A Learning-to-Infer Method for Real-Time Power System Monitoring

Yue Zhao
Stony Brook University

Joint work with Jianshu Chen, Seung-jun Kim, Younghwan Lee, Jiaming Li, and H. Vincent Poor

BigData Tutorial
3/26/2020

Y. Zhao and B. Zhang, *"Deep Learning in Power Systems"*, in *Advanced Data Analytics in Power Systems,* Cambridge University Press, forthcoming.

Motivating Questions in a Bigger Picture

- Deep learning (DL) has seen tremendous recent successes in many areas of artificial intelligence. It has since sparked great interests in its potential use in power systems.

- Potential applications in power system operations
  - Forecasting?
  - Monitoring?
  - Optimization?
  - Control?

Issues with using DL in Power Systems

- Predictability (regardless of complexity)
  - DL has been very successful in highly predictable situations.
    - Image, speech, …
  - Much more challenging for fundamentally unpredictable situations.
    - Stock market, ...
    - Forecasting?
  - Availability of input information/signals

- Complexity (assuming good predictability)
  - DL appears to be very powerful in solving problems of much higher complexity than before.
    - Playing games (Deep RL)
    - **Monitoring,** optimization, and control?
  - Availability of data and labels

Issues with using DL in Power Systems

- Predictability (regardless of complexity)
  - DL has been very successful in highly predictable situations.
    - Image, speech, …
  - Much more challenging for fundamentally unpredictable situations.
    - Stock market, ...
    - Forecasting?
  - Availability of input information/signals

- Complexity (assuming good predictability)
  - DL appears to be very powerful in solving problems of much higher complexity than before.
    - Playing games (Deep RL)
    - **Monitoring**, optimization, and control?
  - Availability of data and labels

Real Time Power System Monitoring: Motivation

- Our power system's efficient, reliable and secure operation is crucially dependent on effective monitoring of the system.

- The real-time information on the current grid status is not only key for optimizing resources to economically maintain normal system operation, but also crucial for identifying current and potential problems that may lead to blackouts.

- With increasing penetration of renewables, EVs, and other DERs, power systems become even more dynamic. Faster power system monitoring that offers more accurate actionable information is needed.

- This tutorial focuses on steady state.

General Formulation

Noiseless measurements

AC PF model

- Observation model

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{s}(t); \boldsymbol{\alpha}) + \boldsymbol{v}(t)$$

noise

Nodal complex voltages

Component statuses

Component parameters

- Desired function $f(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{y}; \boldsymbol{\alpha})$
  - *Situational Awareness*
    - State estimation: $f(\cdot) = \boldsymbol{x}$
    - Outage detection: $f(\cdot) = \boldsymbol{s}$
  - *Preventive Analysis*
    - Voltage stability analysis: $f(\cdot)$ is the voltage stability margin.
    - Contingency analysis: $f(\cdot)$ is whether the system is N-k secure.
- (Approximately) compute $f(\cdot)$ based on ***y*** --- *Inference.*

The Role of Physical Model

- **Data-driven vs. Model-driven**
  - Limitations of data-driven method

    Insufficient real-world data, lack of labels, dealing with “rare events”.

  - Model-driven methods have been mainstream in power systems.

    Information embedded in the physical models is absolutely crucial.

    However, challenges arise for model-driven approaches in providing effective solutions in fundamentally hard problems.

Physical Model based Inference

- State estimation $\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{s})\|^2$
- Joint component outage detection and state estimation $\min_{\boldsymbol{x},\boldsymbol{s}} \|\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{s})\|^2$
- Voltage stability margin estimation
  – Heuristic: Compute the smallest eigenvalue of the Jacobian from solving the power flow equations
- Checking *N* - *k* security
  – Check if the system can continue to operate under every combination of *k*-component outages of interests, by re-solving the power flow equations (or OPF for corrective contingency analysis).

Limitations – Computational Complexity in Real Time

- Non-convexity

- Combinatorial, Fundamentally hard

The Role of Physical Model (cont.)

- **Data-driven + Model-driven**
  - Power system has *very clearly understood physical models* (e.g., the AC power flow model).
  - The quantities that we measure in a power system all follow these physical laws.
  - The information embedded in the physical models is absolutely crucial.
  - How do we maximumly incorporate the information in a physical model in a data-driven approach?

Machine Learning based Monitoring

- Generative model $p(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}) = p(\boldsymbol{x}, \boldsymbol{s}) p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{s})$

  Physical model

- Inference based on ***y***
  - Find a predictor function $F(\boldsymbol{y})$ which, given inputs ***y***, outputs values that are closest to $f(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y})$.

$$\min_{F \in \mathcal{F}} \mathbb{E}_{\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}} \left[L\left(f(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}), F(\boldsymbol{y}) \right) \right]$$

- Target function (assuming no constraints on *F*): $F^*(\boldsymbol{y})$

  [*"Elements of Statistical Learning"*, Hastie, Tibshirani & Friedman 09]

  - E.g., if $L(f, F) = (f - F)^2$, then $F^*(\boldsymbol{y}) = \mathbb{E}_{\boldsymbol{x}, \boldsymbol{s}|\boldsymbol{y}} \left[f(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}) | \boldsymbol{y} \right]$

$F^*(\boldsymbol{y})$ is often computationally intractable to evaluate, or even just to express.

A Learning-to-Infer Method

- We would like to find a predictor function $F(\boldsymbol{y})$ so that
  - $F(\boldsymbol{y})$ matches the target function $F^*(\boldsymbol{y})$ as closely as possible, and
  - $F(\boldsymbol{y})$ takes a form that allows fast evaluation of its value.

$$\min_{F \in \mathcal{F}} \mathbb{E}_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}} \left[L \left(f(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}), F(\boldsymbol{y}) \right) \right]$$

$$\Leftrightarrow \min_{\boldsymbol{\beta}} \mathbb{E}_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}} \left[L \left(f(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}), F_{\boldsymbol{\beta}}(\boldsymbol{y}) \right) \right]$$

A Learning-to-Infer Method

- We would like to find a predictor function $F(\boldsymbol{y})$ so that
  - $F(\boldsymbol{y})$ matches the target function $F^*(\boldsymbol{y})$ as closely as possible, and
  - $F(\boldsymbol{y})$ takes a form that allows fast evaluation of its value.

