The University of Texas at Austin
Electrical and Computer Engineering

NREL AES Workshop
July 2022

Physics-aware and Risk-aware Machine Learning for Power System Operations

Hao Zhu
Assistant Professor of ECE
Kilby/TI Endowed Faculty Fellow
The University of Texas at Austin

Acknowledgements: Shanny Lin, Shaohui Liu, Kyung-bin Kwon, and Yuqi Zhou
NSF Grants 1802319, 1952193, 2130706, DOE-EPRI GREAT-Data

Power of AI/ML

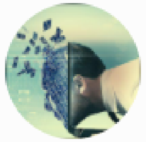
- Unprecedented opportunities offered by diverse sources of data
  - Synchrophasor and IED data
  - Smart meter data
  - Weather data
  - GIS data, ...

How to harness the power of ML to tackle problem-specific challenges in real-time power system decision making?

Forbes

Mar 15, 2019, 07:37am EDT | 21,849 views

How AI Can And Will Predict Disasters



Naveen Joshi Former Contributor
COGNITIVE WORLD Contributor Group
AI

AI could put a stop to electricity theft and meter misreadings

TECHNOLOGY 23 September 2017

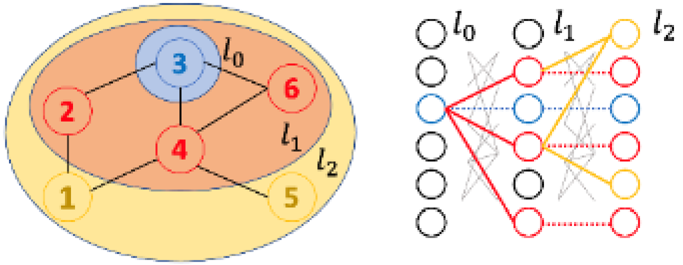
SUSTAINABLE ENERGY

Combining A.I. and human knowledge could transform how power grids work

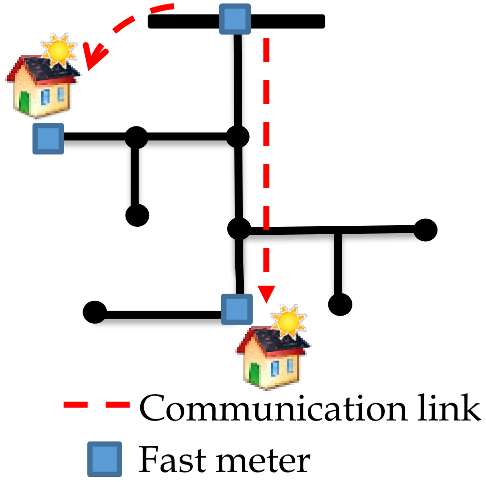
PUBLISHED FRI, SEP 27 2019·4:55 AM EDT | UPDATED FRI, SEP 27 2019·4:55 AM EDT

Overview

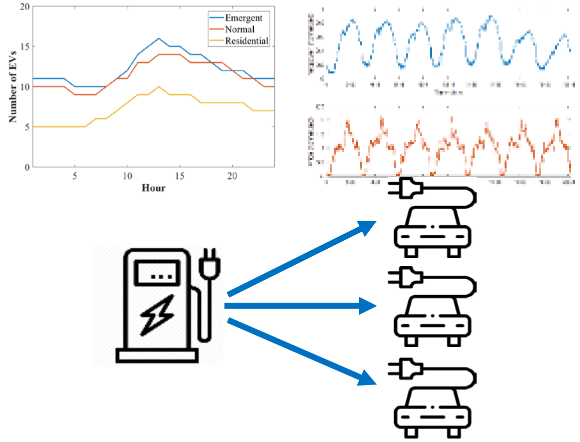
- We visit three problems that use domain knowledge to better design learning models that are physics-informed and risk-aware



Topology-aware learning in large-scale power systems:
Simpler model structure



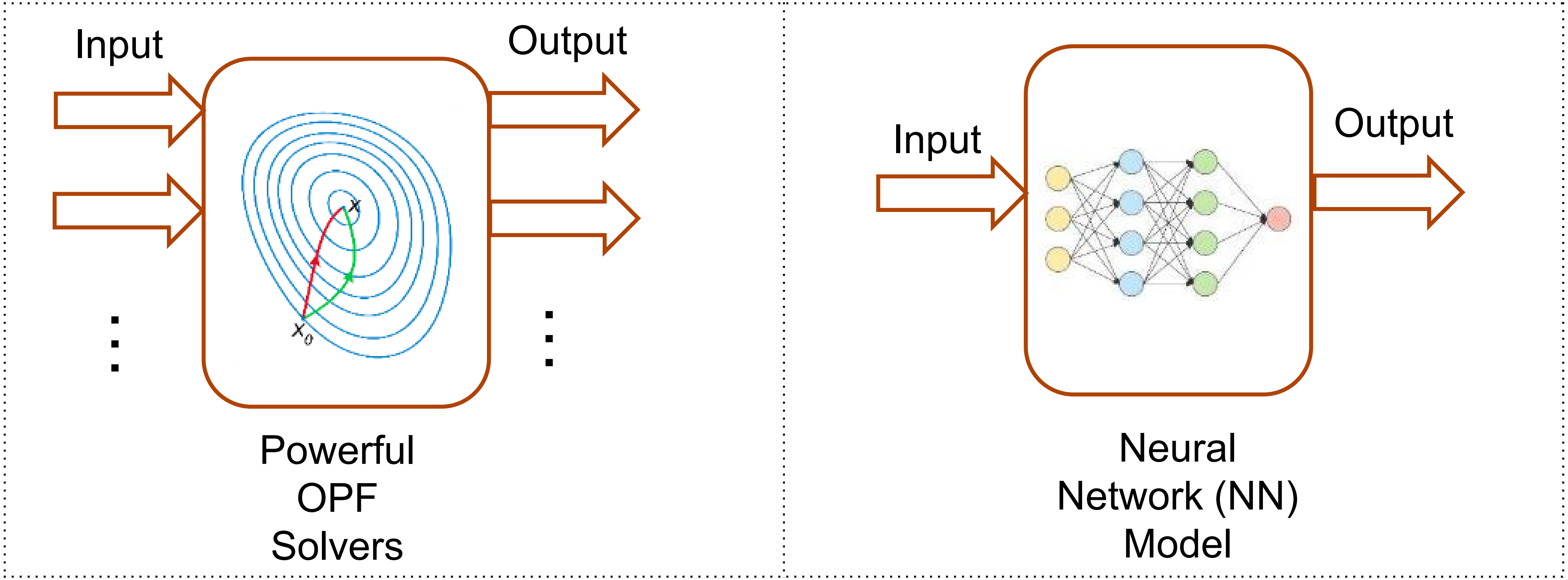
Risk-aware learning for grid-edge coordination:
Reduced risks of voltage violations



Reinforcement learning for dynamical resources:
More efficient representation

Part I: Topology-aware Learning in Large-scale Power Systems

ML for optimal power flow (OPF)



- Attain a pre-trained OPF input-output mapping from available samples

Existing work and our focus

- Integration of renewable, flexible resources increases the grid variability and motivates real-time, feasible OPF via training a neural network (NN)
  - Warm start the search for ac feasible solution [Baker '19]
  - Feasible domain to reduce limit violation [Zamzam et al'20][Zhao et al'21]
  - KKT conditions based regularization [Zhang et al'22] [Nellikkath et al'22]
- Connection to the duality analysis of convex OPF [Chen et al'20] [Singh et al'20]
- Rely on FCNN architecture and cannot adapt to varying topology

Focus: graph learning approach for *complexity reduction & topology adaptivity*

Real-time OPF

- Power network modeled as a graph $G = (\mathcal{V}, \mathcal{E})$ with N nodes
- ac-OPF for all nodal injections

$$
\begin{aligned}
\min_{\mathbf{p},\mathbf{q},\mathbf{v}} \quad & \textstyle\sum_{i=1}^{N} c_i(p_i) \\
\text{s.t.} \quad & \mathbf{p} + \mathrm{j}\mathbf{q} = \mathrm{diag}(\mathbf{v})(\mathbf{Y}\mathbf{v})^* \\
& \underline{\mathbf{V}} \leq |\mathbf{v}| \leq \bar{\mathbf{V}} \\
& \underline{\mathbf{p}} \leq \mathbf{p} \leq \bar{\mathbf{p}} \\
& \underline{\mathbf{q}} \leq \mathbf{q} \leq \bar{\mathbf{q}} \\
& s_{ij}(\mathbf{v}) \leq \bar{s}_{ij}\,, \quad \forall (i,j) \in \mathcal{E}
\end{aligned}
$$

- Nodal input:

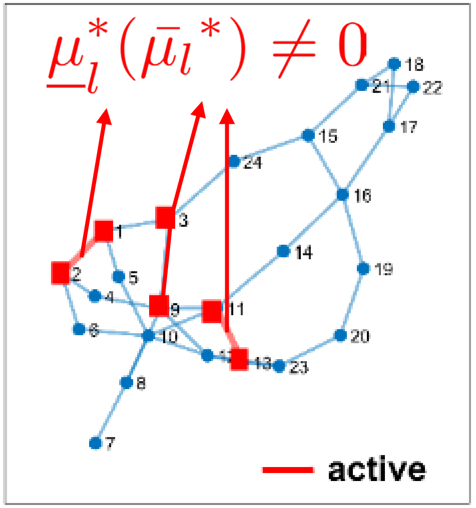
$$\mathbf{x}_i \triangleq [\bar{p}_i, \underline{p}_i, \bar{q}_i, \underline{q}_i, \mathbf{c}_i] \in \mathbb{R}^d$$

power limits + costs

- Nodal output: optimal p/q

Each FCNN layer has $\mathcal{O}(N^2)$ parameters!

Topology dependence



- [Owerko et al'20] using graph learning to predict p/q
- But topology dependence (locality) of output label is crucial!
- Locational marginal price from (very few) congested lines
- Voltage magnitude $|\mathbf{v}|$ approximated using q injection

$$
\begin{aligned}
\min_{\mathbf{p}} \quad & \textstyle\sum_{i=1}^{N} c_i(p_i) \\
\text{s.t.} \quad & \mathbf{1}^\top \mathbf{p} = 0 & : \lambda \\
& \underline{\mathbf{p}} \leq \mathbf{p} \leq \bar{\mathbf{p}} \\
& \underline{\mathbf{f}} \leq \mathbf{S}\mathbf{p} \leq \bar{\mathbf{f}} & : [\underline{\boldsymbol{\mu}};\ \bar{\boldsymbol{\mu}}]
\end{aligned}
$$

$$
\boldsymbol{\pi}^* := \lambda^* \cdot \mathbf{1} - \mathbf{S}^\top (\bar{\boldsymbol{\mu}}^* - \underline{\boldsymbol{\mu}}^*)
$$

