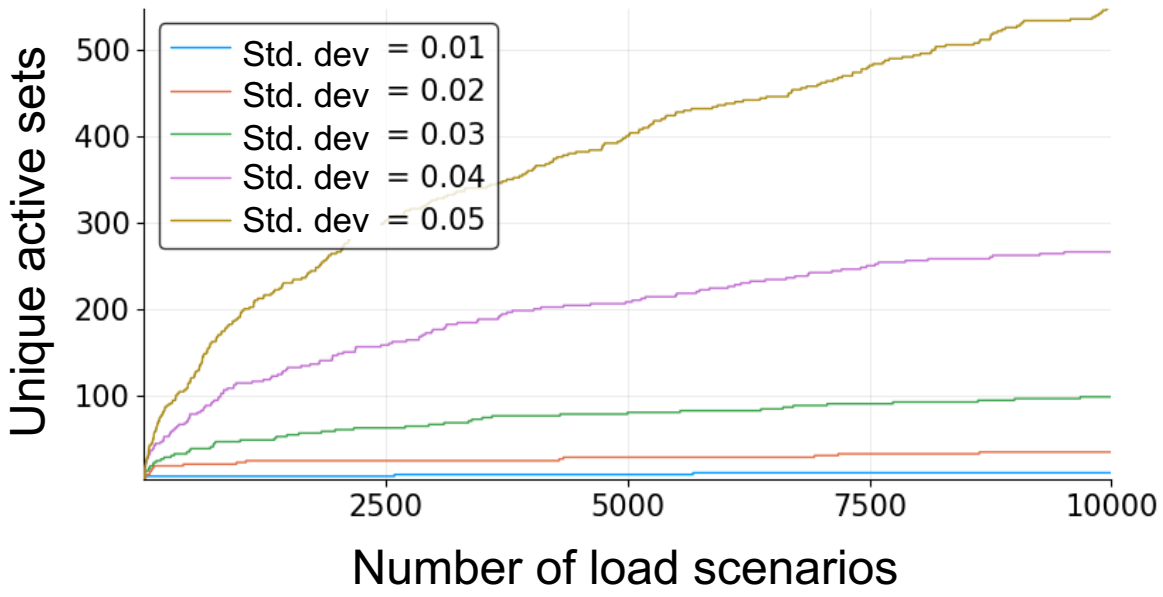




Few optimal active sets with high probability = the algorithm terminates fast!

How do the patterns change with renewable energy?

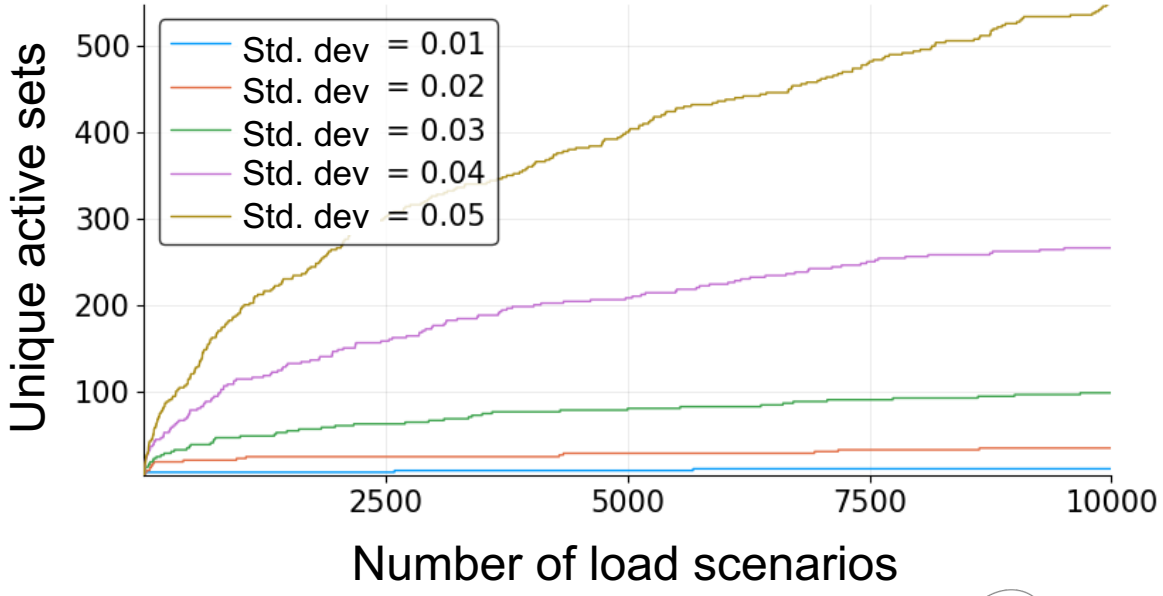
Renewable energy = Increased variability



