DISTRIBUTION SYSTEM STATE ESTIMATION VIA DATA-DRIVEN AND PHYSICS-AWARE DEEP NEURAL NETWORKS

Liang Zhang, Gang Wang, and Georgios B. Giannakis

Dept. of ECE and Digital Tech. Center, Univ. of Minnesota, Mlps, MN 55455, USA
E-mails: {zhan3523, gangwang, georgios}@umn.edu

ABSTRACT

Massive integration of renewables and electric vehicles comes with unknown dynamics — what exemplifies the need for fast, accurate, and robust distribution system state estimation (DSSE). Due to limited real-time measurements however, optimization-oriented DSSE faces major challenges related to convergence, as well as multiple global/local minima. To address these challenges, this paper puts forth a novel deep neural network (DNN)-based computational framework for DSSE that consists of two modules: a deep recurrent neural network (RNN) based pseudo-measurement postulating module, and a prox-linear net-based real-time state estimation module. Both RNN and prox-linear nets learn complex nonlinear functions, and can afford efficient training by leveraging existing deep learning platforms. Numerical tests with semi-real load data demonstrate the merits of the DNN-based DSSE approach.

***Index Terms*—** Distribution system state estimation, pseudo measurement, recurrent neural network, deep neural network.

1. INTRODUCTION

Fast and accurate estimation of power system states is crucial, not only for situational awareness and system protection, but also for energy management [5, 13, 15]. Given limited power measurements acquired by supervisory control and data acquisition (SCADA) and distribution automation systems, state estimation (SE) aims to recover the unknown system state, that is, the complex voltages across the network [13].

Different from transmission systems where metering devices are placed at almost all buses, distribution grids only have a limited number of meters [13], which causes partial observability and also challenges conventional SE schemes. To enhance observability, distribution system state estimation (DSSE) has to rely on the so-called pseudo measurements that can be generated via load, generation, and voltage forecasting tools [13]. Typical pseudo-measurement generation schemes leverage feed-forward neural networks (FNNs) [9], and clustering approaches [4].

Exploiting both actual and pseudo measurements, several DSSE solvers have been proposed. DSSE was posed as a weighted least-squares (WLS) problem, and it is often solved via Gauss-Newton iterations in [1, 2]. Bayesian DSSE using FNNs was suggested in [10]. To improve convergence, single-layer FNNs were employed to obtain a suitable initialization for Gauss-Newton iterations [14]. Successive WLS accounting for forecasted load values has been investigated in [3]. Nonetheless, all these optimization-oriented approaches are computationally demanding, discouraging their implementation in real time. In addition, the role of pseudo measurements on DSSE has not been investigated.

Along with our recent proposal on efficient SE of transmission networks [17, 18], we advocate here a real-time DSSE framework leveraging data- and physics-driven DNNs. The contribution is two-fold. First, we develop physics-specific prox-linear nets for SE of unbalanced distribution grids. The prox-linear net, constructed by unrolling the prox-linear SE solver in [12], requires minimal tuning effort, and features 'skip-connections,' that enable efficient training of DNNs (see e.g., [7]). Second, pseudo measurements such as forecasted loads and voltages are postulated via deep RNNs, that are capable of capturing complex nonlinear dependencies present in time series data. Subsequently, our overall DSSE scheme comprising RNN-based measurement forecasting and prox-linear net based state estimation modules is specified, and tested. Simulated tests using semi-real load data showcase the merits of the proposed DSSE framework.

2. NOVEL DSSE APPROACH

Consider an unbalanced distribution network comprising $N+1$ buses indexed by $n \in \mathcal{N} := \{0, 1, \ldots, N\}$, and phases indicated by $\phi \in \{a, b, c\}$. Suppose that the distribution grid is functionally radial with the substation bus numbered by $n = 0$. Per phase ϕ of bus n, let $v_{n,\phi} := v_{n,\phi}^r + jv_{n,\phi}^r$ denote its associated complex voltages, and $S_{n,\phi} := P_{n,\phi} + jQ_{n,\phi}$ its complex power injections, with $P_{n,\phi}$ $(Q_{n,\phi})$ denoting respectively the active (reactive) injections. For each line (nn', ϕ) that connects phase ϕ of bus n with phase ϕ of bus n', let

This work was supported in part by NSF grants 1508993, 1509040, and 1711471.