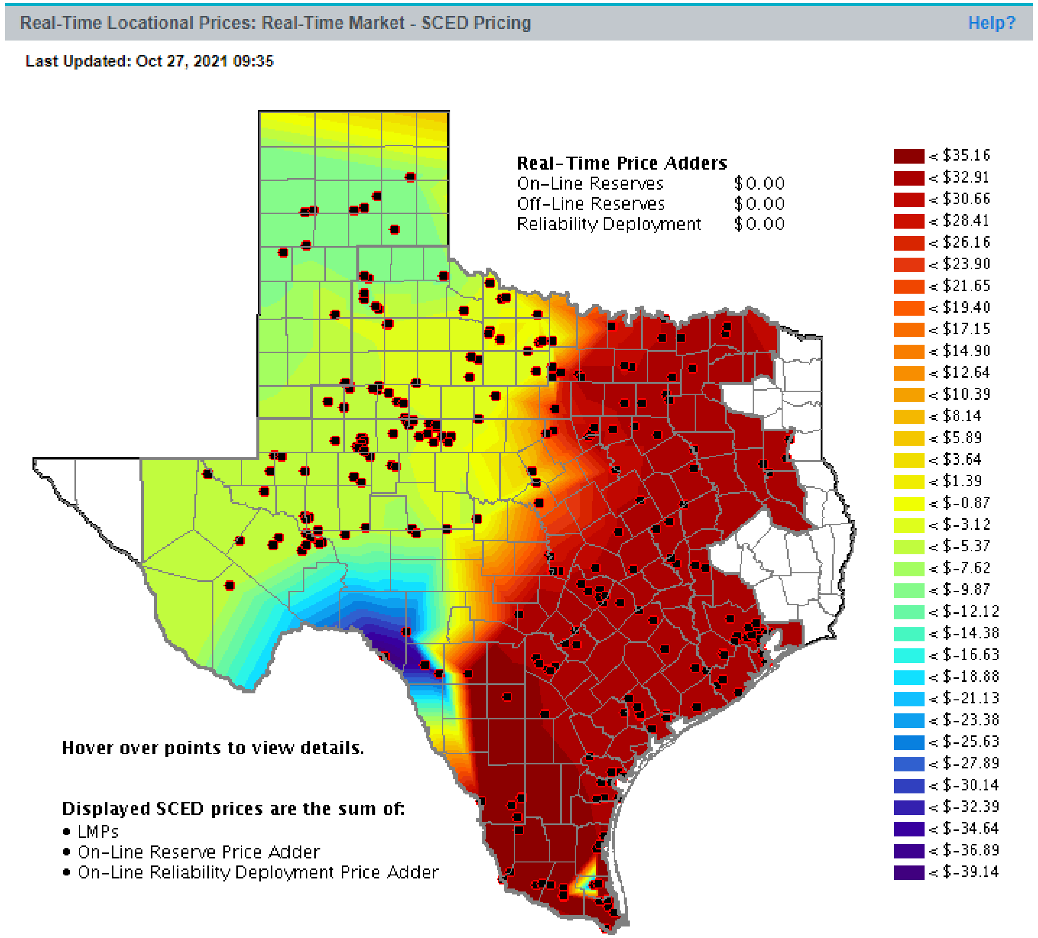
$$
\mathbf{S}^\top = \mathbf{B}^{-1} \mathbf{A}^\top \mathbf{X}^{-1}
$$

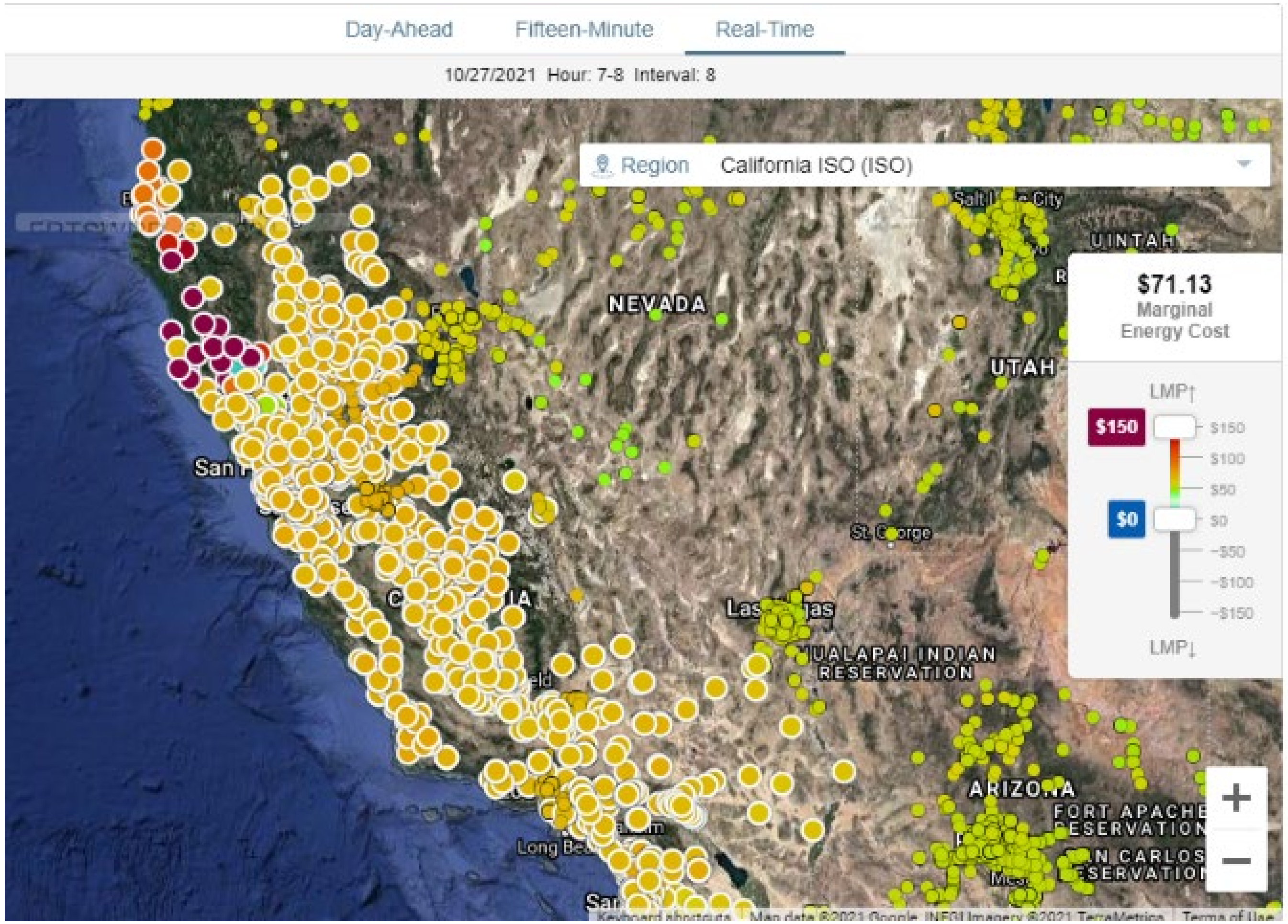
Spanned from the eigen-space of Bbus matrix $\mathbf{B}$ (graph Laplacian)

$$
\mathbf{p} + \mathrm{j}\mathbf{q} = \mathrm{diag}(\mathbf{v})(\mathbf{Y}\mathbf{v})^*
$$

$$
\Delta|\mathbf{v}| \approxeq -\mathbf{B}^{-1} \Delta \mathbf{q}
$$

Locational marginal price (LMP) map





Graph NN (GNN)

- Input formed by nodal features as rows

$$\mathbf{X}^0 = \{\mathbf{x}_i\} \in \mathbb{R}^{N \times d}$$

- GNN layer *l* with learnable parameters

$$\mathbf{X}^{\ell+1} = \sigma\left(\mathbf{W}^\ell \mathbf{X}^\ell \mathbf{H}^\ell + \mathbf{b}^\ell\right)$$

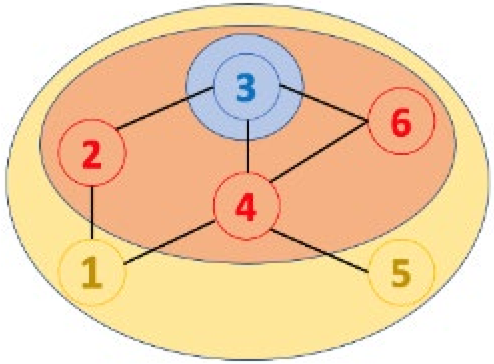
  - Topology-based *graph filter* $\mathbf{W}^\ell \in \mathbb{R}^{N \times N}$

$$[\mathbf{W}^\ell]_{ij} = 0 \text{ if } (i,j) \notin \mathcal{E}$$

  - Feature filters $\{\mathbf{H}^\ell\}$ for higher-dim. nonlinearity

- GNN used for grid fault location [Li-Deka'21]

Hamilton, William L. "Graph representation learning." 2020.
https://www.cs.mcgill.ca/~wlh/grl_book/



$$\mathbf{W} = \begin{bmatrix} W_{11} & W_{12} & 0 & W_{14} & 0 & 0 \\ W_{21} & W_{22} & W_{23} & 0 & 0 & 0 \\ 0 & W_{32} & W_{33} & W_{34} & 0 & W_{36} \\ W_{41} & 0 & W_{43} & W_{44} & W_{45} & W_{46} \\ 0 & 0 & 0 & W_{54} & W_{55} & 0 \\ 0 & 0 & W_{63} & W_{64} & 0 & W_{66} \end{bmatrix}$$

Input feature X^0 is a 6xd matrix

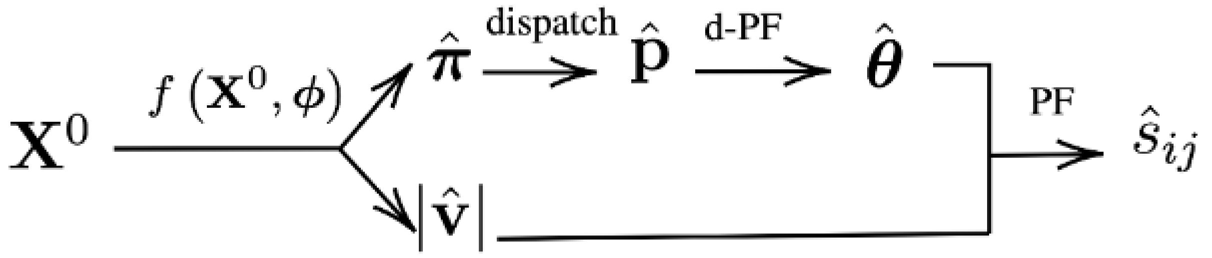
Prop. 1 *(GNN complexity)*:
If lines are sparse $|\mathcal{E}| \sim \mathcal{O}(|\mathcal{V}|)$ and let $D = \max_t\{d_t\}$, then the **number of parameters** for each GNN layer is $\mathcal{O}(N + D^2)$.

Compared to FCNN's $\mathcal{O}(N^2)$

From GNN outputs to OPF variables

- LMP decides (feasible) p from economics $\hat{p}_i = \text{argmin}_{\underline{p}_i \le p_i \le \bar{p}_i} \; c_i(p_i) - \hat{\pi}_i p_i$
- Decoupled (d-)PF approximates angle $\hat{\boldsymbol{\theta}} \approxeq \boldsymbol{\theta}_o + \mathbf{J}_{p\theta}^{-1}(\hat{\mathbf{p}} - \mathbf{p}_o)$

- GNN outputs of LMP and $|\mathbf{v}|$ can fully determine the power flow



Liu, Shaohui, Chengyang Wu, and Hao Zhu. "Topology-aware Graph Neural Networks for Learning Feasible and Adaptive AC-OPF Solutions," submitted. https://arxiv.org/pdf/2205.10129

Feasibility regularization (FR)

- Loss function for predicting LMP and $|\mathbf{v}|$

$$\mathcal{L}(\phi) := \|\boldsymbol{\pi}^* - \hat{\boldsymbol{\pi}}\|_2^2 + \left\||\mathbf{v}^*| - |\hat{\mathbf{v}}|\right\|_2^2 + \lambda_\infty \|\boldsymbol{\pi}^* - \hat{\boldsymbol{\pi}}\|_\infty$$