Renewable energy =
Increasing load variability =
Increasing number of active sets

Results for IEEE 300 bus system with normally distributed load uncertainty [Ng, Misra, Roald, Backhaus, 2018]

Renewable energy = Increased variability



Renewable energy =
Increasing load variability =
Increasing number of active sets

Results for IEEE 300 bus system with normally distributed load uncertainty [Ng, Misra, Roald, Backhaus, 2018]

???

"Operators sometimes make wrong decisions because of increased variability"

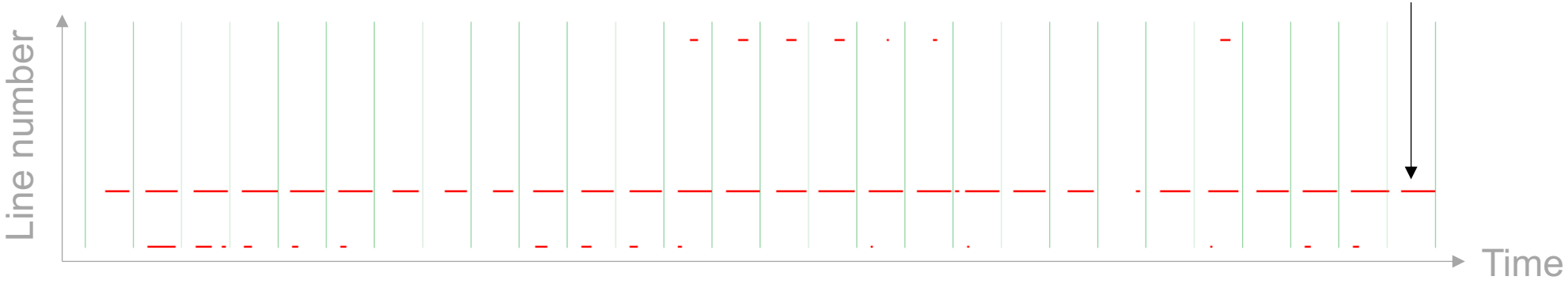
Patrick Panciatici (RTE, France)

Renewable energy = Increased variability

More variety in the active (congested) line constraints!

Red and blue:
Line is congested

Without
Renewable
Energy



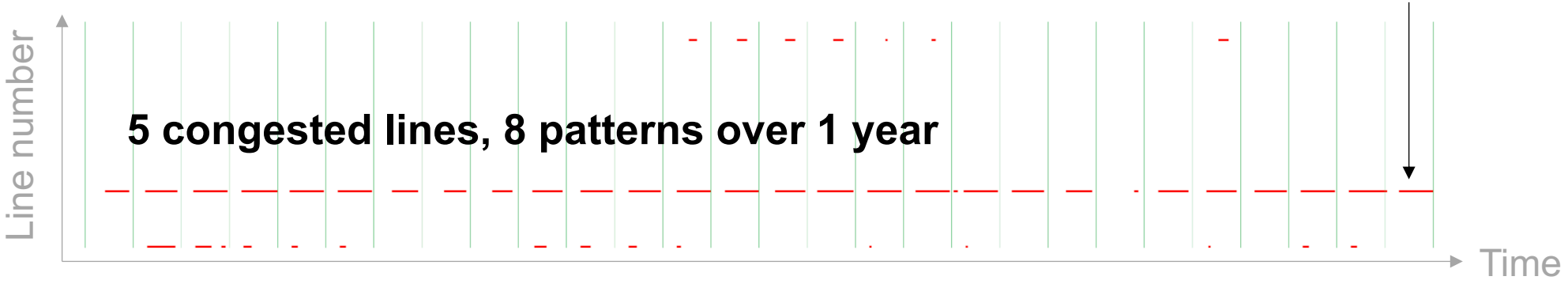
Active line constraints for the month of June in IEEE RTS-GMLC test case (with and without renewable energy).

Renewable energy = Increased variability

More variety in the active (congested) line constraints!