$$\min_{F \in \mathcal{F}} \mathbb{E}_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}} \left[L\left(f(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}), F(\boldsymbol{y}) \right) \right]$$

$$\Leftrightarrow \min_{\boldsymbol{\beta}} \mathbb{E}_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}} \left[L\left(f(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}), F_{\boldsymbol{\beta}}(\boldsymbol{y}) \right) \right] \approx \min_{\boldsymbol{\beta}} \frac{1}{I} \sum_{i=1}^{I} L\left(f(\boldsymbol{x}^i, \boldsymbol{s}^i, \boldsymbol{y}^i), F_{\boldsymbol{\beta}}(\boldsymbol{y}^i) \right)$$

Monte Carlo samples drawn from $p(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}) = p(\boldsymbol{x}, \boldsymbol{s}) p(\boldsymbol{y} | \boldsymbol{x}, \boldsymbol{s})$

A Learning-to-Infer Method

- We would like to find a predictor function $F(\boldsymbol{y})$ so that
  - $F(\boldsymbol{y})$ matches the target function $F^*(\boldsymbol{y})$ as closely as possible, and
  - $F(\boldsymbol{y})$ takes a form that allows fast evaluation of its value.

$$\min_{F \in \mathcal{F}} \mathbb{E}_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}} \left[L\left(f(\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}), F(\boldsymbol{y}) \right) \right]$$

$$\Leftrightarrow \min_{\boldsymbol{\beta}} \mathbb{E}_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}} \left[L\left(f(\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}), F_{\boldsymbol{\beta}}(\boldsymbol{y}) \right) \right] \approx \min_{\boldsymbol{\beta}} \frac{1}{I} \sum_{i=1}^{I} L\left(f(\boldsymbol{x}^i,\boldsymbol{s}^i,\boldsymbol{y}^i), F_{\boldsymbol{\beta}}(\boldsymbol{y}^i) \right)$$

Monte Carlo samples drawn from $p(\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}) = p(\boldsymbol{x},\boldsymbol{s})p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s})$

- *An empirical risk minimization problem: Training a discriminative model* $F_{\boldsymbol{\beta}}(\boldsymbol{y})$ *with MC samples from a generative model.*

General Method

Table 1.1 The Learning-to-Infer Method

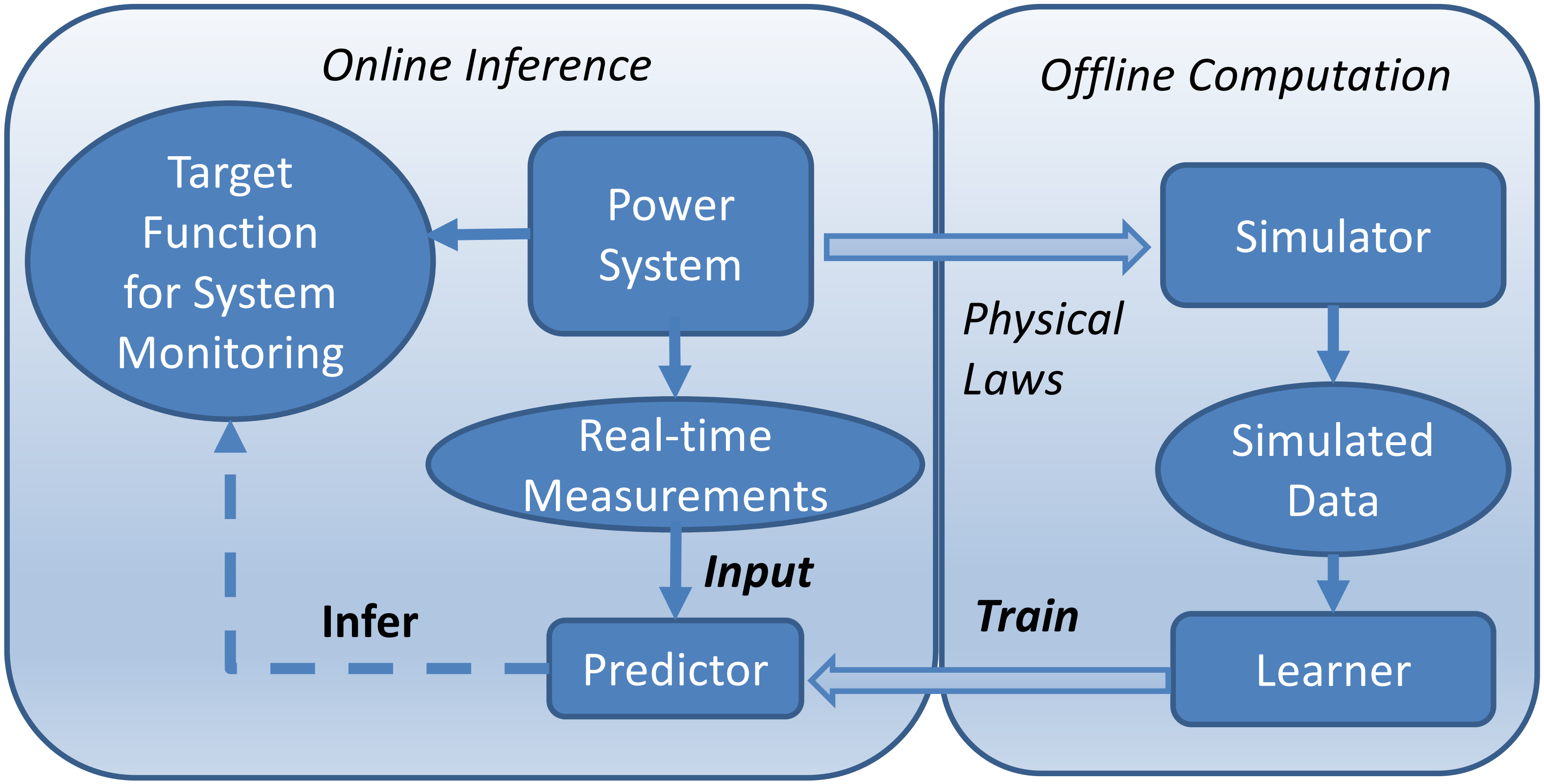
Offline computation:

1. Generate a data set $\{\boldsymbol{x}^i, \boldsymbol{s}^i, \boldsymbol{y}^i\}$ using Monte Carlo simulations with the power flow and sensor models.
2. Select a parametrized function class $\mathcal{F} = \{F_{\boldsymbol{\beta}}(\boldsymbol{y})\}$.
3. Train the function parameters $\boldsymbol{\beta}$ with $\{\boldsymbol{x}^i, \boldsymbol{s}^i, \boldsymbol{y}^i\}$.