$S^f_{nn',\phi} := P^f_{nn',\phi} + jQ^f_{nn',\phi}$ be the complex power flow at the 'forwarding' end with $P^f_{nn',\phi}$ ($Q^f_{nn',\phi}$) denoting the active (reactive) flow. Likewise, let $S^e_{nn',\phi}$, $P^e_{nn',\phi}$, and $Q^e_{nn',\phi}$ represent the complex, active, and reactive flows at the 'terminal' end of line (nn',ϕ). To perform DSSE, we measure M system variables collected in the vector $\mathbf{z} := [z_1, \ldots, z_M]^\top$.

DSSE aims to retrieve the state vector $\mathbf{v} := [v_1^r, v_1^i, \ldots, v_N^r, v_N^i]^\top \in \mathbb{R}^{2N}$ from the generally noisy vector $\mathbf{z}$. Concretely, DSSE can be stated as follows. Given $\mathbf{z}$ and the associated linear or quadratic functions $\{h_m(\cdot)\}_{m=1}^M$ obeying

$$z_m = h_m(\mathbf{v}) + \epsilon_m, \qquad \forall m = 1, \ldots, M \tag{1}$$

where ϵ_m accounts for the measurement noise and modeling errors, our goal is to find $\mathbf{v} \in \mathbb{R}^{2N}$. To endow our estimate with resilience to bad data due to e.g. cyber attacks, the following least-absolute-value estimate is sought (see e.g., [12])

$$\arg \min_{\mathbf{v}\in\mathbb{R}^{2N}} \frac{1}{M} \sum_{m=1}^{M} |z_m - h_m(\mathbf{v})| \,. \tag{2}$$

Existing solvers for (2) are often computationally heavy, and several encounter even convergence issues especially when the number of measurements M is small. With growing network sizes as well as unpredictable dynamics of renewable generation, real-time DSSE schemes are highly desirable. To this end, the ensuing section develops learning-based DSSE.

2.1. Prox-linear net for power system state estimation

Learning-based DSSE seeks the function mapping from the measurement vector $\mathbf{z}$ to $\mathbf{v}$ based on historical/simulated data. As the relationship between $\mathbf{z}$ and $\mathbf{v}$ is complex, the function mapping from $\mathbf{z}$ to $\mathbf{v}$ is arguably nonlinear; see e.g. [14, 18]. To render nonlinear estimators computationally tractable, DNN [17] or kernel-based [16] approaches provide viable solutions. Specifically, our physics-aware prox-linear neural networks have been demonstrated empirically to be successful in estimating the states of transmission networks [18].

The prox-linear net is devised by unrolling a recently proposed prox-linear solver for (2) [12]. Here we will design related prox-linear nets, but for DSSE. The architecture of our $3(I+1)$-layer prox-linear net is depicted in Fig. 2, where $\mathbf{z}$ denotes the input measurement vector, f in all the greenish boxes is the pre-selected nonlinear activation function (understood entry-wise when applied to vector inputs), such as 'tanh', soft-thresholding operator, or rectified linear unit (ReLU). The vector $\{\mathbf{b}_i^k\}_{0<=i<=I}^{0<=k<=2}$, and matrices $\{\mathbf{A}_i\}_{0\le i\le 1}$, $\{\mathbf{W}_i^k\}_{0\le i\le I}^{1\le l\le 3}$, $\mathbf{B}_I^u$, and $\mathbf{B}_I^z$ contain weights learned from data in the training phase. The number of hidden units per-layer is set equal to the dimension of the input vector $\mathbf{z}$. Relative to the conventional feed-forward NN illustrated in Fig. 1, the prox-linear net features 'skip-connections' through the bluish lines in Fig. 2, connecting input $\mathbf{z}$ directly to intermediate/output layers. Empirical tests have demonstrated that skip connections help avoid the so-termed 'vanishing/exploding' gradient problems, thus facilitating the training process of DNNs [7].

Upon learning the weight coefficients during the off-line training stage, the prox-linear net can be employed for inferring the distribution system states in real time. Unlike transmission networks where metering devices are adequate, distribution grids suffer from partial observability due to limited instrumentation. To enhance observability, data vector $\mathbf{z}$ have to be augmented with pseudo measurements. In the ensuing subsection, deep recurrent neural networks will be advocated for predicting pseudo measurements.