Red and blue:
Line is congested

Without
Renewable
Energy



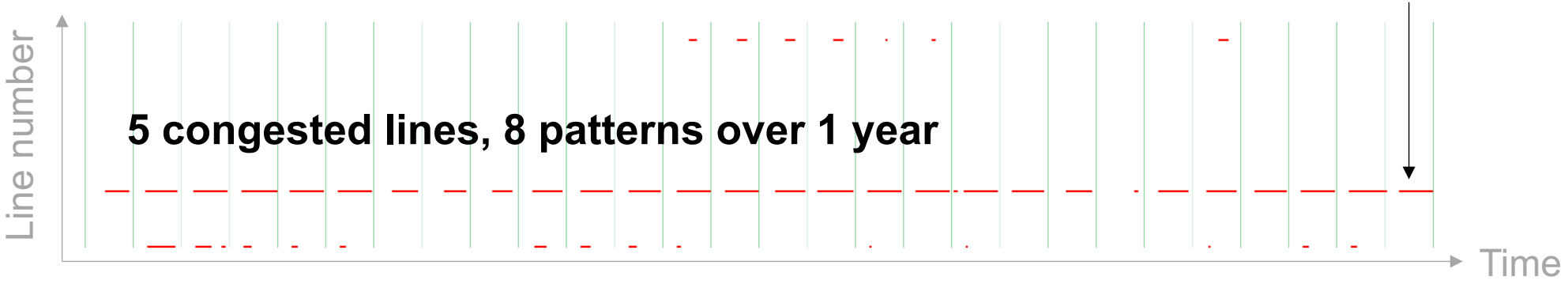
Active line constraints for the month of June in IEEE RTS-GMLC test case (with and without renewable energy).

Renewable energy = Increased variability

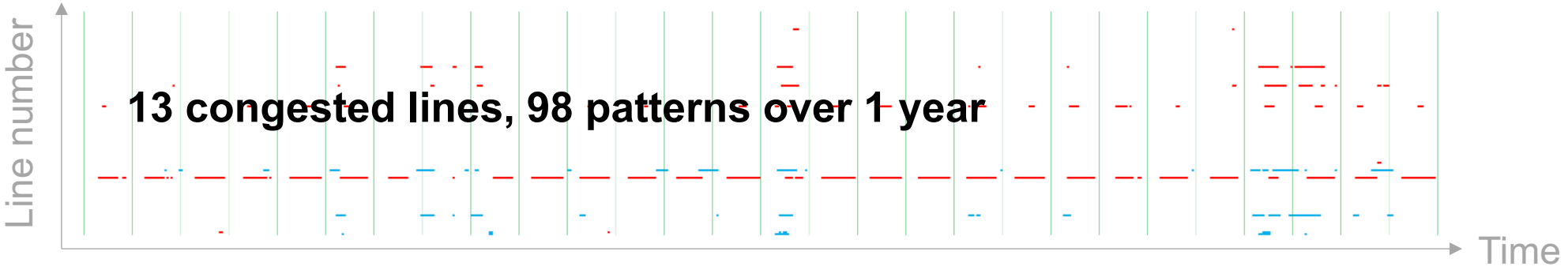
More variety in the active (congested) line constraints!

Red and blue:
Line is congested

Without
Renewable
Energy



With
Renewable
Energy



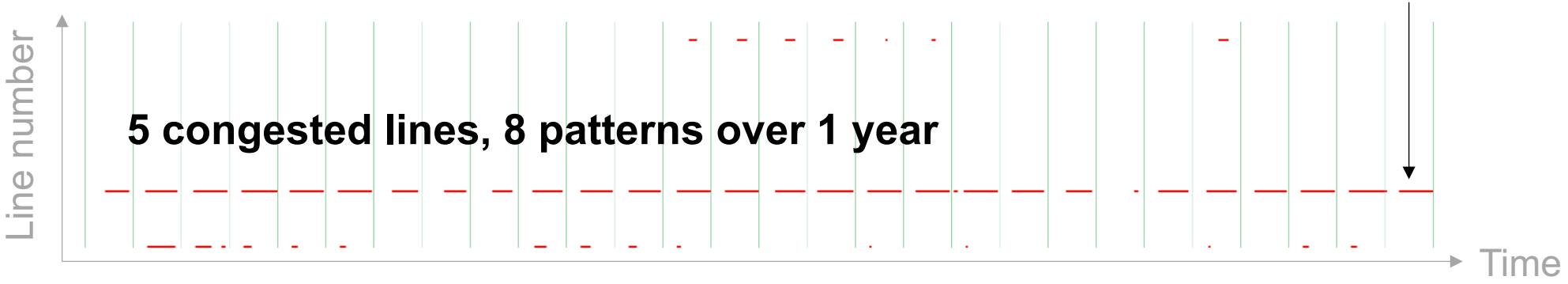
Active line constraints for the month of June in IEEE RTS-GMLC test case (with and without renewable energy).