Online inference (in real time):

1. Collect instant measurements $\boldsymbol{y}$ from the system.
2. Compute the prediction $F_{\boldsymbol{\beta}^*}(\boldsymbol{y})$.

General Method



Key Advantages

- Labeled data can be generated in an arbitrarily large amount, often efficiently.

  Many types of labels are built-in.

- Very complex predictor models can be trained to ensure good approximation of $F^*(\boldsymbol{y})$.

  Overfitting is much less of an issue as additional data can always be generated from the physical model.

- Offline computation can be maximumly exploited to offer the best real-time inference performance.
- Information from complex physical models, represented by the simulated data, are seamlessly integrated with that from real-time sensor measurements.
- Accelerate key procedures in power system operation, e.g., contingency analysis, security constrained OPF.

Case Study 1: Multi-Line Outage Identification

- Motivation
  - Missing/incorrect information of multi-component failures is a major cause of large-scale blackouts in power systems.
  - E.g., cascading failures can develop very quickly in wide-area power networks.

    Case: The cascading failures in the 2011 Southwest blackout caused 7 million people out of power in 11 minutes.
  - The depth of cascading failures can be much beyond "N-1".
  - System operators need ***real-time and accurate*** information about complex failure scenarios to effectively contain failures.
  - At any time instant, given all the available measurements ***y***, how do we ***infer*** the current topology/failure scenario ***s***?
- One big hypothesis testing problem

  The complexity grows exponentially with the number of unsure line statuses.

Related work

- *Real-time* line outage identification (use instant measurements)
  - Exhaustive search based: [Tate & Overbye 08], [Zhao et al. 14], [Garcia et al. 16]
  - Exploiting sparsity: [Zhu & Giannakis 12]
  - Graphical model based: [Chen et al. 14]
  - Sequential detection: [Heydari & Tajer 17]

- *Non-real-time* topology identification (collect data over a certain period)
  - [He & Zhang 11] [Bolognani et al. 13] [Li et al. 13] [Yuan et al. 16] [Kekatos et al. 16] [Gera et al. 17] [Weng et al. 17] [Deka et al. 17]

Challenges

- **Complexity:** The complexity grows exponentially with the number of unsure line statuses.
- Real-time requirement: very low complexity algorithm needed.
- Use only instant measurements.
- High accuracy.

Learning to Infer

- A Probabilistic formulation

  Optimal inference of grid topology ***s*** depends on $p(\boldsymbol{s}|\boldsymbol{y}), \forall s$.

  – Even listing them has an exponential complexity.

- We focus on *marginal* inference, i.e., compute $p(s_l|\boldsymbol{y})$.

  – Still, summing out all $s_k, k \neq l$ is exponentially complex.

  – The MAP detector $\text{argmax}_{s_l \in \{0,1\}} p(s_l|\boldsymbol{y})$ has a very complicated decision boundary.

- Learning to Infer:

  – **Desired function** $f(\cdot) = p(s_l|\boldsymbol{y}), \forall l$

  – Approximate, by offline training, the posterior $p(s_l|\boldsymbol{y})$ with functions that enable real-time online inference.

Learning to infer

- Goal: Find a variational distribution $q(\boldsymbol{s}|\boldsymbol{y})$ to approximate $p(\boldsymbol{s}|\boldsymbol{y})$, s.t.
  - The model of $q(\boldsymbol{s}|\boldsymbol{y})$ has *sufficient expressive power* to closely represent complicated $p(\boldsymbol{s}|\boldsymbol{y})$.
  - Given $\boldsymbol{y}$, $q(s_l|\boldsymbol{y})$ can be easily computed for real-time inference.
- Optimize the variational distribution **via empirical risk minimization**:

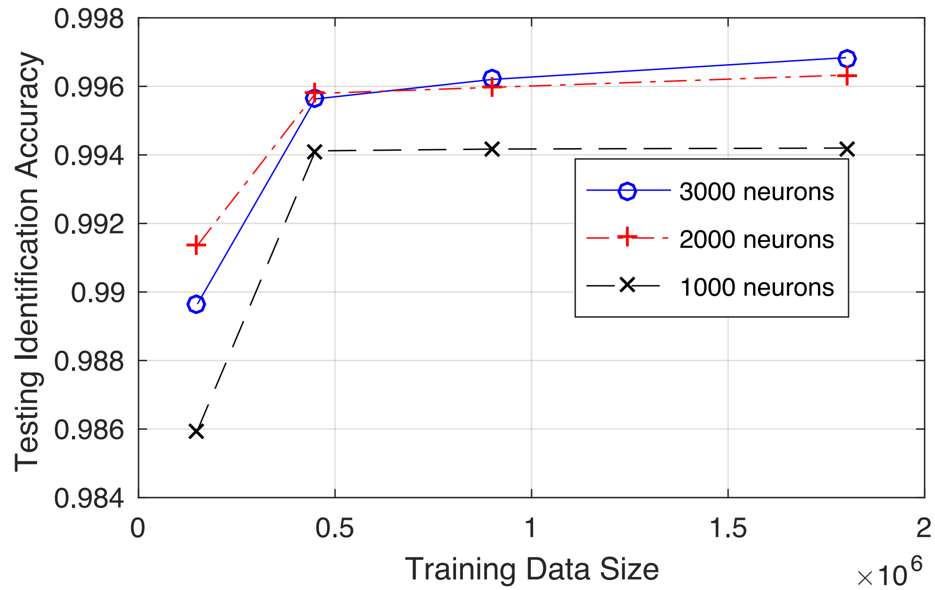
$$\min_{\boldsymbol{\beta}} \ \mathbb{E}_{\boldsymbol{y}}\left[D(p\|q_{\boldsymbol{\beta}})\right] \Leftrightarrow \ \max_{\boldsymbol{\beta}} \ \mathbb{E}_{\boldsymbol{s},\boldsymbol{y}}\left[\log q_{\boldsymbol{\beta}}(\boldsymbol{s}|\boldsymbol{y})\right] \approx \max_{\boldsymbol{\beta}} \frac{1}{I}\sum_{i=1}^{I} \log q_{\boldsymbol{\beta}}(\boldsymbol{s}^i|\boldsymbol{y}^i)]$$

MC samples drawn from the generative model.

Y. Zhao, J. Chen and H. V. Poor, "A Learning-to-Infer Method for Real-Time Power Grid Multi-Line Outage Identification," in IEEE Transactions on Smart Grid, vol. 11, no. 1, pp. 555-564, Jan. 2020.