2.2. RNNs for state predictions as pseudo-measurements

Besides the actual measurements acquired by smart meters, load, generation, and voltage forecasting approaches can be employed to generate pseudo-measurements. The forecasted measurements $\tilde{\mathbf{z}}_t$ at time t, obey the model

$$\tilde{\mathbf{z}}_t = \tilde{\mathbf{h}}(\mathbf{v}) + \boldsymbol{\xi}_t \tag{3}$$

where $\boldsymbol{\xi}_t$ accounts for the foresting error, and $\tilde{\mathbf{h}}(\cdot)$ as in (1) represents linear or quadratic functions dictated by the physics. Clearly, $\tilde{\mathbf{z}}_t$ provides additional equations that augment (1), thus enhancing system observability. Given time series $\{\tilde{\mathbf{z}}_\tau\}_{\tau=0}^{t-1}$, the following model is typically adopted

$$\tilde{\mathbf{z}}_t = \boldsymbol{\phi}(\tilde{\mathbf{z}}_{t-1}, \tilde{\mathbf{z}}_{t-2}, \ldots, \tilde{\mathbf{z}}_{t-r}) + \boldsymbol{\eta}_t \tag{4}$$

where r is the number of past measurements used for predicting $\tilde{\mathbf{z}}_t$, $\boldsymbol{\eta}_t$ captures the modeling inaccuracies, and $\boldsymbol{\phi}$ represents the unknown function representing the load/state transitions. To obtain pseudo-measurements, we will capture $\boldsymbol{\phi}$ using a deep RNN, whose parameters will be estimated next.

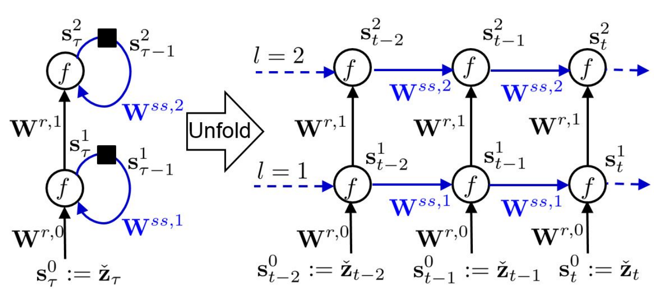

Fig. 3: An unfolded deep RNN with no outputs.

Deep RNNs are RNNs having multiple processing layers designed to learn from correlated time series data with hierarchical nonlinear transformations. They are not only scalable to sequence inputs with large r, but also capable of capturing complex dependencies within time series. This characteristic has allowed deep RNNs to improve upon state-of-the-art in several applications, including machine translation and music modeling [6].

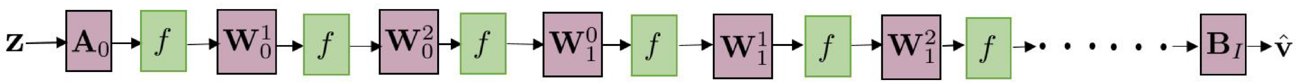

Fig. 1: Plain-vanilla feed-forward net whose number of hidden units per layer is the same as in our prox-linear net.

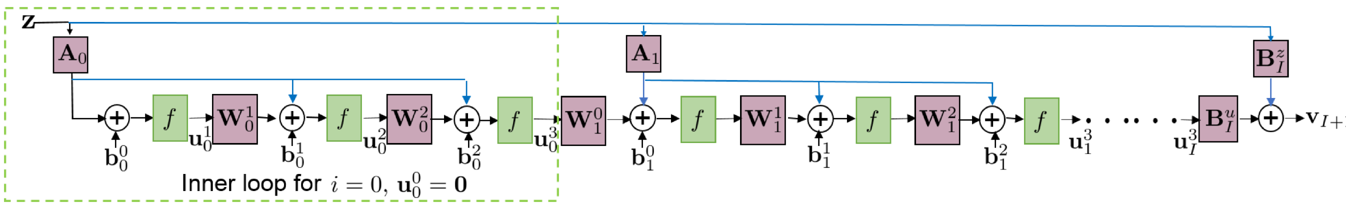

Fig. 2: Prox-linear net with $K = 3$ blocks.