Renewable energy = Increased variability

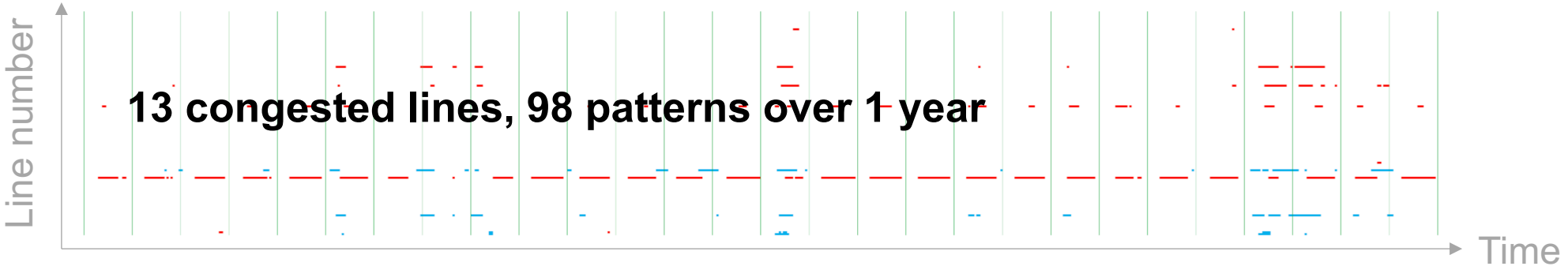
More variety in the active (congested) line constraints!

Red and blue:
Line is congested

Without
Renewable
Energy



With
Renewable
Energy



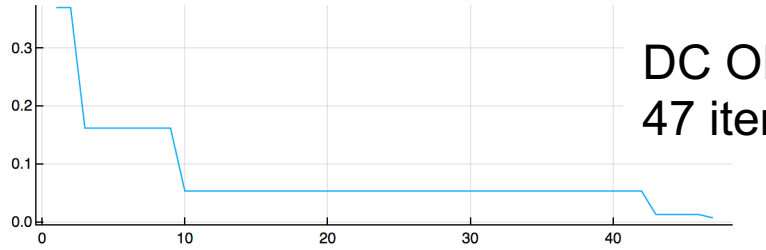
With more variable power injections, it gets harder to learn all the relevant patterns!

How about AC optimal power flow?

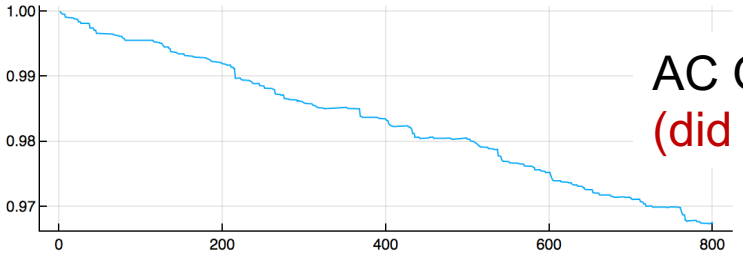
Learning AC congestion patterns

iteration

RTE 1951 bus test case

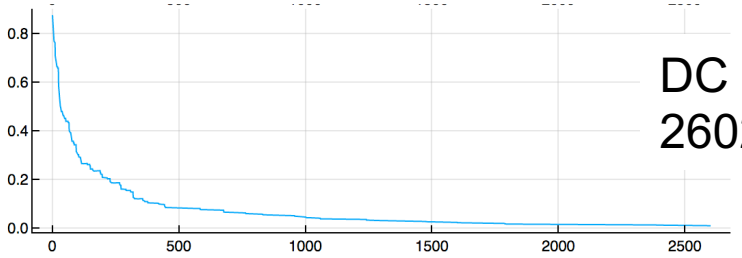


DC OPF
47 iterations

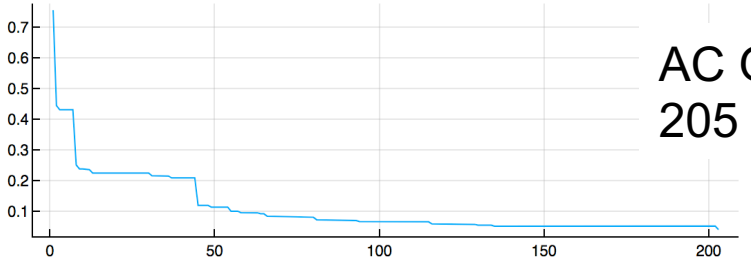


AC OPF ?
(did not terminate)