  - Infinity-norm on LMP due to its larger variability than $|\mathbf{v}|$

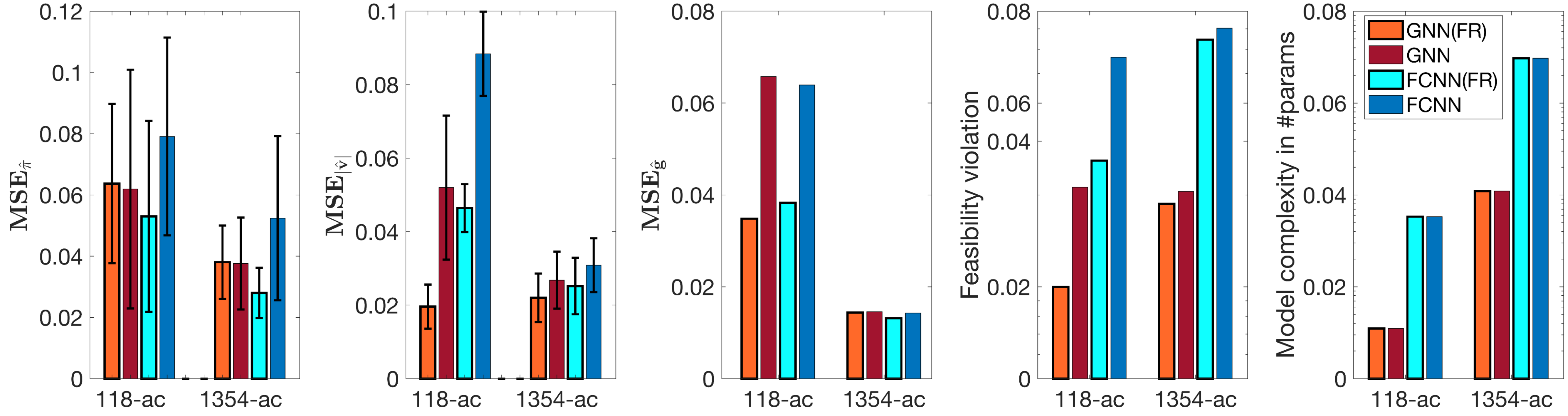
- Network-wide line limits are difficult to satisfy
- FR to reduce line flow violations: $\mathcal{L}'(\phi) := \mathcal{L}(\phi) + \lambda \left\|\mathbb{P}_{[\mathbf{0},\infty]}[\hat{\mathbf{s}} - \bar{\mathbf{s}}]\right\|_1$

Prop. 2 *(Feasibility):* ac-FR based OPF learning is a *fully feed-forward* NN. The proposed FR term still allows for efficient using ***autograd*** and ***backpropagation***. The feasibility of both predicted $|\hat{\mathbf{v}}|$ and $\hat{\mathbf{p}}$ can be strictly enforced via projections, as well.

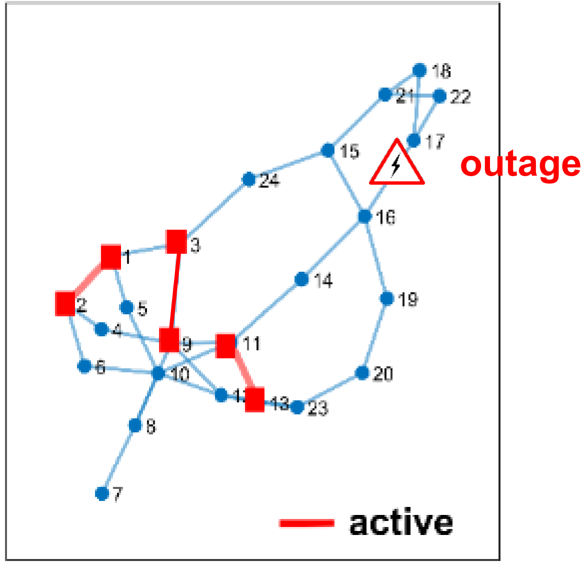
Liu, Shaohui, Chengyang Wu, and Hao Zhu. "Topology-aware Graph Neural Networks for Learning Feasible and Adaptive AC-OPF Solutions," submitted. https://arxiv.org/pdf/2205.10129

Benchmark results

- 118-bus and 1354-bus for ac-opf
- **Metrics:** normalized MSE; line flow limit violation rate; model complexity
- GNN, FCNN, both + feasibility regularization (FR)

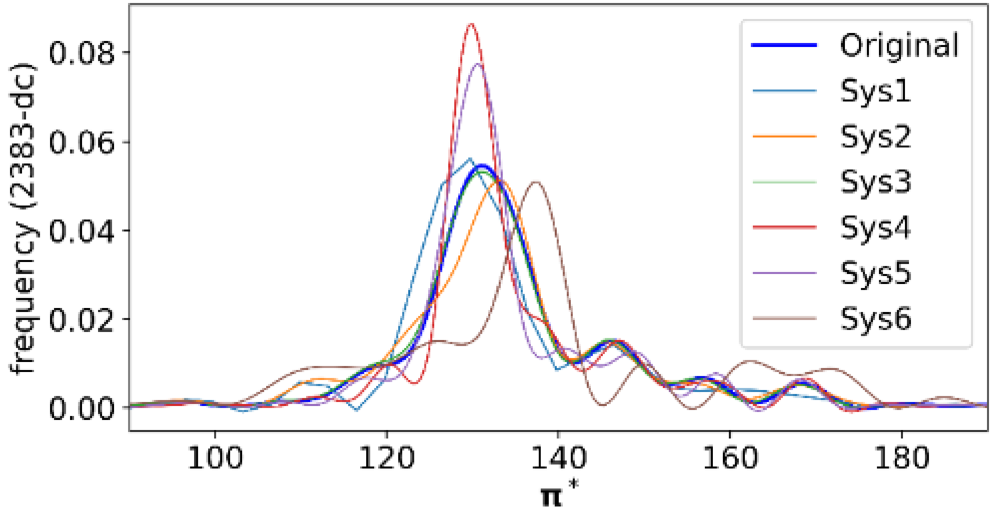


OPF learning under contingency



- Topology-agnostic NNs lack in transfer capability
  - Sample re-generation and re-training are time-consuming
- OPF outputs tend to be stable under line outages
  - Thanks to stability of the eigen-space

    $span\{\mathbf{U}\}$ with $\mathbf{B}^{-1} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\top$
  - LMP outputs slightly vary with the outages of multiple lines (of high capacity)
- We have established analytical bounds for this perturbation on graph subspace



Stability analysis

- Under line k outage, rank-one perturbations on $\mathbf{B}' = \mathbf{B} - \frac{1}{x_k}\mathbf{a}_k\mathbf{a}_k^\top$ and its inverse

$$(\mathbf{B}')^{-1} = \mathbf{B}^{-1} + \mathbf{\Delta}_k = \mathbf{B}^{-1} + \frac{\mathbf{B}^{-1}\mathbf{a}_k\mathbf{a}_k^\top\mathbf{B}^{-1}}{x_k - \mathbf{a}_k^\top\mathbf{B}^{-1}\mathbf{a}_k}$$

[Matrix inversion lemma]

- Difference between the corresponding (sub)spaces

$$d(span\{\mathbf{U}\}, span\{\mathbf{U}'\}) := \|\text{diag}\,(\sin\theta_1, \cdots, \sin\theta_{N-1})\,\|_F \text{ with } \quad \theta_i := \theta(\mathbf{u}_i, \mathbf{u}_i')$$

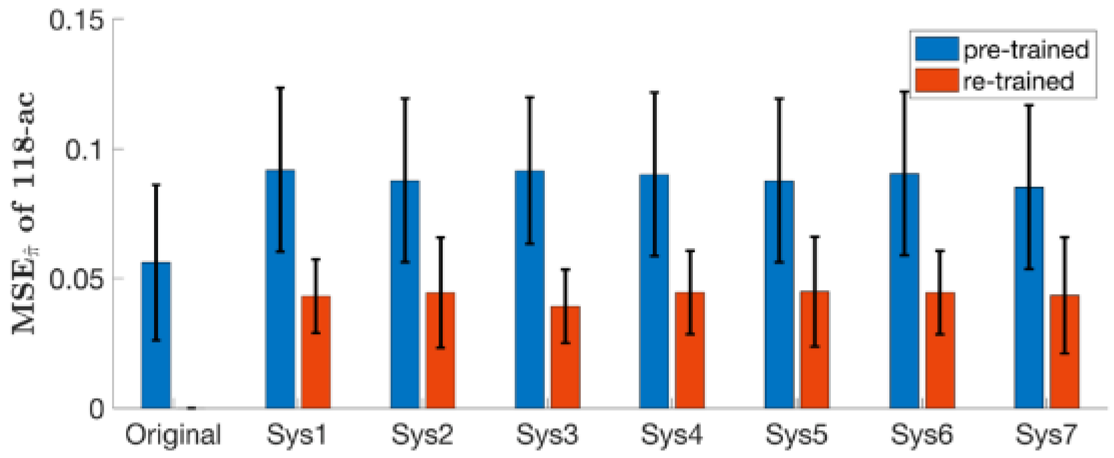
Prop. 3 *(Bounded subspace difference):* Let $\{\lambda_i\}_{i=1}^{N-1}$ and $\{\lambda_i'\}_{i=1}^{N-1}$ represent the respective positive eigenvalues of $\mathbf{B}^{-1}$ and $(\mathbf{B}')^{-1}$ in non-increasing order. Consider the first s eigenvalues with the minimum separations $\delta \triangleq \min_{1\leq i\leq s-1}(\lambda_i - \lambda_{i+1})$ and $\delta' \triangleq \min_{1\leq i\leq s-1}(\frac{1}{\lambda_{i+1}} - \frac{1}{\lambda_i})$.

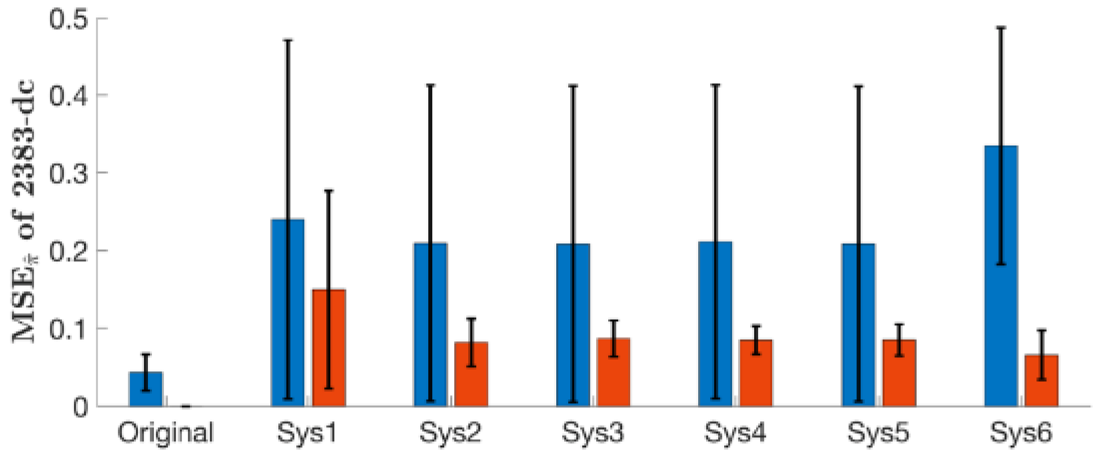
We can bound the difference between the leading sub-spaces $span(\mathbf{U}_s) \triangleq span([\mathbf{u}_1, \cdots, \mathbf{u}_s])$

$$d(span\{\mathbf{U}_s\}, span\{\mathbf{U}_s'\}) \leq \min\left(\frac{\|\mathbf{\Delta}_k\|_F}{\delta}, \frac{2}{x_k \cdot \delta'}\right).$$