Given initial states $\{\mathbf{s}_{t-r-1}^l\}_{l\geq 1}$, and the input time series $\{\tilde{\mathbf{z}}_\tau\}_{\tau=t-r}^{t-1}$, deep RNNs find for $\tau = t-r, \ldots, t$ [11]

$$\mathbf{s}_\tau^l = f\big(\mathbf{W}^{l-1}\mathbf{s}_\tau^{l-1} + \mathbf{W}^{ss,l}\mathbf{s}_{\tau-1}^l + \mathbf{b}^{l-1}\big), \quad l \geq 1 \qquad (5)$$

where l is the layer index, $\{\mathbf{W}^{r,l}, \mathbf{W}^{ss,l}, \mathbf{b}^l\}$ consist of unknown weights, and $\mathbf{s}_\tau^l$ is the 'so-termed' hidden state of the l-th layer at time τ with $\mathbf{s}_\tau^0 := \tilde{\mathbf{z}}_\tau$. The computational graph representing (5) for $l = 2$ is depicted in the left panel of Fig. 3, where the black filled squares indicate one-step delay units, and the bias vectors $\mathbf{b}^l = \mathbf{0}\ \forall l$. The right panel of Fig. 3 shows the unfolded version of this graph with columns representing time slots, and rows corresponding to layers. The output of our deep RNN has the form

$$\check{\mathbf{z}}_t = \mathbf{W}^{out}\mathbf{s}_{t-1}^l + \mathbf{b}^{out} \qquad (6)$$

where $\mathbf{W}^{out}$ and $\mathbf{b}^{out}$ collect unknown weights, and $\check{\mathbf{z}}_t$ is the predicted values of $\tilde{\mathbf{z}}_t$. Given historical measurement series $\{\mathbf{z}_\tau\}$, weight coefficients $\{\mathbf{W}^{out}, \mathbf{b}^{out}, \mathbf{W}^{r,l}, \mathbf{W}^{ss,l}, \mathbf{b}^l\}$ can be learned end-to-end via back-propagation through time [6].

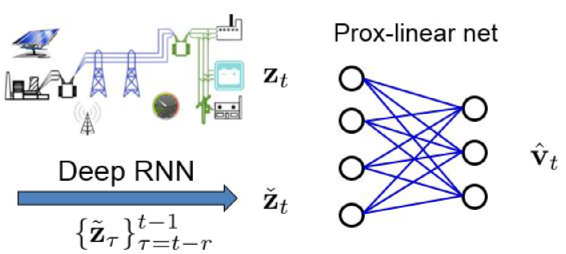

Fig. 4: DNN-based real-time DSSE.

To summarize, our proposed real-time distribution grid monitoring approach is made up of two modules: a real-time DSSE scheme based on the prox-linear net, and a pseudo-measurement generation scheme via deep RNNs. The overall flow chart of the proposed approach is depicted in Fig. 4.

3. NUMERICAL TESTS

Performance of our novel DSSE approach was tested using the unbalanced IEEE 13-bus test feeder. To prepare the training and testing datasets, real load data from the National Renewable Energy Laboratory [1] were used, where all loads were first normalized to match the scale of active power demands in the IEEE 13-bus system. The actual voltage profile $\mathbf{v}$ was obtained by solving the AC power flow equations using the forward-backward sweep algorithm [8].

The obtained measurements include: i) PMU measurements at bus 6. All voltage and current phasors of the three phases were measured, adding up to 14 complex measurements, or 28 real measurements. ii) SCADA measurements at branches 4, 5, and 13. In particular, magnitudes of current flows on all phases were measured, yielding 8 real measurements. iii) Load measurements at buses $3, 5, 7, 9, 10, 11$ and 13. Specifically, we considered two setups. In the first setup, all load measurements were obtained from metering devices, whereas in the second one only a fraction of load measurements were obtained using our RNN-based forecasting scheme. All measurements were further corrupted with zero-mean additive white Gaussian noise, where the standard deviations for PMU and SCADA measurements were 10^{-3} and 10^{-2}, respectively. In total, $8,760$ measurement-voltage $(\mathbf{z}, \mathbf{v})$ pairs were generated, among which the first $6,132$ (70% of the total) pairs were used for training, and the remaining $2,628$ (30%) pairs for testing.