PSERC 200 bus test case



DC OPF
2602 iterations



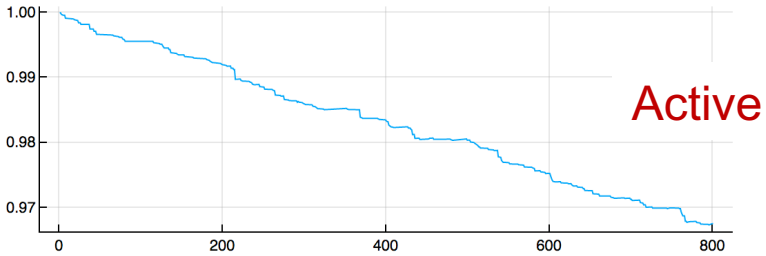
AC OPF
205 iterations

AC OPF is **typically** harder (voltage and reactive power constraints)

but not always

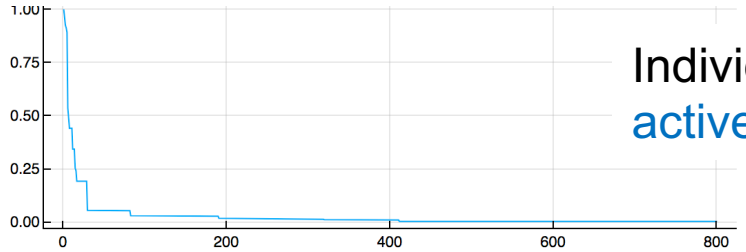
Learning AC active constraints

RTE 1951 bus test case



Active sets

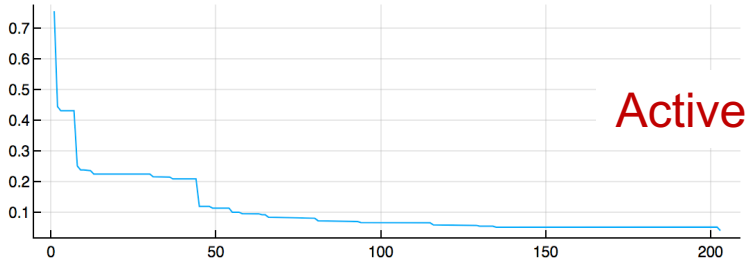
0.967



Individual
active constraints

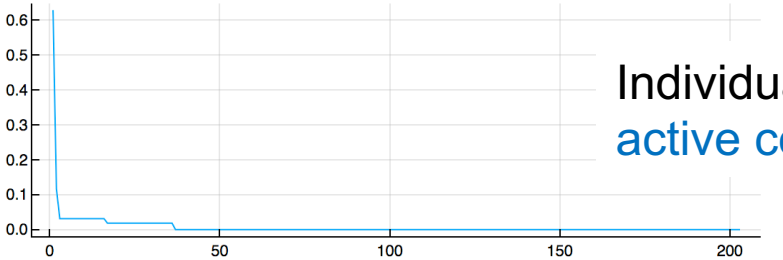
0.0035

PSERC 200 bus test case



Active sets

0.039



Individual
active constraints

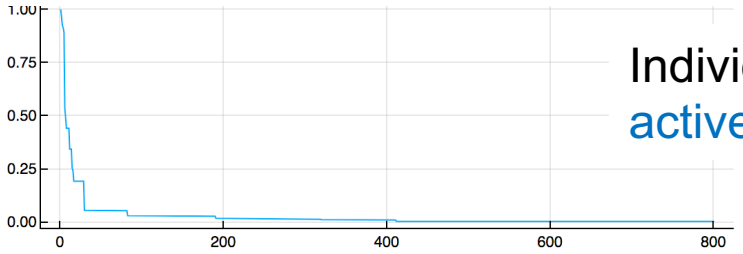
0.0001

If there are too many active sets, we can still learn **active constraints**.