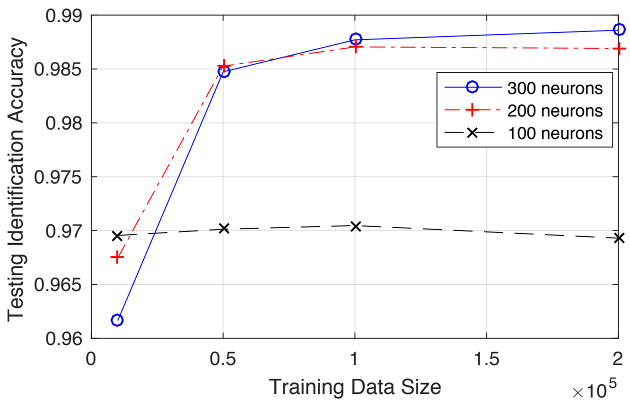
Numerical experiments

- IEEE 300 bus
  ~10^21 candidate topologies
- Non-sparse outages
  Average cardinality 11.6
- Highly variable operating conditions
- Noisy measurements of voltage phase angles at all the buses

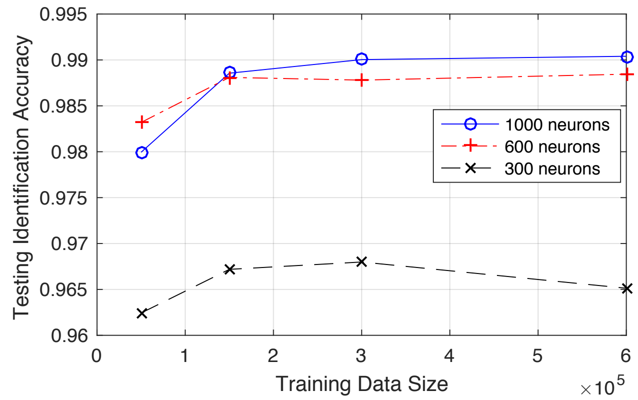


- Neural networks with shared features are used as the predictor models.
- **1.8M** data for training are sufficient.
- 200K testing data contains **100% unseen** topologies from the training data.
- Training time: 3 hours on a GPU. Testing time: **under a millisecond per data sample**.
- Results: the accuracy is **0.997** --- on average **1.0** misidentified line status label
  - Great generalizability.

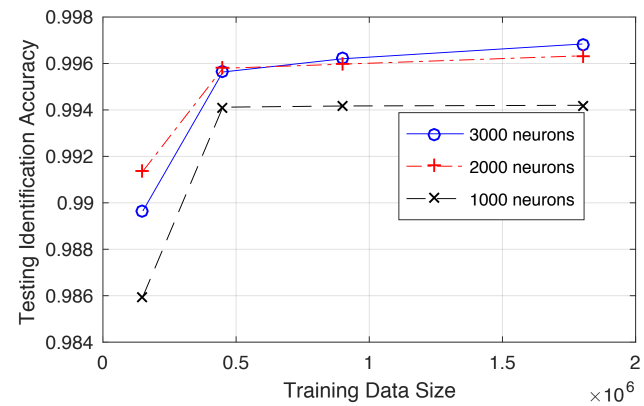
Numerical experiments (cont.)



IEEE 30-bus



IEEE 118-bus



IEEE 300-bus

- Scalability

  The required training data size increases approximately *linearly* with the problem size.

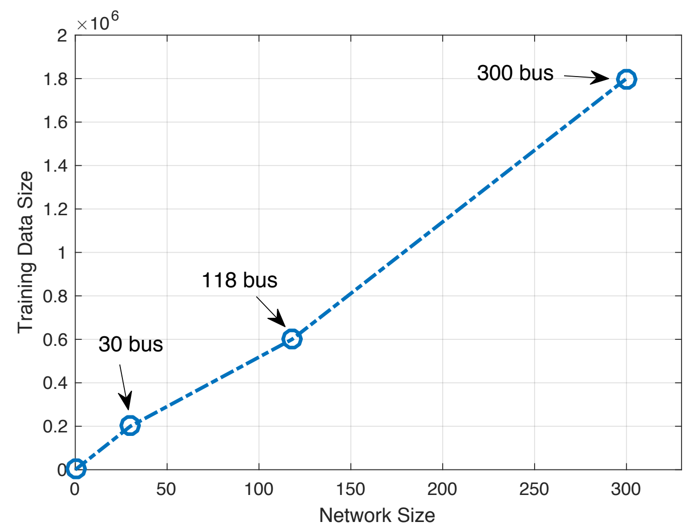


Figure 5: Scalability of the Learning-to-Infer method, from the IEEE 30 bus system to the IEEE 300 bus system.

Numerical experiments (cont.)

- IEEE 118-bus system with AC power flow model

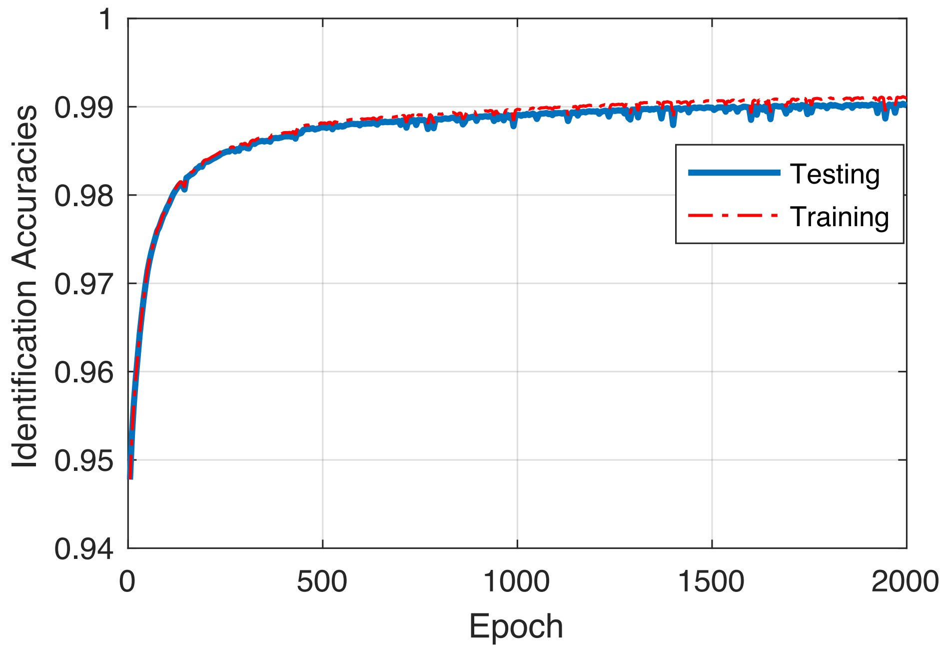


Figure 7: Progressions of the training and testing accuracies, the IEEE 118-bus system, with the AC power flow model employed.