GNN topology transfer learning

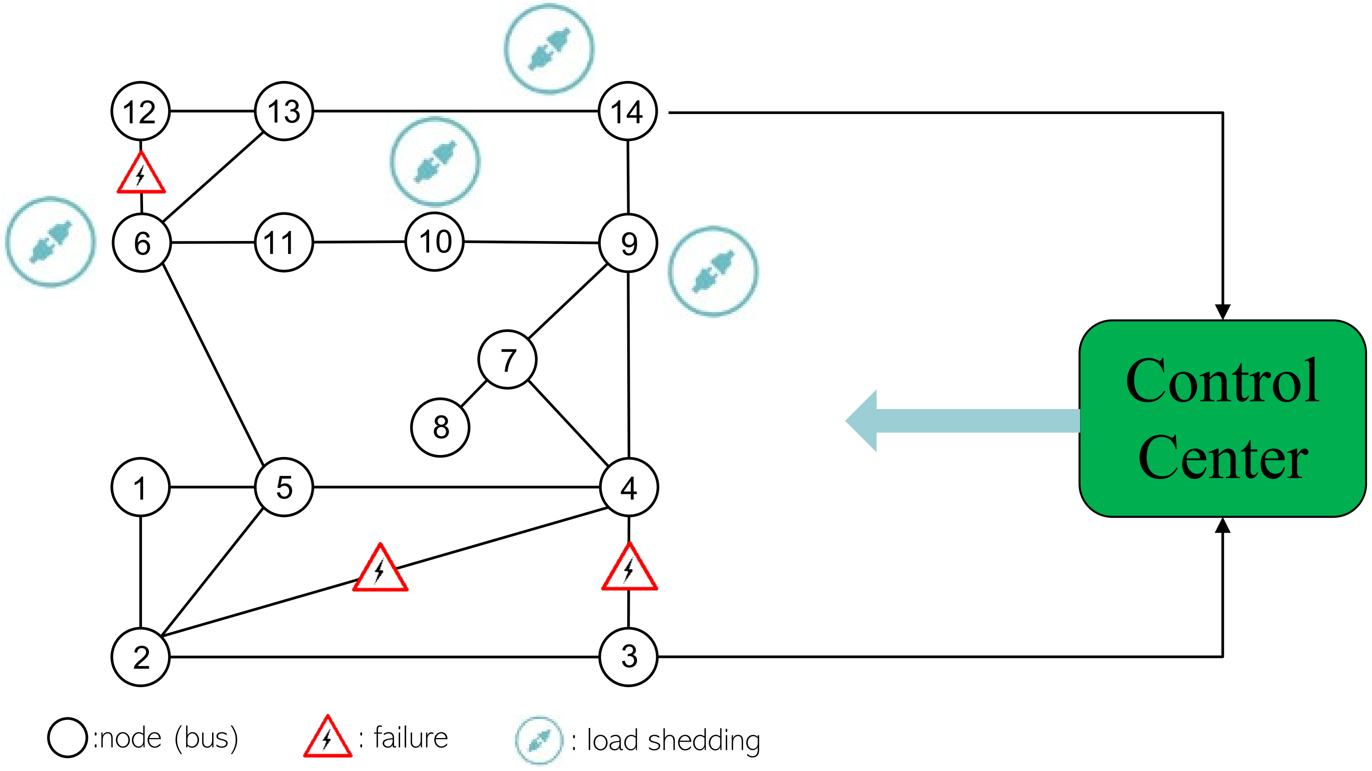
- Perturb the original system with the outages of 2-4 lines of high capacity
- Pre-trained GNN for the original system has reasonable error rates
  - warm-start the re-training using only half of samples
- GNN exhibits excellent adaptivity to the varying grid topology
  - Re-training takes *only 3-5 epochs* to converge to the original performance





Centralized load shedding

- Grid resilience challenged by resource variability and extreme weather
- Optimal load shedding (OLS) is a special case of ac-OPF

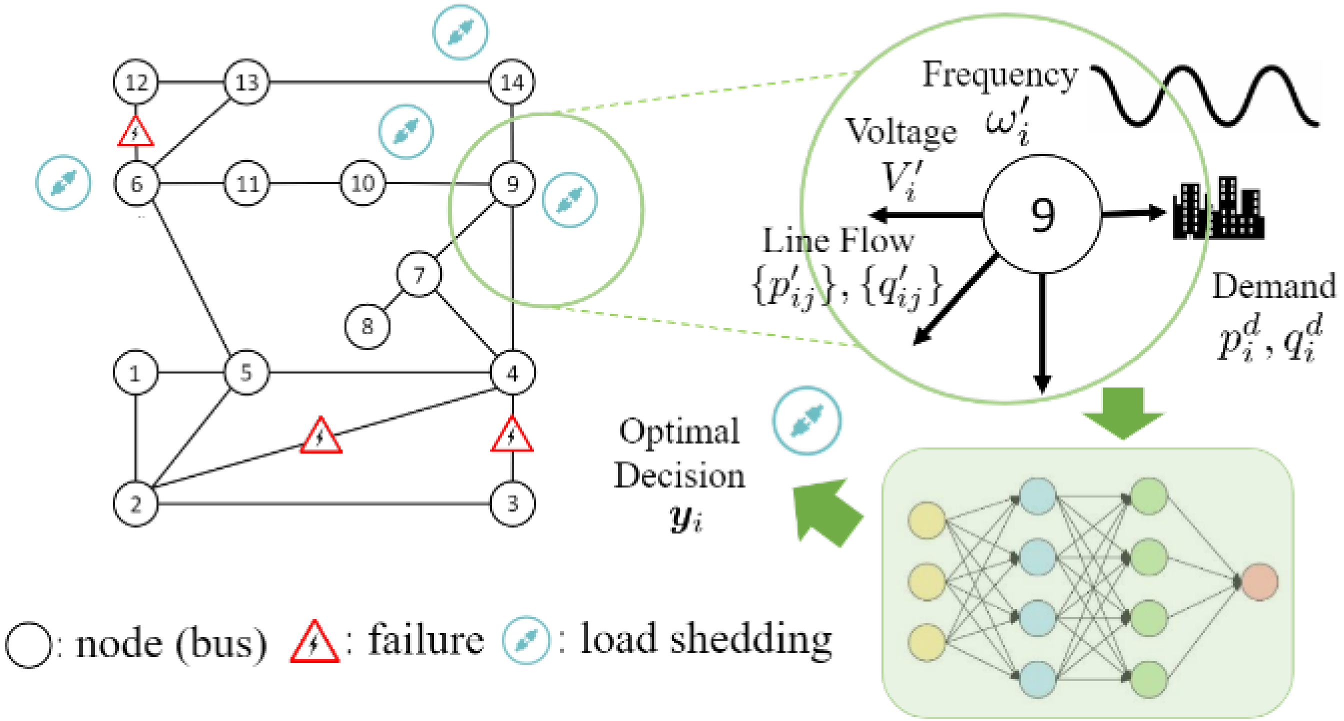


- Centralized optimization using system-wide information
- However, need very fast-speed communication links and computation capability

- Can we use ML to enable *scalable OLS* at each node using *local information* only?

ML for decentralized load shedding

- Each load center learns the decision rule from historical or synthetic scenarios



- Input feature:

$$\boldsymbol{x}_i = [p^d_i, q^d_i, V'_i, \{p'_{ij}\}, \{q'_{ij}\}, \omega'_i]$$

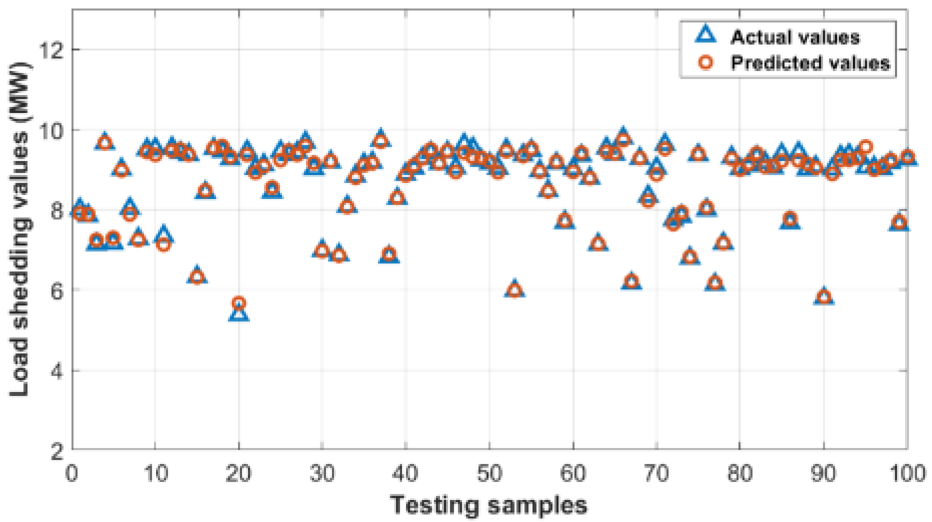
- Local reduction solutions:

$$\boldsymbol{y}_i = [p^s_i, q^s_i]$$

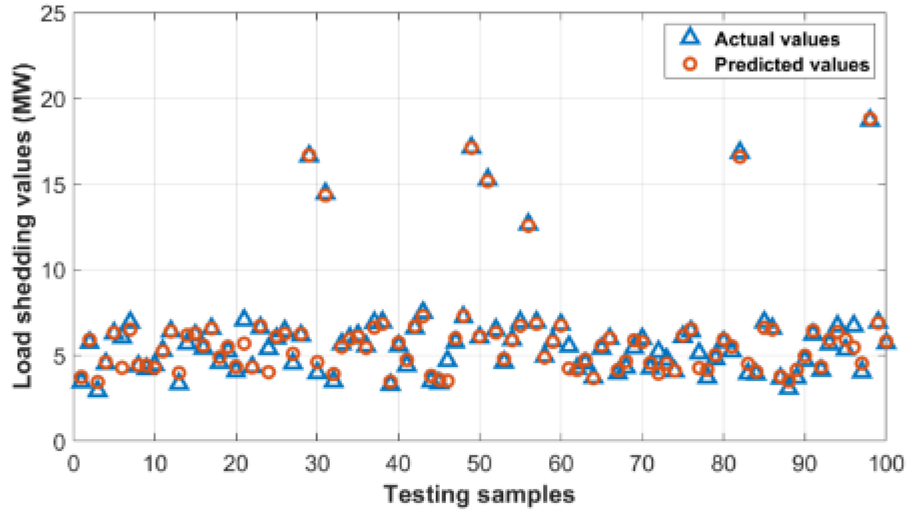
Yuqi Zhou, Jeehyun Park, and Hao Zhu, “Scalable Learning for Optimal Load Shedding Under Power Grid Emergency Operations,” *PES General Meeting (PESGM)* 2022 (accepted) https://arxiv.org/abs/2111.11980

Prediction under single line outage

- IEEE 14-bus system; quadratic cost functions
- All $(N-1)$ contingency scenarios, under different load conditions (1000 samples for each scenario)



(a) Load center at bus 10

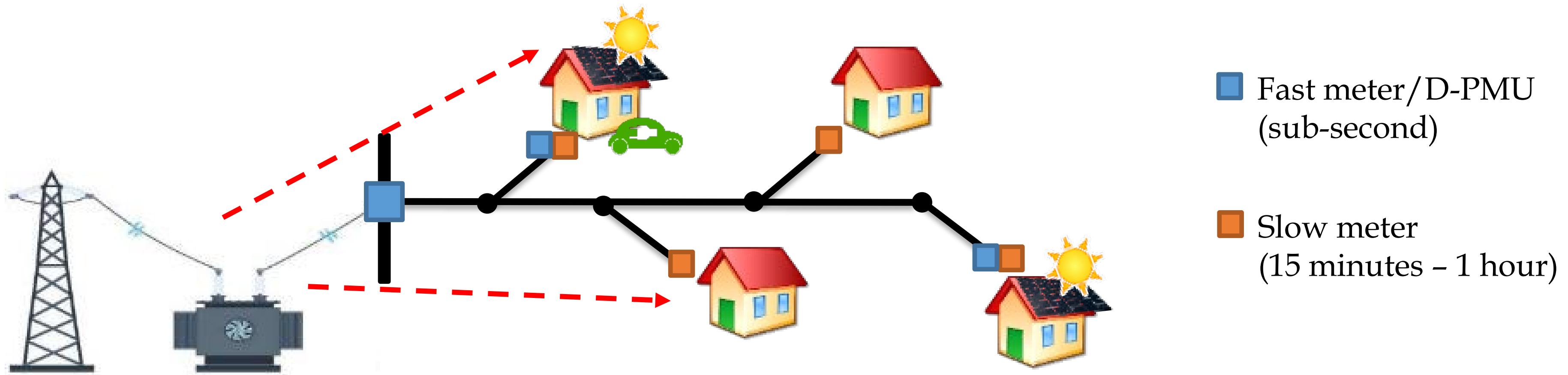


(b) Load center at bus 14

Part II: Risk-aware Learning for Voltage Safety in Distribution Grids

ML for distributed energy resources (DERs)

- Rising DERs at grid edge motivate scalable and efficient coordination to support the operations of connected distribution grids
  - Lack of *frequent, real-time* communications
  - Distribution control center may broadcast messages to every DER



Distribution Substation

Liu, Hao Jan, Wei Shi, and Hao Zhu. "Hybrid voltage control in distribution networks under limited communication rates." *IEEE Transactions on Smart Grid* 10.3 (2018): 2416-2427.
Molzahn, Daniel K., et al. "A survey of distributed optimization and control algorithms for electric power systems." *IEEE Transactions on Smart Grid* 8.6 (2017): 2941-2962.