Learning AC active constraints

RTE 1951 bus test case

Only 164 of 5192 transmission line constraints ever active

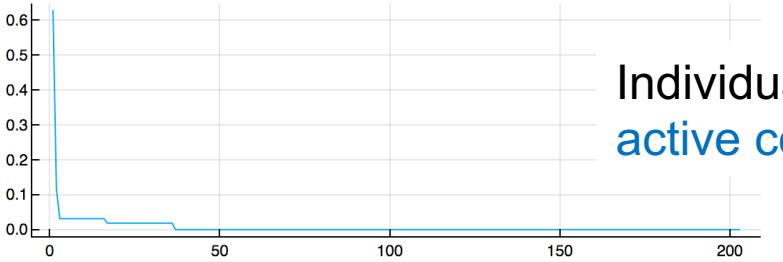


Individual
active constraints

0.0035

PSERC 200 bus test case

Only 28 of 490 transmission line constraints ever active



Individual
active constraints

0.0001

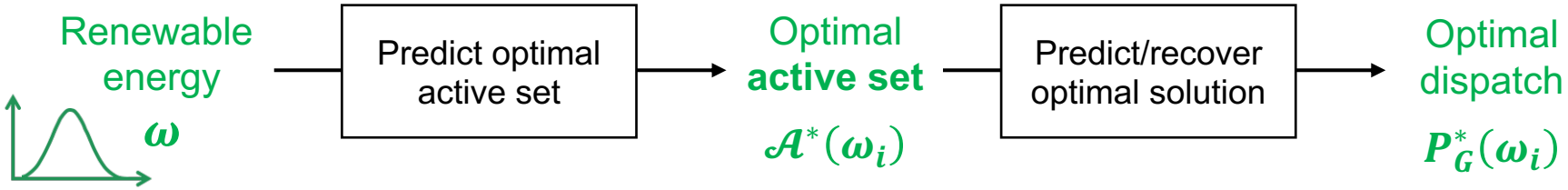
If there are too many active sets, we can still learn **active constraints**.

Only very few constraints seem to be at the optimal solution

Are there constraints that can never be active at all?

Learning the active constraints

Previous part of the talk: If you know the *set of active constraints* at optimum, you can recover the optimal solution!



- Optimal active set = “minimal” information we need to recover optimal solution
- Inherently encodes information about physical constraints and technical limits
- Finite, low dimensional object
- Nice practical interpretation as “operational patterns”

Learning the active constraints

If you know the *set of active constraints* at optimum, you can recover the optimal solution!

If you know that some constraints are *never active*, you can solve a smaller optimization problem!

Previously, we used learning to identify constraints that are probably not active.

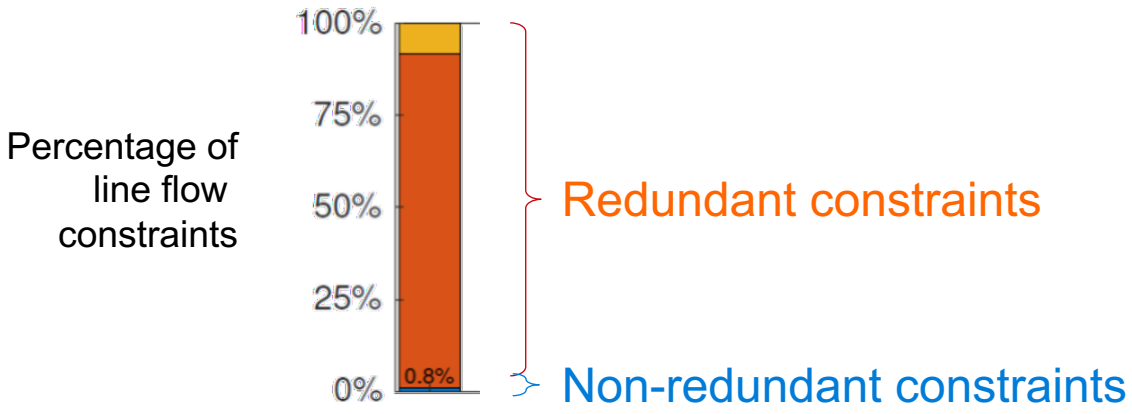
Let's see if we can find transmission constraints that will never be active (i.e., are redundant).

How many constraints can ever be active?

Previous part of the talk: If you know the *set of active constraints* at optimum, you can recover the optimal solution!

If you know that some constraints are *never active*,

you can solve a solve a smaller optimization problem!



2848 bus RTE test case

Only 0.8% constraints are non redundant!

1) Parallel lines

- For **parallel lines**, one flow limit may always be more restrictive.
- **Evaluating analytic condition** is sufficient to identifying redundant flow limits on parallel lines using only the lines' parameters.

$$\frac{b_{ij}^{(k)}}{f_{ij}^{max\ (k)}} < \frac{b_{ij}^{(l)}}{f_{ij}^{max\ (l)}}$$

- For some test cases, **approximately 20% of parallel lines** are identified as redundant.