Case Study 2: Voltage Stability Margin Estimation

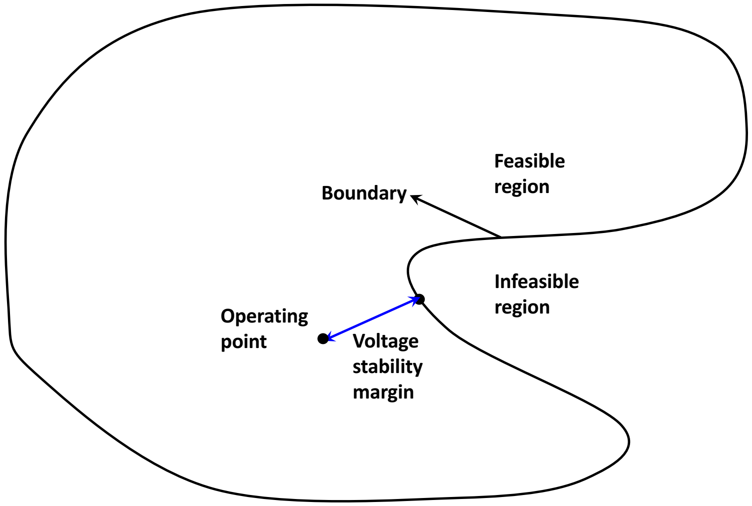
- Motivation
  - Voltage collapse is one of the major causes of large-scale black-outs [Kundur et al. 04].
  - Increasing penetration of renewables brings higher variabilities into power system operations.
  - Determining the system stability margin in real time is greatly valuable for system operators to maintain situational awareness of the system and a safe operating margin.
  - We study voltage stability from a static analysis perspective.
    - Bifurcation [Ajjarapu & Lee 92].
    - ACPF's Jacobian becomes singular.

Problem

- Goal: Given any stable operating condition, estimate its **"distance"** to voltage instability in real time.
- Stability region ***C***: all the power profiles ***S*** that do not induce voltage instability.
- Voltage stability margin of any ***S***: the (Euclidean) distance from ***S*** to the instability region ***C^c***:

$$\mathrm{dist}\left(\boldsymbol{S}, \mathcal{C}^c\right) \triangleq \min_{\boldsymbol{S}' \in \mathcal{C}^c} \|\boldsymbol{S} - \boldsymbol{S}'\|_2$$

- The desired function $f(\cdot) = \mathrm{dist}\left(\boldsymbol{S}, \mathcal{C}^c\right)$
- Challenges of computing $f(\cdot)$
  - No computationally efficient representation of ***C^c***
  - High dimensional problem
  - Non-convexity of ***C*** and ***C^c***



Related work

- Continuation power flow (CPF) method [Ajjarapu & Christy 92]
- Iterative method to find locally worst case margin [Dobson& Lu 93]

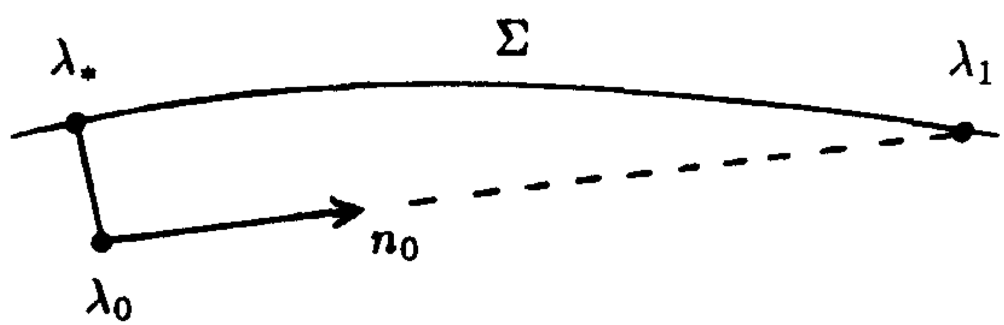


Figure 2. Insecure point λ_0 in load power parameter space

- Power flow's Jacobian's smallest singular value [Tiranuchit & Thomas 88] [Lof et al. 92]

Observations

- It is computationally easy to verify if any ***S*** is stable or not, i.e., obtain the 0/1 label of ***S***, (AC power flow).

- It is computationally easy to compute the distance from ***S*** to C^c along any given direction (CPF).

- It is computationally costly to (approximately) compute the voltage stability margin of ***S***, $f(\cdot) = \mathrm{dist}\,(\boldsymbol{S}, \mathcal{C}^c)$, as in principle all directions need to be explored.

Learning to Infer

- General procedure
  - Generate samples based on the power system model
  - Train a predictor of the stability margin based on the generated samples
  - Use the offline trained predictor for online inference of the margin

- ***Challenge: Lack of Labels***
  - Computing accurate approximation of stability margins is very computationally heavy --- *Even for offline computation*, generating such labeled data for training is very time consuming.

Learning to Infer

- Observation
  - Although computing a stability margin is hard, verifying if an operating condition is stable or not is fast.
- **Solution: Transfer Learning**
  - Generate a sufficient **large** dataset of data/operating conditions with **0/1** stability classification labels only.
  - Learn “as much as possible” from this 0/1 labeled data set.
  - Generate a relatively **small** dataset of data/operating conditions with stability **margin** labels.
  - Transfer what we learn from the large 0/1 labeled data set to further learning from the much smaller margin labeled data set.