Prior work

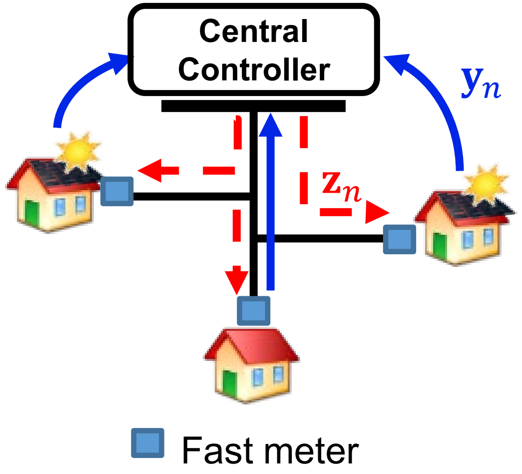
- Scalable DER coordination as an instance of optimal power flow (OPF)
  - Kernel support vector machines [Karagiannopoulos et al' 19] [Jalali et al' 20]
  - Deep neural network for ac-OPF [Zamzam et al' 20][Gupta et al' 21] [Nellikkath et al' 21]
  - Deal with worst-case dc-OPF guarantees by post-analysis [Venzke et al' 20]
  - Reinforcement learning for dynamic coordination [Yang et al' 20] [Cao et al' 21]

- Enforcing network-wide constraints is challenging
  - Project OPF solutions for global learning [Zamzam et al' 20] [Jalali et al' 20]
  - Penalize the constraint violation via regularization [Karagiannopoulos et al' 19] [Pan et al' 19] [Yang et al' 20]
  - Chance-constrained formulation for optimization-and-learn [Gupta et al' 21]

- **Focus:** Use *statistical risks* to improve the safety of DER actions for (network-wide) limits

Centralized DER coordination

- Controllable DER reactive power for voltage optimization



$$\mathbf{z} = \min_{\mathbf{q}\in\mathcal{Q}} \; loss(\mathbf{q})$$
$$\text{s. to} \quad \begin{bmatrix} \mathbf{X}\mathbf{q} + \mathbf{h}(\mathbf{y}) - \overline{\mathbf{v}} \\ -\mathbf{X}\mathbf{q} - \mathbf{h}(\mathbf{y}) + \underline{\mathbf{v}} \end{bmatrix} \leq \mathbf{0}$$

- $\mathcal{Q}$: available reactive power
- $\mathbf{X}$: network matrix
- $\mathbf{y}$: operating condition
- $\underline{\mathbf{v}}, \; \overline{\mathbf{v}}$: voltage limits

- Linearized DistFlow (LDF) model, even for multiphase systems, leads to quadratic programs

- Centralized solutions require high communication rates and communication availability

Scalable design

- Aim to predict from data to the optimal $\Phi(\mathbf{y}) \rightarrow \mathbf{z}$
- *Scalable* neural network (NN) architecture to obtain the mapping for each individual node n

$$\mathbf{y}_n^{t+1} = \sigma(\mathbf{W}_n^t \mathbf{y}_n^t + \mathbf{b}_n^t)$$

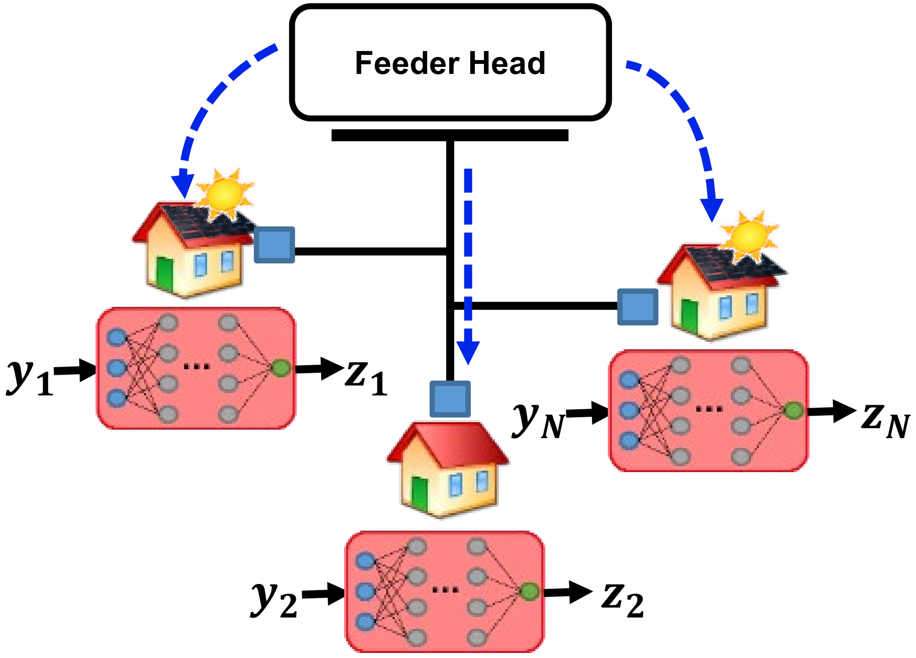
  - $\varphi := \{\mathbf{W}_n^t, \mathbf{b}_n^t\}$: NN parameters

- Convergence analysis in our recent work[Kwon et al'22]

- The average loss under mean squared error (MSE)

$$\min_{\varphi} f(\varphi) := \frac{1}{K} \sum_{k=1}^{K} \ell\left(\Phi(\mathbf{y}_k;\varphi), \mathbf{z}_k\right)$$

with $\ell\left(\Phi(\mathbf{y}_k;\varphi), \mathbf{z}_k\right) = \|\Phi(\mathbf{y}_k;\varphi) - \mathbf{z}_k\|_2^2$ for each sample k

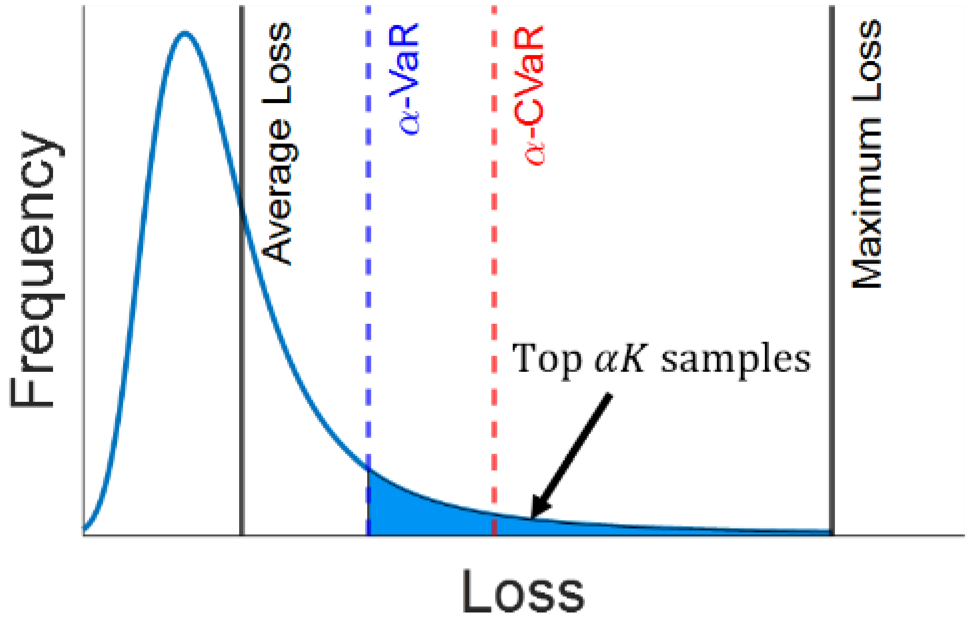


Risk-aware learning

- Conditional value-at-risk (CVaR) metric (empirical approx. of the worst-case mean)

$$\gamma_\alpha(\boldsymbol{\varphi}) := \frac{1}{\alpha K}\sum_{k=1}^{K} \ell(\Phi(\mathbf{y}_k;\boldsymbol{\varphi}),\mathbf{z}_k) \times \mathbb{1}\{\ell(\Phi(\mathbf{y}_k;\boldsymbol{\varphi}),\mathbf{z}_k) \geq v\}$$

  - $\alpha \in (0,1)$: significance level
  - v : α-VaR



- Risk-aware learning going beyond MSE

$$\min_{\boldsymbol{\varphi}} \; f(\boldsymbol{\varphi}) + \lambda\gamma_\alpha(\boldsymbol{\varphi})$$

  - $\lambda > 0$: hyperparameter balances between average and worst-case performances

Shanny Lin, Shaohui Liu, and Hao Zhu. "Risk-Aware Learning for Scalable Voltage Optimization in Distribution Grids," *Power Systems Computation Conference (PSCC) 2022 (accepted)*, https://arxiv.org/abs/2110.01490

Features of CVaR metric

- In addition to $\boldsymbol{q}$, consider *voltage deviation risk* (turns out numerically powerful)

$$\gamma_\alpha^v(\boldsymbol{\varphi}) := \frac{1}{\alpha K}\sum_{k=1}^{K} |v_n(\Phi(\mathbf{y}_k;\boldsymbol{\varphi}))| \times \mathbb{1}\{|v_n(\Phi(\mathbf{y}_k;\boldsymbol{\varphi}))| \geq v\}$$

- CVaR can be recast as a convex problem, using the projection operator $[a]_+ \triangleq \max\{0, a\}$:

$$\gamma_\alpha(\boldsymbol{\varphi}) := \min_{\beta\in\mathbb{R}} \left\{\beta + \frac{1}{\alpha K}\sum_{k=1}^{K}\left[\ell\left(\Phi(\mathbf{y}_k;\boldsymbol{\varphi}), \mathbf{z}_k\right) - \beta\right]_+\right\}$$

  - The optimal β turns out to be the α-VaR value
  - But risk-aware learning is not convex due to nonlinear $\Phi(\cdot\,;\boldsymbol{\varphi})$

- CVaR gradient evaluation can be simplified by replacing $[a]_+$ with soft projection (softplus)