Case Name	Num. Lines	Num. Parallel Lines	Num. Redundant Limits
PL-2383wp	2896	20	6
PL-2736sp	3269	12	4
PL-2737sop	3269	12	4
PL-2746wop	3307	16	5
PL-2746wp	3279	12	4
PL-3012wp	3572	12	5
PL-3120sp	3693	18	8
PL-3375wp	4161	178	5
PEGASE-89	210	8	2
PEGASE-1354	1991	519	203
PEGASE-2869	4582	1157	316
PEGASE-9241	16049	3503	650

[Molzahn '18]

2) Optimization-based constraint screening

$\min/\max_{p_G, \boldsymbol{p_D}}$	$f_l = M_{(l,\cdot)}(p_G - \boldsymbol{p_D})$	Minimize/maximize line flows	
s.t.	$\sum_{i=1}^{N_B}(p_{G(i)} - \boldsymbol{p_{D(i)}}) = 0$	Power balance	Non-redundant
	$\mathbf{0} \leq p_G \leq p_G^{max}$,	Generation constraints	Typically non-redundant
	$-f_L^{max} \leq M(p_G - \boldsymbol{p_D}) \leq f_L^{max}$,	Transmission constraints	**Often redundant**
	$(1 - X)\, p_D^{nom} \leq p_D \leq (1 + X)\, p_D^{nom}$,	Considered load variation	

Allow power demand $\boldsymbol{p_D}$ to vary $\pm X \cdot 100\%$ where $0 \leq X \leq 1$

Relax generator lower bounds to $\mathbf{0}$ (applicable to unit commitment!)

Find **maximum and minimum achievable flows** $\overline{\boldsymbol{f_l}}$, $\underline{\boldsymbol{f_l}}$

2) Optimization-based constraint screening

$\min/\max_{p_G, \boldsymbol{p_D}}$	$f_l = M_{(l,\cdot)}(p_G - \boldsymbol{p_D})$	Minimize/maximize line flows	
s.t.	$\sum_{i=1}^{N_B}(p_{G(i)} - \boldsymbol{p_{D(i)}}) = 0$	Power balance	Non-redundant
	$\mathbf{0} \leq p_G \leq p_G^{max}$,	Generation constraints	Typically non-redundant
	$-f_L^{max} \leq M(p_G - \boldsymbol{p_D}) \leq f_L^{max}$,	Transmission constraints	**Often redundant**
	$(1-X)\, p_D^{nom} \leq p_D \leq (1+X)\, p_D^{nom}$,	Considered load variation	

Allow power demand $\boldsymbol{p_D}$ to vary $\pm X \cdot 100\%$ where $0 \leq X \leq 1$

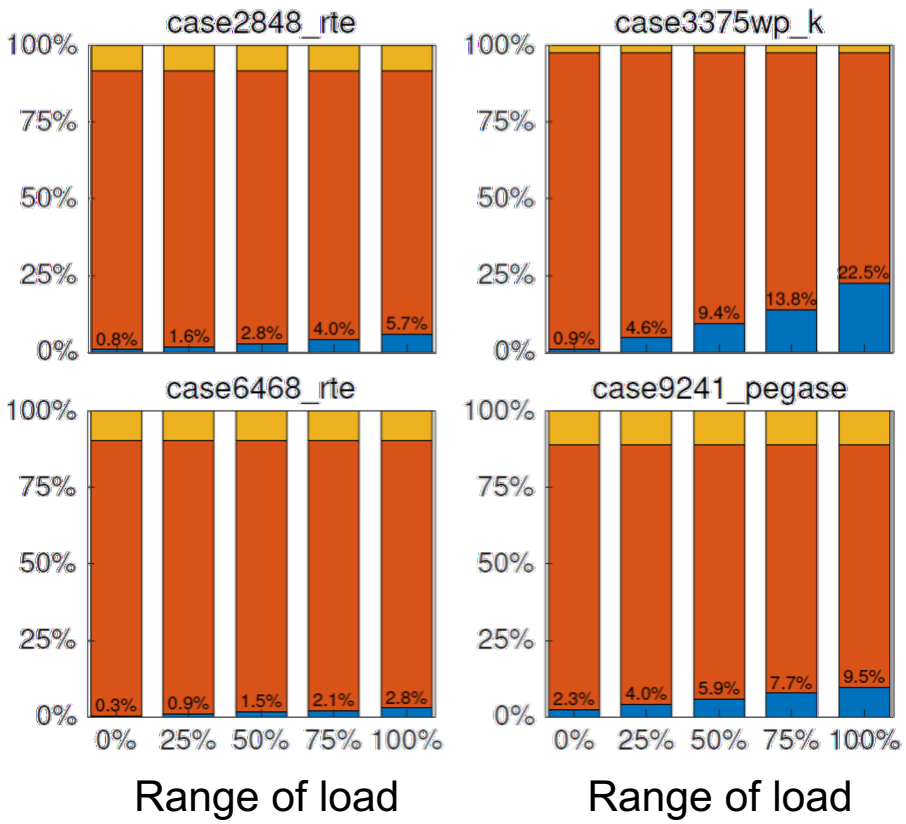
Relax generator lower bounds to $\mathbf{0}$ (applicable to unit commitment!)