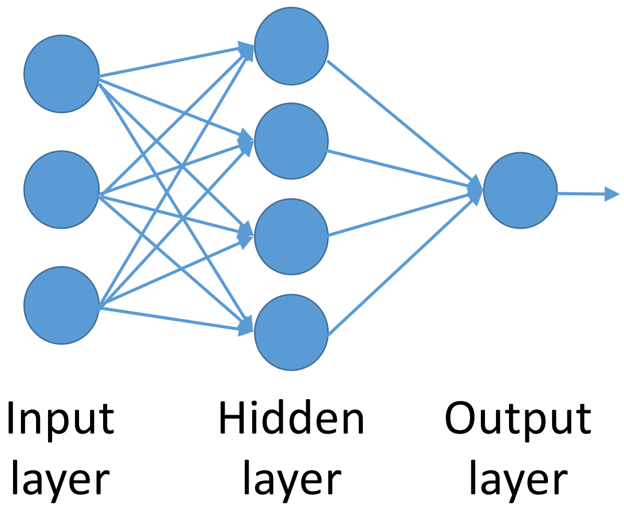
J. Li, Y. Zhao, Y. Lee and S. Kim, "Learning to Infer Voltage Stability Margin Using Transfer Learning," 2019 IEEE Data Science Workshop (DSW), pp. 270-274, 2019.

Overall training structure

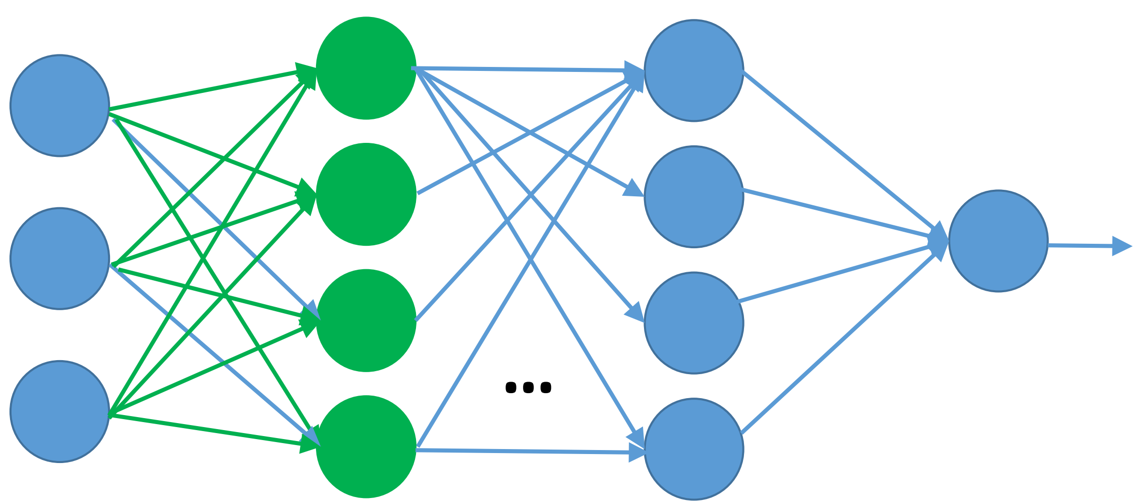
- Offline: Sample a very large number of ***S*** with its 0/1 labels.
- Offline: Sample only a small number of ***S*** with $f(\cdot) = \mathrm{dist}\left(\boldsymbol{S}, \mathcal{C}^c\right)$ (approximately) computed.
- Offline: Use the large 0/1 labeled data set of labeled ***S*** to learn a binary classifier $\hat{h}(\boldsymbol{S})$ that characterize the boundary of C.
- Offline: Use the small data set of ***S*** labeled with voltage stability margins, and employ the intermediate features learned by the classifier $\hat{h}(\boldsymbol{S})$ to further learn a margin estimator using regression.
- Online: Apply the learned margin estimator to any newly observed ***S*** to estimate its voltage stability margin in real time.

Predictor Design

- Learning voltage stability boundary

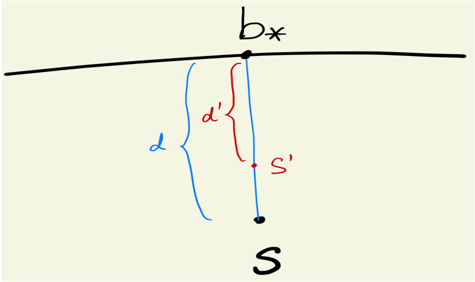


- Predicting voltage stability margin
  - **Reuse the features from classification**



Data Set

- The IEEE 300-bus test case.
- Generate directions: starting with one base case, multiply each P and Q with i.i.d. U[0,1]. **720K** directions generated.
- Generate **1.4M** (720K feasible & 720K infeasible) points using CPF, close to the stability boundary.
- Generate **11.4K** points with voltage stability margin approximately computed.
  - Search along every coordinate followed by the iterative method in [Dobson & Lu 93].
  - Data augmentation.



Numerical Experiments

- Learning voltage stability boundary
  - 1M samples for training, 440K for testing
  - Testing classification accuracy of 99.11%

- Predicting voltage stability margin
  - Transfer Learning
    - 10K samples for training, 1.4K for testing
    - Testing MSE: **0.001979214**
    - Testing R2: **0.9989559**
    - Testing computation time: **2.46ms/profile**

  - Baseline: Jacobian's SSV
    - Testing MSE: 1.503076
    - Testing R2: 0.204
    - Testing computation time: 89.65ms/prof

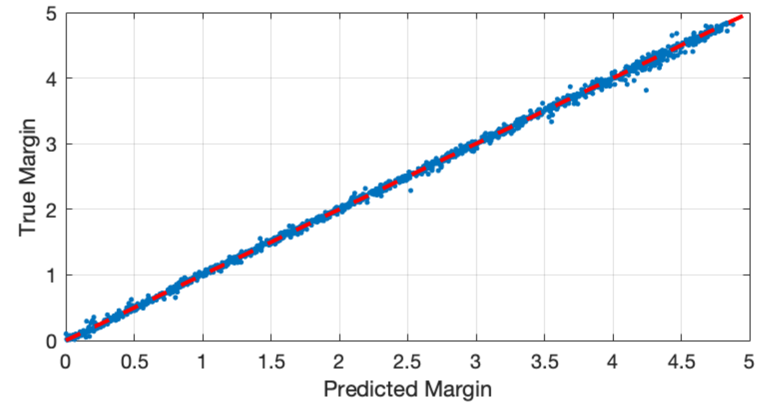


Fig. 1 Scatter plot from using transfer learning.