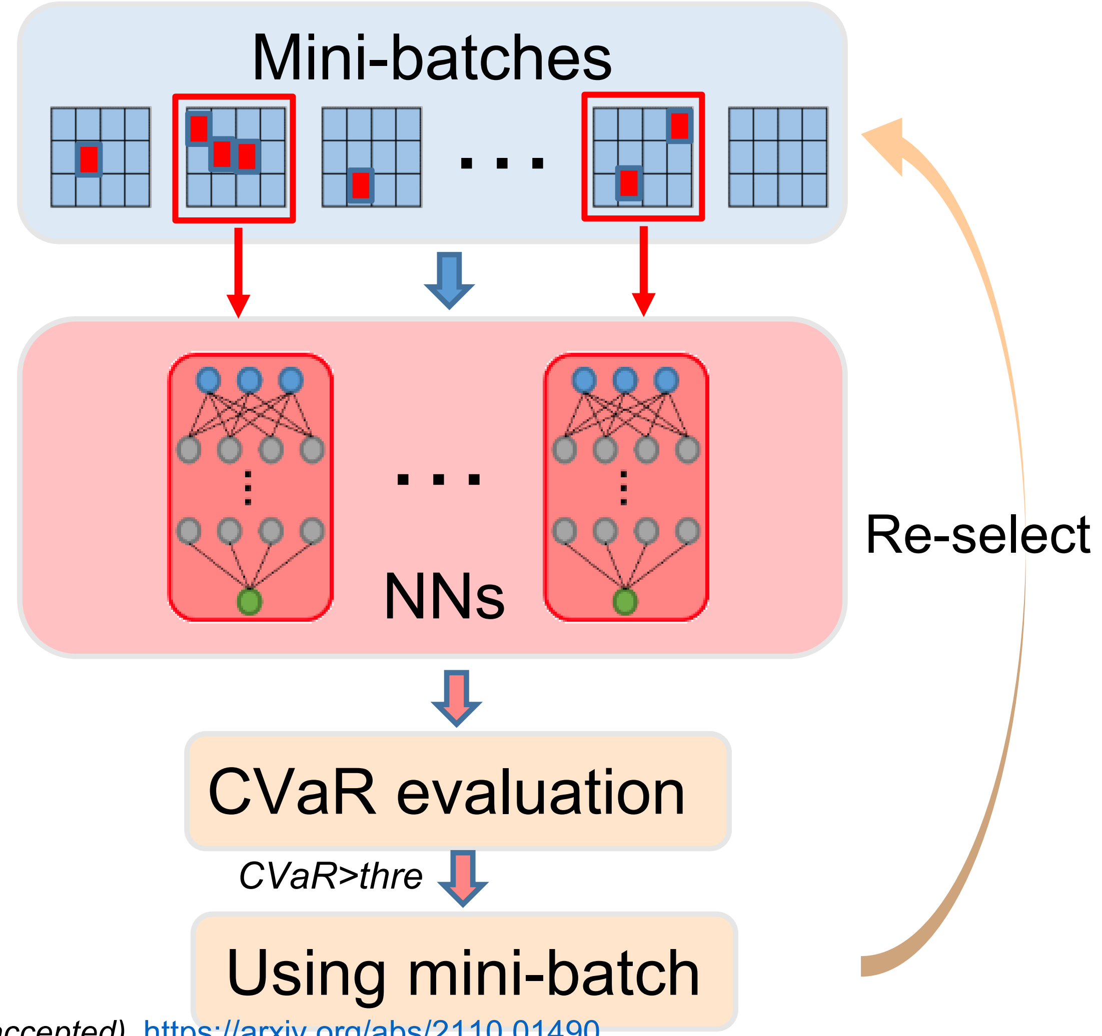
Accelerated CVaR learning

Key challenge: the training efficiency with CVaR is worse than that of average loss

- Reduced sample number affecting the statistical significance of sample-based gradient estimation
- Gradient computation cost increased, as well

➢ Typical NN training uses subset of samples per iteration like the mini-batch method

➢ Accelerating the CVaR training by *selecting mini-batches* with sufficient statistical significance



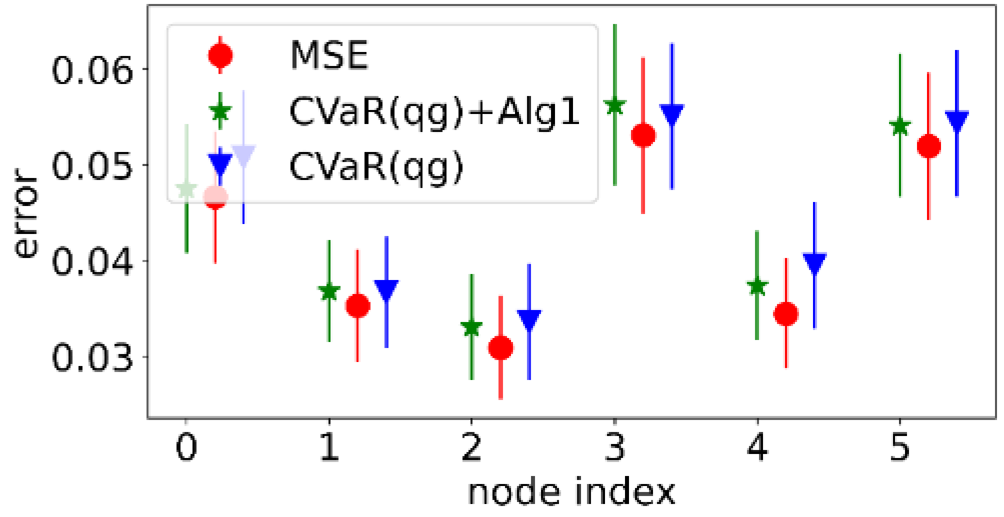
Shanny Lin, Shaohui Liu, and Hao Zhu. "Risk-Aware Learning for Scalable Voltage Optimization in Distribution Grids," *Power Systems Computation Conference (PSCC) 2022 (accepted)*, https://arxiv.org/abs/2110.01490

Predicting reactive power *q*

- IEEE 123-bus test case with 6 DERs, each with its own controllable ***q***
  - Decision rules are learned using local nodal measurements and some feeder head broadcast

- Prediction error is very close among all three approaches due to high prediction accuracy

- Proposed mini-batch selection algorithm (Alg1) reduces training time for CVaR

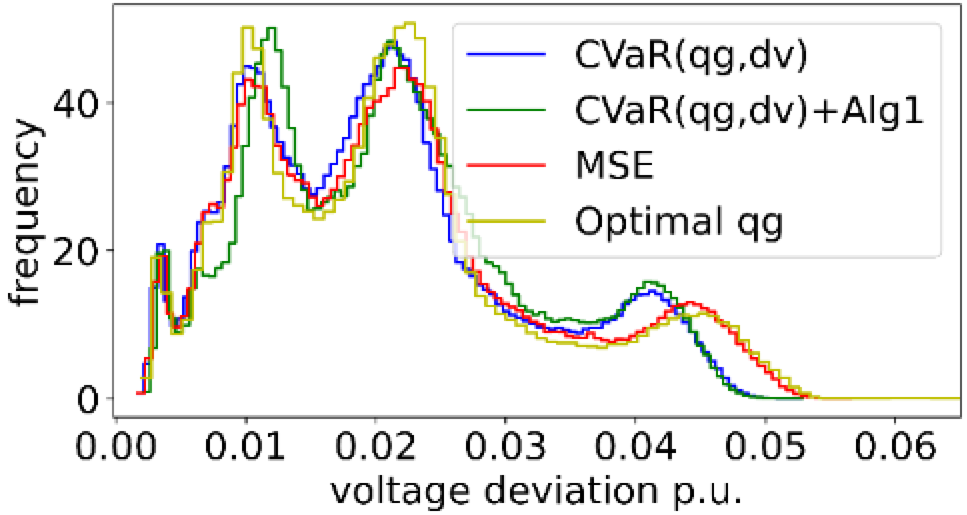


Computational Time

Loss obj.	Epoch [s]	Total [s]
MSE	0.52	46.48
CVaR(qg)	1.07	38.70
CVaR(qg)+Alg 1	0.61	35.63

Reducing voltage deviation risk

- Further incorporated CVaR regularization on voltage deviation error
- CVaR metric can reduces the worst-case voltage deviation and leads to improved system safety
- Training time is accelerated using the proposed mini-batch selection algorithm (Alg1)
  - Evan faster than the *q* prediction CVaR only case

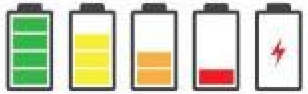


Computational Time

Loss obj.	Epoch [s]	Total [s]
MSE	0.54	44.89
CVaR(qg,dv)	0.77	31.73
CVaR(qg,dv)+Alg 1	0.51	25.93

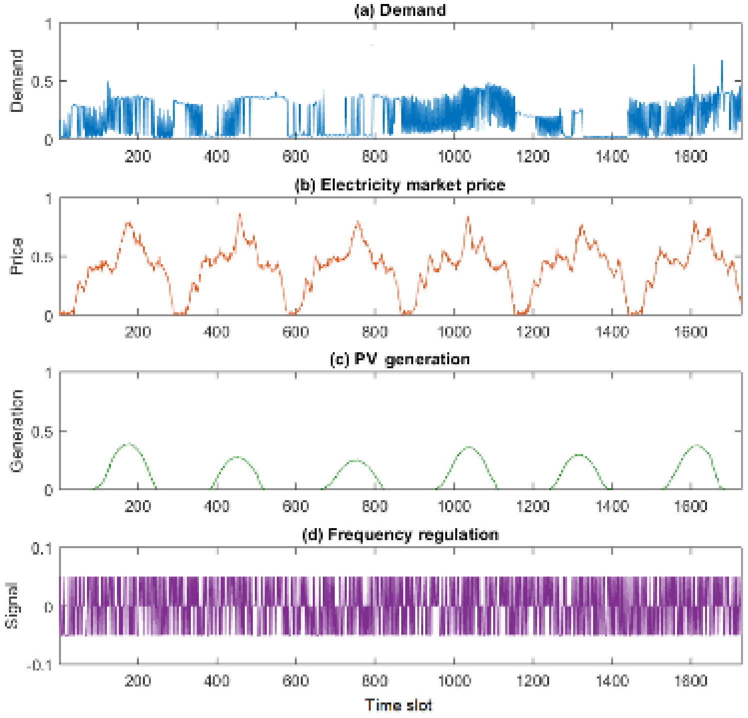
Part III: RL for Dynamical Resources using Efficient Representation