The first experiment evaluates performance of our prox-linear net based DSSE, in which all measurements come from metering devices. Estimation performance was assessed via the normalized root mean-square error (RMSE) $\|\hat{\mathbf{v}} - \mathbf{v}\|_2/N$, with $\hat{\mathbf{v}}$ being our estimate, and $\mathbf{v}$ the ground-truth. In particular, a 3-layer prox-linear net with $I = 0$ was implemented. A 3-layer feed-forward NN (cf. Fig. 1) having the same depth as our prox-linear net was also simulated as a baseline. For comparison, all NNs used the ReLU activation functions, and

[1] https://openei.org/datasets/files/968/pub/.

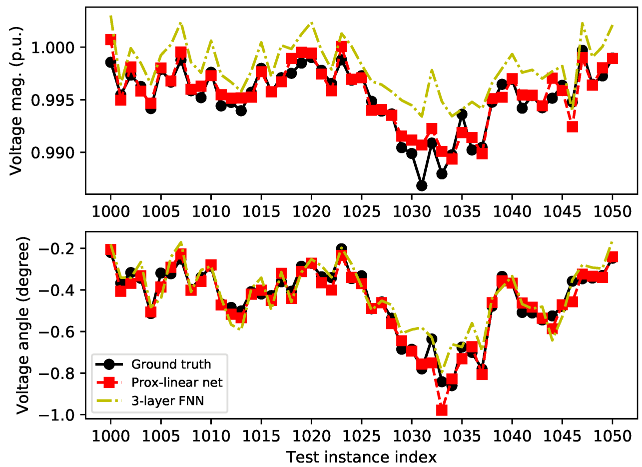

Fig. 5: Estimation errors in voltage magnitudes and angles of phase a of bus 4 from instances $1,000$ to $1,050$.

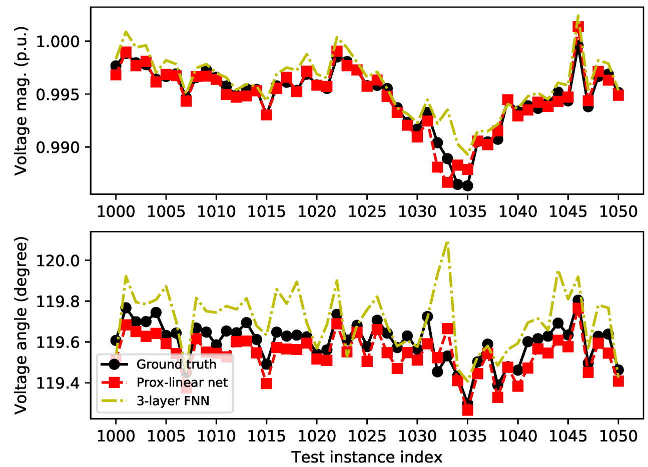

Fig. 6: Estimation errors for voltage magnitudes and angles of phase c of bus 7 from instances $1,000$ to $1,050$.

they were trained by the 'Adam' optimizer for 200 epochs, with a starting step size of 10^{-3}. The training process was performed using the TensorFlow interface on an NVIDIA Titan X GPU.

Estimation performance in terms of the RMSE averaged over $2,628$ testing examples, for the prox-linear net, and 3-layer NN are respectively 5.99×10^{-4} and 7.08×10^{-4}, which showcase competitive performance of our prox-linear net. The actual system states along with the estimated ones provided by the prox-linear net and a 3-layer NN, for phase a of bus 4 and phase c of bus 7 in the test set, are depicted in Figs. 5 and 6, respectively.

The second experiment tests the efficacy of our DSSE scheme, where only a fraction of measurements are provided via deep RNN forecasts. For a fixed number of pseudo measurements, 20 independent trials were run. For each trial, locations of the pseudo measurements were sampled uniformly at random. Fig. 7 presents the sample mean and standard derivation of RMSE averaged over the 20 trials for different numbers of pseudo measurements. Two pseudo-measurement generation schemes were compared: i) pseudo-measurements obtained using a 3-layer RNN; and, ii) pseudo measurements set equal to their values at time $t-1$. It is observed that the DSSE with RNN generated pseudo-measurements offers improved performance in all cases.