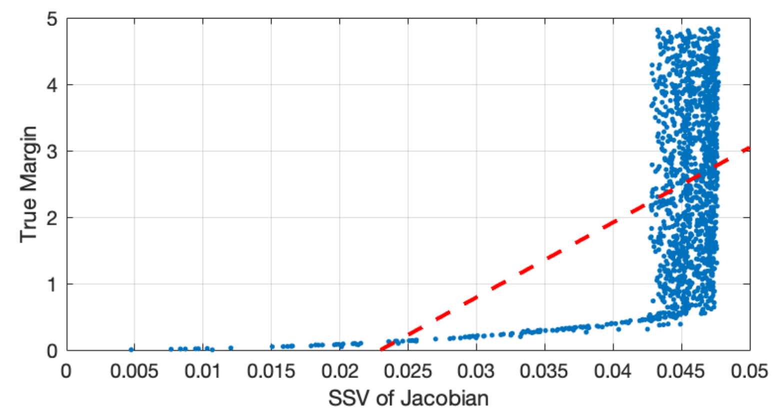


Fig. 2 Scatter plot from using Jacobian's SSVs.

Numerical Experiments (cont.)

- Zooming in for points *very close to the boundary*.
  - Transfer Learning
    - Testing MSE: **0.000725**.
    - Testing computation time: **2.45ms**
  - Baseline: Jacobian's SSV
    - Testing MSE: 0.001017
- Observation: SSV's predictive accuracy improves as the operating condition moves toward the boundary; Transfer learning still outperforms SSV.

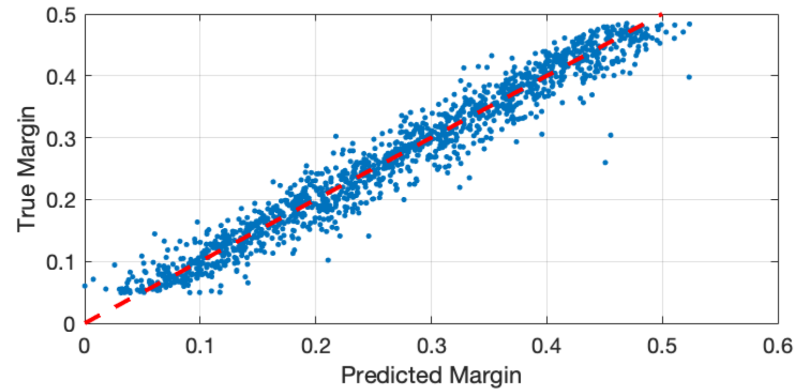


Fig. 1 Scatter plot from using transfer learning.

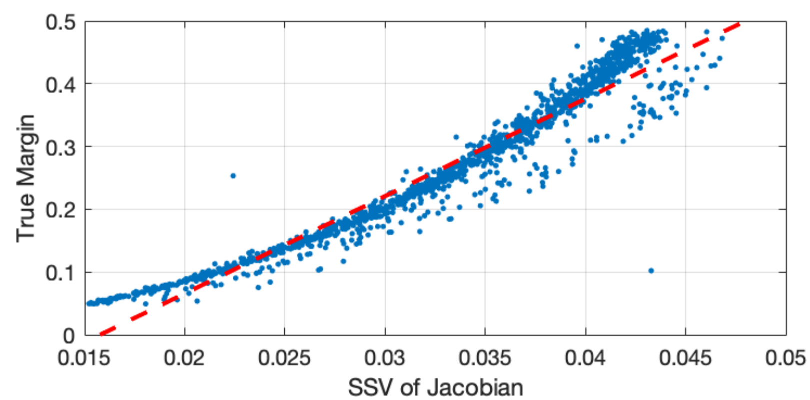


Fig. 2 Scatter plot from using Jacobian's SSVs.

Applying to OPF and contingency analysis

$$
\begin{aligned}
\min_{\boldsymbol{v}} \ & \sum_i C_i(P_i) \\
\text{s.t. } & P_i + jQ_i = \textstyle\sum_{k\sim i} V_i(V_i - V_k)^* y_{ik}^*, \ \forall i \\
& \underline{P}_i \le P_i \le \overline{P}_i, \ \forall i \in \text{Source} \\
& \underline{Q}_i \le Q_i \le \overline{Q}_i, \ \forall i \in \text{Source} \\
& P_i + jQ_i = L_{p,i} + jL_{q,i}, \ \forall i \in \text{Load} \\
& \text{other operational constraints} \\
& \boxed{f(\boldsymbol{P}, \boldsymbol{Q}) \ge \delta}
\end{aligned}
$$

- Solving OPF with stability margin guarantees.
  - Related work: [Tiranuchit & Thomas 88], “optimal posturing”.

- Very fast screening of contingencies with stability margin requirements.

Other Applications

- AC State estimation
  - [Zhang et al. 19], unrolling

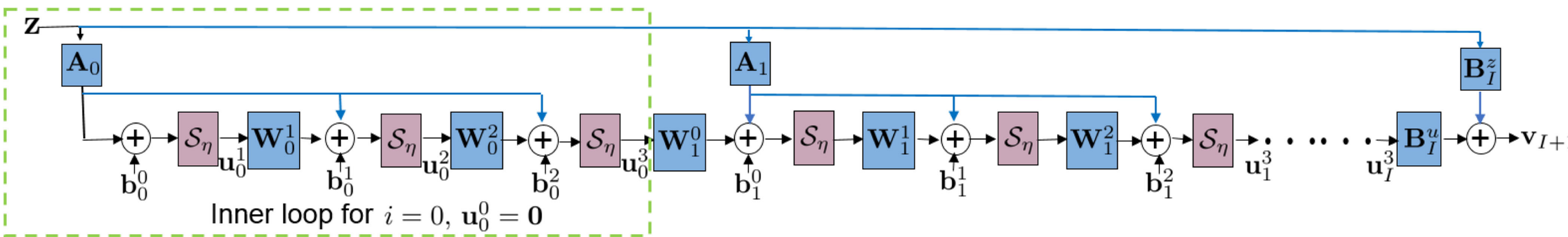


Fig. 2. Prox-linear net with $K = 3$ blocks.

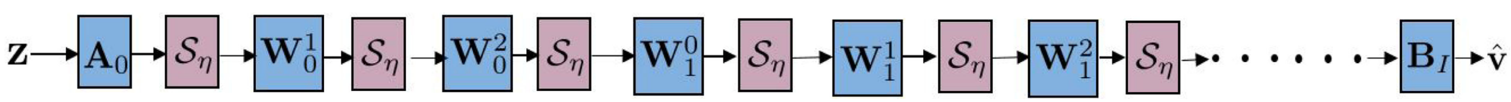


Fig. 3. Plain-vanilla FNN which has the same per-layer number of hidden units as the prox-linear net.

- Dynamic security assessment
  - See the previous tutorial by Tindemans and Cremer.