RL for dynamical grid resources



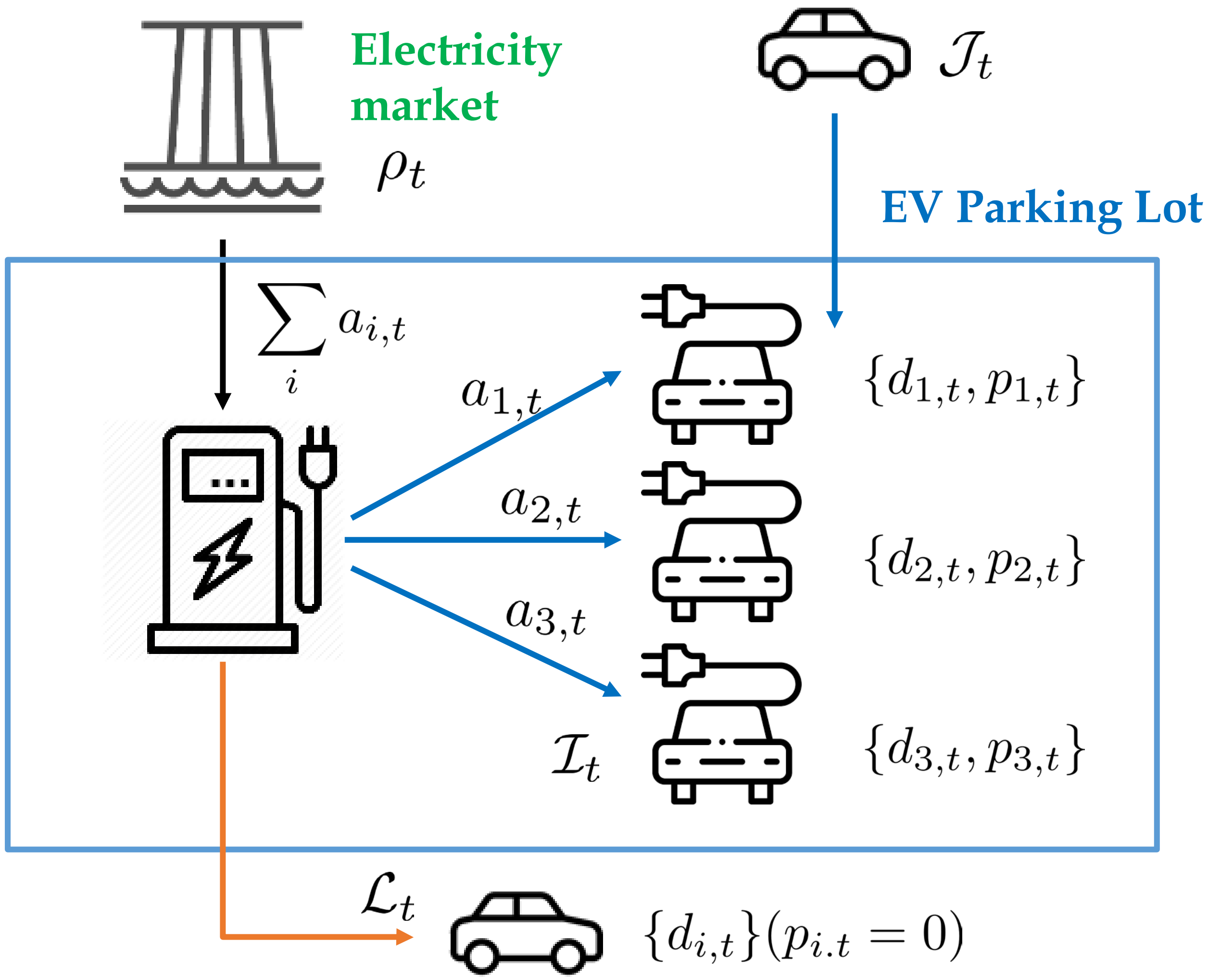
- DERs (energy storage/loads) and external inputs (price/weather) are dynamical
- Motivate a RL approach to learn $a_t \leftarrow \pi(s_t)$
  - data-driven, not requiring the probability
  - adaptive to varying online conditions

Key Challenge: *abundant, heterogenous* resources at grid edge need powerful state/action *representation*



Chen, Xin, Guannan Qu, Yujie Tang, Steven Low, and Na Li. "Reinforcement learning for decision-making and control in power systems: Tutorial, review, and vision." https://arxiv.org/abs/2102.01168

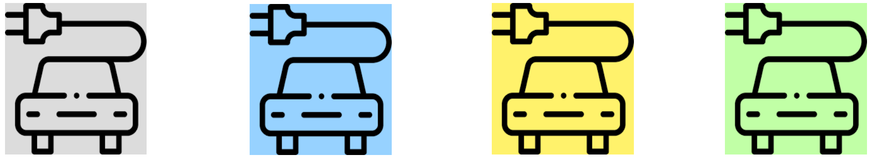
Electrical vehicle charging station (EVCS) problem



- Arriving EV i with demand $d_{i,t}$ and parking time $p_{i,t}$
- EVCS decides which EVs to charge ($a_{i,t} = 1$) from electricity at price ρ_t
- Clearly, the state/action space *incurs high complexity* due to large, *time-varying* dimensionality
- *How to represent state/action to allow for efficient RL training?*

An aggregation scheme

$p_{i,t}$	5	2	3	4
$d_{i,t}$	3	2	2	3
$\ell_{i,t}$	2	0	1	1



<Original State>

EV i with $d_{i,t}, p_{i,t}$

Laxity (priority): $\ell_{i,t} = p_{i,t} - d_{i,t}$

<Aggregated State>

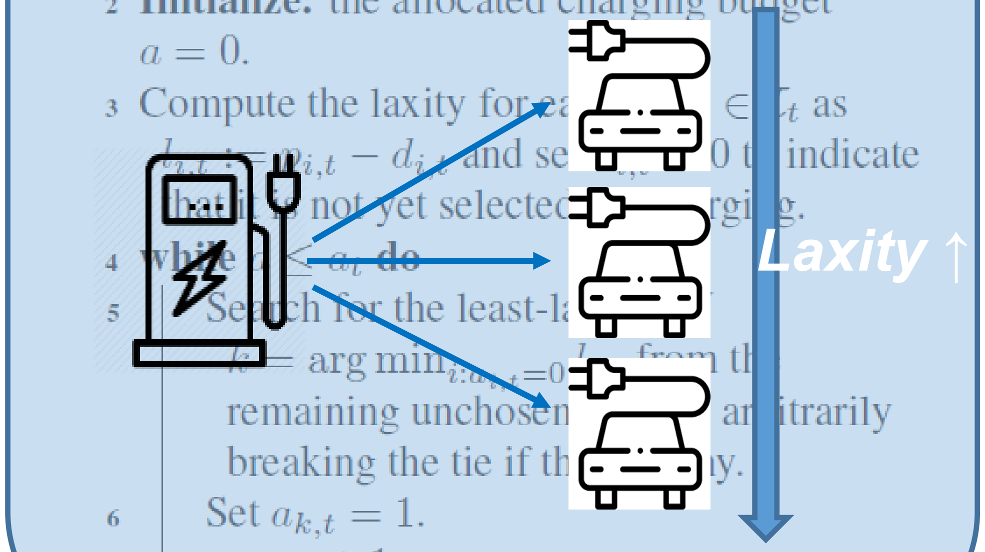
$n_t^{(0)}$	$n_t^{(1)}$	$n_t^{(2)}$	$n_t^{(3)}$
1	2	1	0

State representing the number of EVs with the same laxity (max L)

$$s'_t = [\rho_t, n_t^{(0)}, n_t^{(1)}, \cdots, n_t^{(L)}]$$

Kyung-Bin Kwon and H. Zhu, "Efficient representation for electric vehicle charging station operations using reinforcement learning," HICCS 2022 https://arxiv.org/abs/2108.03236

Laxity-based action reduction



Algorithm 1: Least-laxity first (LLF) rule

1 **Inputs:** Total charging power a_t, the set of EVs in $\mathcal{I}_t$ along with their remaining demand $d_{i,t}$ and parking time $p_{i,t}$.

2 **Initialize:** the allocated charging budget $a = 0$.

3 Compute the laxity for ea ... $\in \mathcal{I}_t$ as ... $p_{i,t} - d_{i,t}$ and se ... 0 t ... indicate ... not yet selecte ... rgi ... g.

4 **while** ... $< a_t$ **do**

5 Se ... ch for the least-l ... arg min ... from th ... remaining unchosen ... arbitrarily breaking the tie if th ... y.

6 Set $a_{k,t} = 1$.

7 $a \leftarrow a + 1$

8 **end**

Prop 1: *If the EVCS total charging schedule $\{a_t\}_{t \in T}$ is feasible (corresponds to some feasible schedule for individual EVs that ensure all fully charged before departure), then Algorithm 1 can produce such a feasible schedule for all EVs.*

- Basically, LLF ensures the **feasibility** of the recovered actions

Proof idea: Any feasible schedule equivalent to one satisfying LLF [Wang et al'21]

Equivalence of state aggregation

- Ideally, we want the new state represents the same MDP
- This equivalence requires two conditions:
  - (i) Reward homogeneity: same reward for any states aggregated into the same new state
  - (ii) Dynamic Homogeneity: same transition kernel for any aggregated states

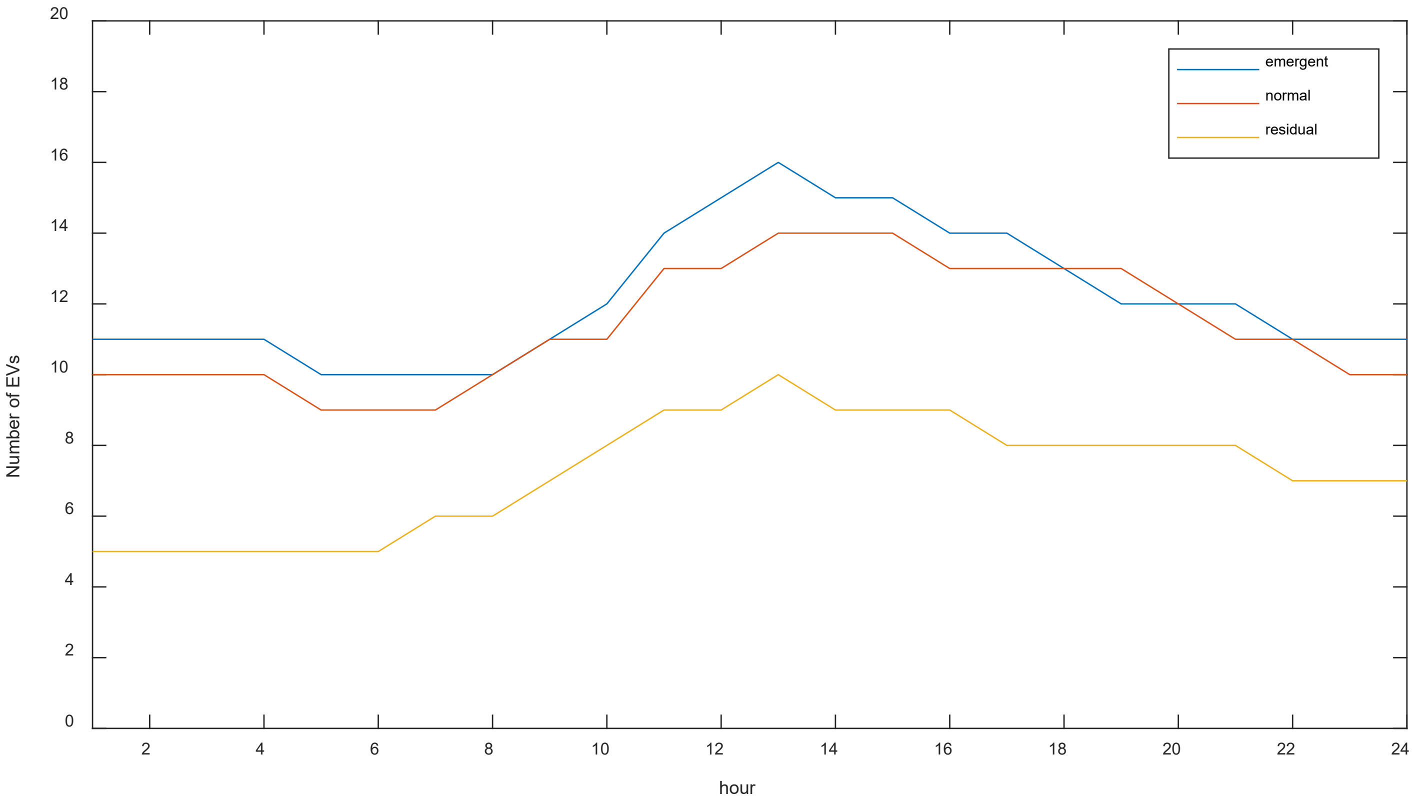
Prop 2: *The original MDP for s_t/a_t is equivalent to the new one for s_t'/a_t using the total charging action. Accordingly, the optimal policy (or action) obtained from the new MDP through aggregation are equivalent to that for the original one.*

Intuitions for dyn. homogeneity:

Under the LLF rule, charge either one of 2 EVs at the same laxity leads to the same transition of new state or aggregated state