4. CONCLUSIONS

This paper dealt with DSSE drawing from advances in contemporary deep learning. In particular, physics-aware prox-linear nets that are obtained by unrolling the state-of-the-art prox-linear optimization solver were developed for DSSE. To enhance DSSE performance, load and voltage pseudo-

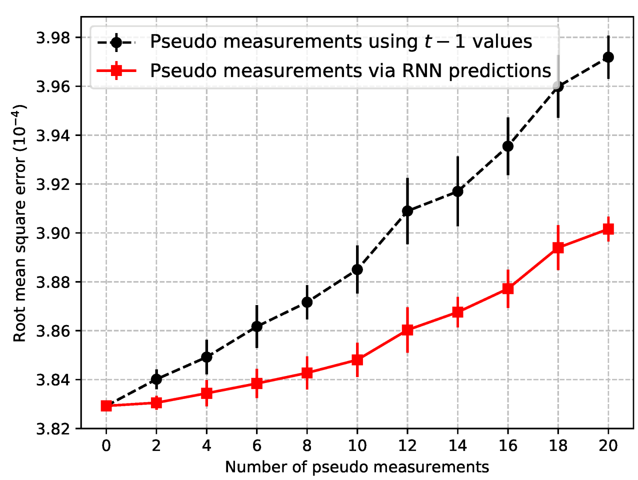

Fig. 7: Average RMSE w.r.t. the number of predicted values.

measurements were obtained using RNNs. The proposed DNN-based DSSE framework is computationally inexpensive, and easy to implement. Preliminary tests on the unbalanced IEEE 13-bus test feeder showcase the merits of our DNN-based DSSE scheme. Our agenda includes DNN-based solvers for optimal power flow and generalized DSSE.

Acknowledgment

The authors would like to thank Z. Wang for preparing the data used in the numerical tests.

5. REFERENCES

[1] M. E. Baran and A. W. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1601–1609, Aug. 1994.

[2] ——, "A branch-current-based state estimation method for distribution systems," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 483–491, Feb. 1995.

[3] I. Dvzafić and R. A. Jabr, "Real time multiphase state estimation in weakly meshed distribution networks with distributed generation," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4560–4569, Nov. 2017.

[4] Y. R. Gahrooei, A. Khodabakhshian, and R.-A. Hooshmand, "A new pseudo load profile determination approach in low voltage distribution networks," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 463–472, Apr. 2018.

[5] G. B. Giannakis, V. Kekatos, N. Gatsis, S.-J. Kim, H. Zhu, and B. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, Sep. 2013.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016, http://www.deeplearningbook.org.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European Conf. Comput. Vision*. Amsterdam, Netherlands: Springer, Oct. 2016, pp. 630–645.

[8] W. H. Kersting, *Distribution System Modeling and Analysis*. Boca Raton, FL: CRC Press, 2002.

[9] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1888–1896, Nov. 2012.

[10] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *arXiv:1811.02756*, 2018.

[11] R. Pascanu, C. Gulcehre, K. Cho, and Y. Bengio, "How to construct deep recurrent neural networks," in *Proc. Intl. Conf. on Learning Representations*, Banff, Canada, Apr. 2014.

[12] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," *IEEE Trans. Smart Grid*, 2019, doi: 10.1109/TSG.2019.2897100.

[13] G. Wang, G. B. Giannakis, J. Chen, and J. Sun, "Distribution system state estimation: An overview of recent developments," *Front. Inf. Technol. Electron. Eng.*, vol. 20, no. 1, pp. 4–17, Jan. 2019.

[14] A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, "Data-driven learning-based optimization for distribution system state estimation," *arXiv:1807.01671*, 2018.

[15] L. Zhang, V. Kekatos, and G. B. Giannakis, "Scalable electric vehicle charging protocols," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1451–1462, Mar. 2017.

[16] L. Zhang, G. Wang, and G. B. Giannakis, "Going beyond linear dependencies to unveil connectivity of meshed grids," in *Proc. IEEE Wkshp. on Comput. Adv. Multi-Sensor Adaptive Process.*, Curacao, Dutch Antilles, Dec. 2017.

[17] ——, "Real-time power system state estimation and forecasting via deep neural networks," *arXiv:1811.06146*, Nov. 2018.

[18] ——, "Real-time power system state estimation via deep unrolled neural networks," in *Proc. Global Conf. on Signal and Info. Process.*, Anaheim, CA, USA, Nov. 2018.