Other Applications

- Learning-based OPF
  - [Ng et al. 18] [Pan et al. 19][Chen et al. 20]

- Learning-based *N-k* security check (ongoing)

Summary

- A Learning-to-Infer Method is developed for addressing fundamentally hard problems in power system monitoring, by exploiting the information in the physical model via a data-driven approach.

- Two case studies, multi-line outage identification and voltage stability margin estimation, demonstrate the power of this methodology.

- Many more applications
  - Each has its own specific challenges and requires novel learning algorithm design.

References

- V. Ajjarapu and C. Christy, "The continuation power flow: a tool for steady state voltage stability analysis," [Proceedings] Conference Papers 1991 Power Industry Computer Application Conference, Baltimore, MD, USA, 1991, pp. 304-311.
- V. Ajjarapu and B. Lee, "Bifurcation theory and its application to nonlinear dynamical phenomena in an electrical power system," in IEEE Transactions on Power Systems, vol. 7, no. 1, pp. 424-431, Feb. 1992.
- S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in Proc. IEEE Conference on Decision and Control, 2013, pp. 1659–1664.
- J. Chen, Y. Zhao, A. Goldsmith, and H. V. Poor, "Line outage detection in power transmission networks via message passing algorithms," in Proc. 48th Asilomar Conference on Signals, Systems and Computers, 2014, pp. 350–354.
- Y. Chen, Y. Shi, B. Zhang, "Input Convex Neural Networks for Optimal Voltage Regulation," arXiv preprint arXiv:2002.08684. 2020.
- D. Deka, M. Chertkov, and S. Backhaus, "Structure learning in power distribution networks," IEEE Transactions on Control of Network Systems, 2017.
- I. Dobson and L. Lu, "New methods for computing a closest saddle node bifurcation and worst case load power margin for voltage collapse," in IEEE Transactions on Power Systems, vol. 8, no. 3, pp. 905-913, Aug. 1993.

References

- M. Garcia, T. Catanach, S. Vander Wiel, R. Bent, and E. Lawrence, “Line outage localization using phasor measurement data in transient state,” IEEE Transactions on Power Systems, vol. 31, no. 4, pp. 3019– 3027, 2016.
- I. Gera, Y. Yakoby, and T. Routtenberg, “Blind estimation of states and topology (best) in power systems,” in Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2017, pp. 1080–1084.
- T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning, 2nd Ed*. New York: Springer series in statistics, 2009.
- M. He and J. Zhang, “A dependency graph approach for fault detection and localization towards secure smart grid,” IEEE Transactions on Smart Grid, vol. 2, no. 2, pp. 342–351, Jun. 2011.
- J. Heydari and A. Tajer, "Quickest Localization of Anomalies in Power Grids: A Stochastic Graphical Framework," in IEEE Transactions on Smart Grid, vol. 9, no. 5, pp. 4679-4688, Sep. 2018.
- V. Kekatos, G. B. Giannakis, and R. Baldick, “Online energy price matrix factorization for power grid topology tracking,” IEEE Transactions on Smart Grid, vol. 7, no. 3, pp. 1239–1248, 2016.
- P. Kundur et al., "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," in IEEE Transactions on Power Systems, vol. 19, no. 3, pp. 1387-1401, Aug. 2004.

References

- J. Li, Y. Zhao, Y. Lee and S. Kim, "Learning to Infer Voltage Stability Margin Using Transfer Learning," 2019 IEEE Data Science Workshop (DSW), pp. 270-274, 2019.
- X. Li, H. V. Poor, and A. Scaglione, “Blind topology identification for power systems,” in Proc. the IEEE International Conference on Smart Grid Communications, 2013, pp. 91–96.
- P. -. Lof, T. Smed, G. Andersson and D. J. Hill, "Fast calculation of a voltage stability index," in IEEE Transactions on Power Systems, vol. 7, no. 1, pp. 54-64, Feb. 1992.
- Y. Ng, S. Misra, L. A. Roald and S. Backhaus, "Statistical Learning for DC Optimal Power Flow," 2018 Power Systems Computation Conference (PSCC), Dublin, 2018, pp. 1-7.
- X. Pan, T. Zhao, M. Chen, “DeepOPF: A Deep Neural Network Approach for Security-Constrained DC Optimal Power Flow,” arXiv preprint arXiv:1910.14448, 2019.
- J. E. Tate and T. J. Overbye, “Line outage detection using phasor angle measurements,” IEEE Transactions on Power Systems, vol. 23, no. 4, pp. 1644 – 1652, Nov. 2008.
- A. Tiranuchit and R. J. Thomas, "A posturing strategy against voltage instabilities in electric power systems," in IEEE Transactions on Power Systems, vol. 3, no. 1, pp. 87-93, Feb. 1988.
- Y. Weng, Y. Liao, and R. Rajagopal, “Distributed energy resources topology identification via graphical modeling,” IEEE Transactions on Power Systems, vol. 32, no. 4, pp. 2682–2694, 2017.
- Y. Yuan, O. Ardakanian, S. Low, and C. Tomlin, “On the inverse power flow problem,” arXiv preprint arXiv:1610.06631, 2016.

References

- L. Zhang, G. Wang and G. B. Giannakis, "Real-Time Power System State Estimation and Forecasting via Deep Unrolled Neural Networks," in IEEE Transactions on Signal Processing, vol. 67, no. 15, pp. 4069-4077, 1 Aug.1, 2019.
- Y. Zhao and B. Zhang, “Deep Learning in Power Systems,” in *Advanced Data Analytics in Power Systems,* Cambridge University Press, forthcoming.
- Y. Zhao, J. Chen, A. Goldsmith, and H. V. Poor, “Identification of outages in power systems with uncertain states and optimal sensor locations,” IEEE Journal of Selected Topics in Signal Processing, vol. 8, no. 6, pp. 1140–1153, Dec. 2014.
- Y. Zhao, J. Chen and H. V. Poor, “Learning to infer: a new variational inference approach for power grid topology identification,” Proc. IEEE Workshop on Statistical Signal Processing (SSP), Jun. 2016.
- Y. Zhao, J. Chen and H. V. Poor, "A Learning-to-Infer Method for Real-Time Power Grid Multi-Line Outage Identification," in IEEE Transactions on Smart Grid, vol. 11, no. 1, pp. 555-564, Jan. 2020.
- H. Zhu and G. B. Giannakis, “Sparse overcomplete representations for efficient identification of power line outages,” IEEE Transactions on Power Systems, vol. 27, no. 4, pp. 2215–2224, Nov. 2012.

Thanks!