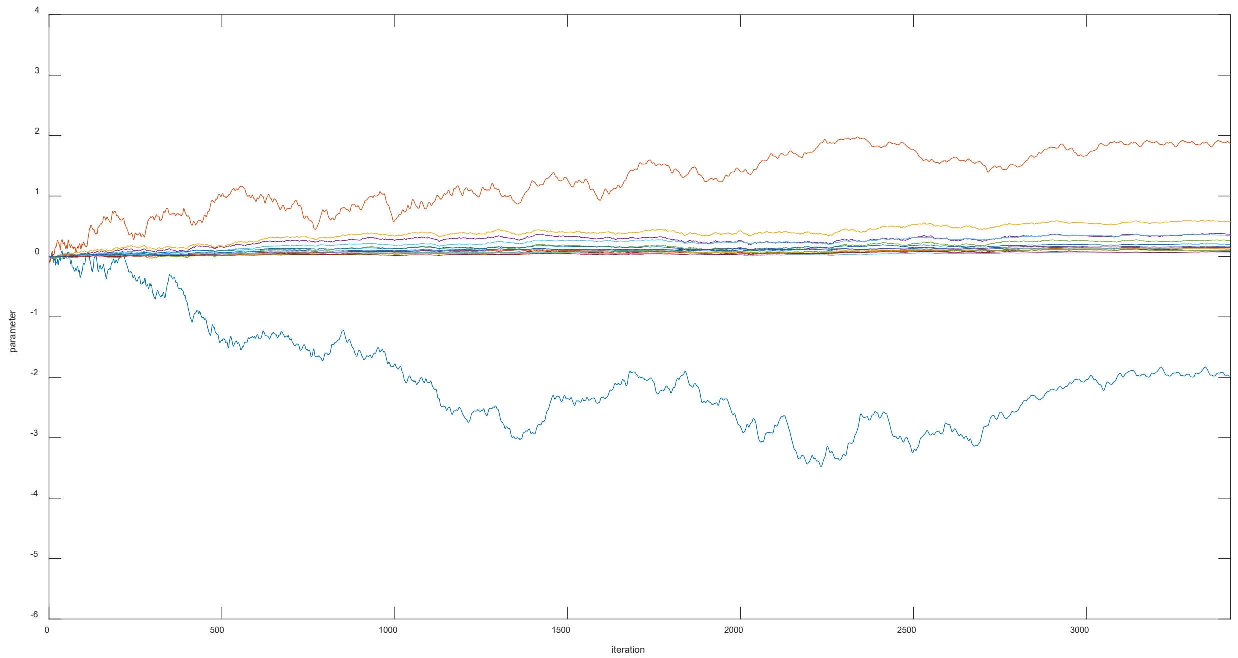
Kyung-Bin Kwon and H. Zhu, "Efficient representation for electric vehicle charging station operations using reinforcement learning," HICCS 2022
https://arxiv.org/abs/2108.03236

Numerical tests



- Daily charging at 15-min intervals ($T = 96$)
  - Realistic EV arrival model
  - ERCOT real-time price
- 20 daily scenarios for training; 5 for testing
- Comparing proposed Alg 2 with Alg. QE in [Wang et al'21] by approximating the Q-function

Convergence => further aggregation



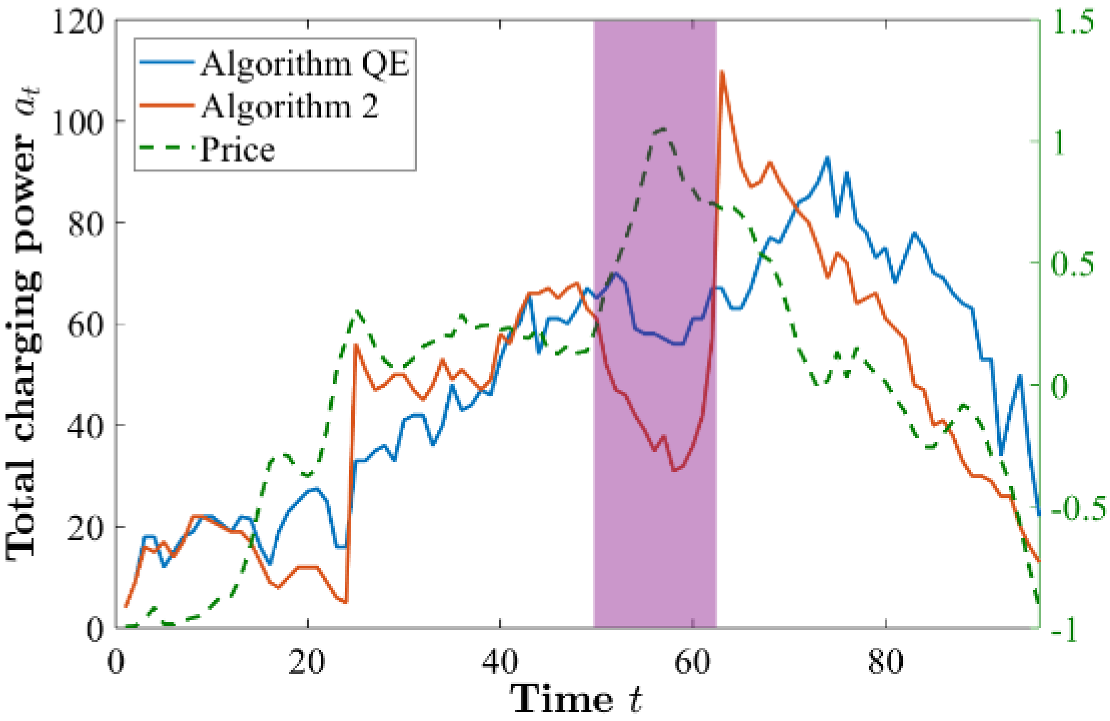
ρ_t	$n_t^{(0)}$	$n_t^{(1)}$	$n_t^{(2)}$	$n_t^{(3)}$	$n_t^{(4)}$	$n_t^{(5)}$
-1.9735	**1.8628**	0.5772	0.3674	0.2651	0.3485	0.1191
$n_t^{(6)}$	$n_t^{(7)}$	$n_t^{(8)}$	$n_t^{(9)}$	$n_t^{(10)}$	$n_t^{(11)}$	$n_t^{(12)}$
0.2021	0.1404	01386	0.1592	0.0975	0.0693	0.0797

- Parameter convergence
- Most weight parameters are very small; except for ρ_t and $n_t^{(0)}$
- **Remark:** we can further reduce # of states by grouping higher-laxity EVs!
- Possible to consider nonlinear policies as well

Case study – testing result

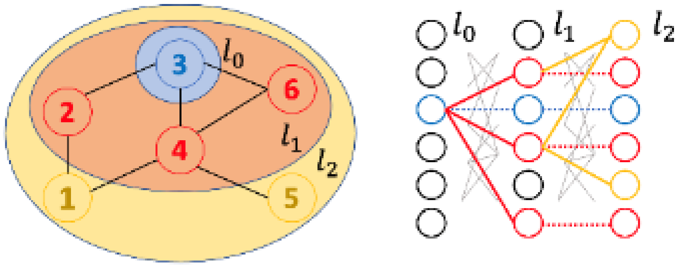
	Test 1	Test 2	Test 3	Test 4	Test 5	Average
Alg. 2	-5016.2	-5022.6	-5009.5	-5012.8	-5007.8	-5013.8
Alg. QE	-5240.1	-5240.3	-5234.2	-5239.3	-5230.6	-5236.9
Increase (%)	4.27	4.15	4.29	4.32	4.26	4.26

Algorithm QE: Estimating an approximate Q-function [Wang et al'21]

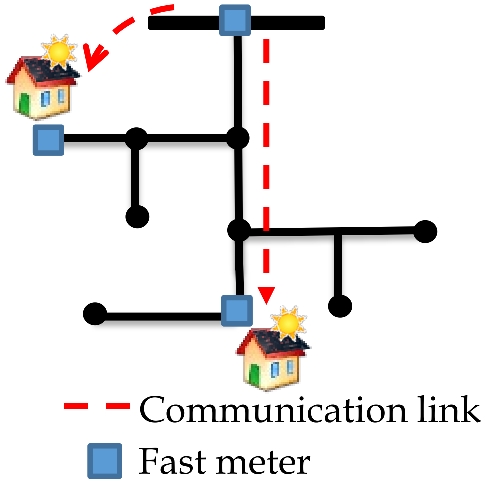


- Alg. 2 improves the reward of Alg. QE [Wang et al'21] by ~4.2%
- Example charging profile indicates Alg. 2 very sensitive to price peaks and strategically reducing a_t, while Alg. QE fails to do so

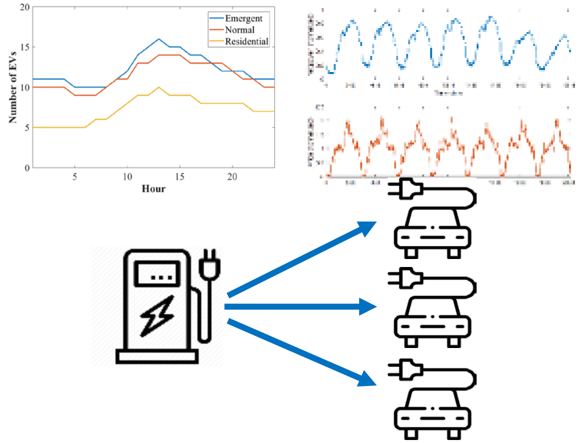
Summary



Topology-aware learning in large-scale power systems:
Simpler model structure



Risk-aware learning for grid-edge coordination:
Reduced risks of voltage violations



Reinforcement learning for dynamical resources:
More efficient representation

- I: Generalized transfer capability in graph-based learning
- II: Convergence analysis and strengthened safety guarantees
- III: Comprehensive grid-edge resource coordination

Related Work

- Shaohui Liu, Chengyang Wu, and Hao Zhu. "Topology-aware Graph Neural Networks for Learning Feasible and Adaptive AC-OPF Solutions," submitted. https://arxiv.org/pdf/2205.10129
- Kyung-bin Kwon and Hao Zhu, "Reinforcement Learning Based Optimal Battery Control Under Cycle-based Degradation Cost," *IEEE Trans. Smart Grid (accepted)*, https://ieeexplore.ieee.org/abstract/document/9789478
- Kyung-bin Kwon, Lintao Ye, Vijay Gupta, and Hao Zhu, "Model-free Learning for Risk-constrained Linear Quadratic Regulator with Structured Feedback in Networked Systems," submitted https://arxiv.org/abs/2204.01779
- M. Jalali, V. Kekatos, S. Bhela, and H. Zhu, "Inferring power system dynamics from synchrophasor data using Gaussian processes," *IEEE Trans. Power Systems*, https://ieeexplore.ieee.org/abstract/document/9693288
- Yuqi Zhou, Jeehyun Park, and Hao Zhu, "Scalable Learning for Optimal Load Shedding Under Power Grid Emergency Operations," *PES General Meeting (PESGM)* 2022 (accepted) https://arxiv.org/abs/2111.11980
- Shanny Lin, Shaohui Liu, and Hao Zhu. "Risk-Aware Learning for Scalable Voltage Optimization in Distribution Grids," *Power Systems Computation Conference (PSCC) 2022 (accepted)*, https://arxiv.org/abs/2110.01490
- S. Lin and H. Zhu, "Enhancing the Spatio-temporal Observability of Grid-Edge Resources in Distribution Grids," *IEEE Trans. Smart Grid*, 2021. DOI: 10.1109/TSG.2021.3107239
- Kyung-Bin Kwon and H. Zhu, "Efficient representation for electric vehicle charging station operations using reinforcement learning," HICCS 2022 https://arxiv.org/abs/2108.03236
- Liu, Shaohui, Chengyang Wu, and Hao Zhu. "Graph Neural Networks for Learning Real-Time Prices in Electricity Market." *ICML Workshop on Tackling Climate Change with Machine Learning*, 2021. https://arxiv.org/abs/2106.10529

Learning and Optimization for Smarter Electricity Infrastructure

Hao Zhu
haozhu@utexas.edu
http://sites.utexas.edu/haozhu/
@HaoZhu6

Learning for grid resilience

Learning for dynamical resources

Learning for inverter-based resources

….

Thank you!