Find **maximum and minimum achievable flows** $\overline{f_l}$, $\underline{f_l}$

If $\overline{f_l} < f_{L,l}^{\max}$ or $\underline{f_l} > f_{L,l}^{\min}$ $\Rightarrow$ Constraint is redundant!

Many redundant constraints…

Percentage of line flow constraints



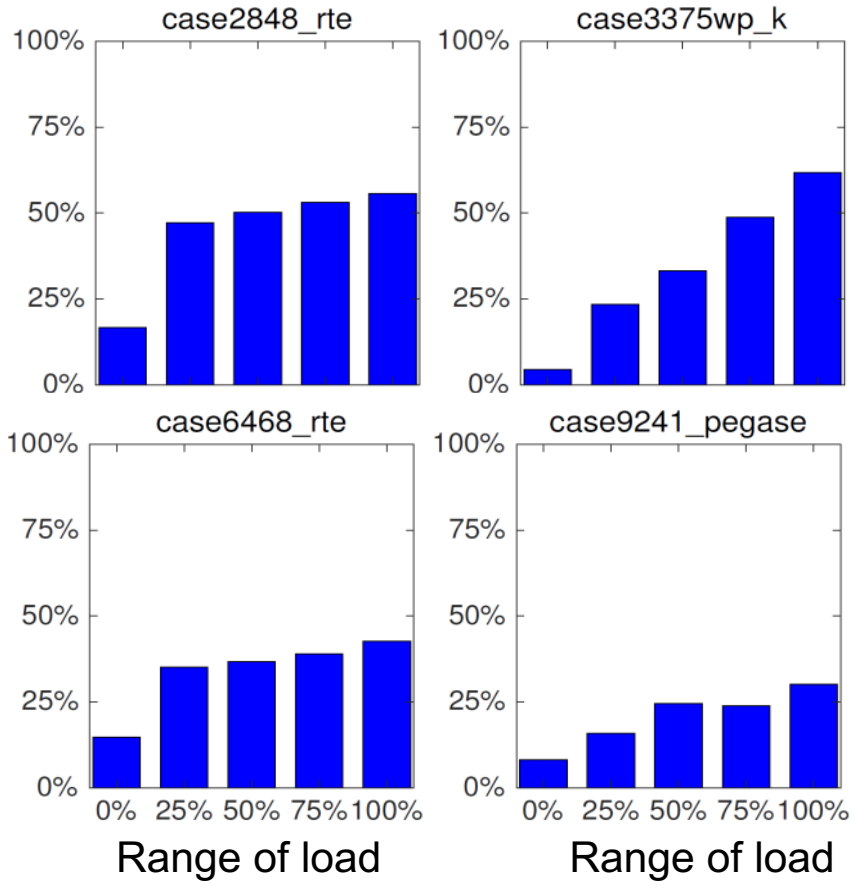
Remaining Constraints
Optimization Elimination
Parallel Line Elimination

Constraint screening results for test cases from PGLib v.17.08

…even for large load variations!

Significant reduction in computation…

Relative computational time



Single-period Unit Commitment problem

Results based on 100 computations of UC for test cases from PGLib v.17.08

…even for large load variations!

Summary

Congestion pattern = Optimal active set of OPF

In traditional operations, only a few constraints are relevant

… our algorithms can identify them!

Renewable energy variability increases complexity of system operations = increasing number of operational patterns!

A good case for optimization? ☺

Thank you!

roald@wisc.edu

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

Sidhant Misra, Line Roald and Yee Sian Ng, «Learning for Constrained Optimization», https://arxiv.org/abs/1802.09639

L.A. Roald and D.K. Molzahn, "Implied Constraint Satisfaction in Power System Optimization: The Impacts of Load Variations," *57th Annual Allerton Conference on Communication, Control, and Computing,* 2019.