Optimizing State Estimation Error with the LinDist3Flow Model

Jeslyn Ero, Kieran Slattery, Xianhe Qin, Samuel Talkington, Daniel K. Molzahn

Abstract—**This research paper presents an algorithmic approach to optimizing state estimation errors in unbalanced distribution networks as the integration of renewable energy sources such as solar and wind increases. These resources introduce uncertainty into grid models, challenging the satisfaction of engineering constraints.**

The study focuses on the impacts of this randomness on power flow equations and aims to enhance grid state estimation by combining electric grid physics with techniques from probability theory. The goal is to develop methodologies that effectively utilize uncertain measurement data, optimizing the accuracy and efficiency of smart meter data streams.

The algorithm implementation aims to learn the optimal line parameters (resistance and reactance) in a three-phase unbalanced distribution system by minimizing the discrepancy between predicted and measured voltage values. It leverages the LinDist3Flow power flow approximation model to simulate voltage magnitudes, and formulates a loss function representing the squared error between the predicted and noisy measurements. The ADAM optimization algorithm iteratively minimizes this loss function, adjusting the parameters to achieve the best fit to observed data. This contributes to improving the accuracy of power flow modeling under noisy and unbalanced conditions.

Index Terms—**estimation; distribution systems; smart meter measurements; probability theory**

I. Introduction

Efficient use of smart meter bandwidth is an emerging challenge in distribution systems engineering. Ongoing research has explored this topic in distribution networks with limited communication bandwidth [1]. By applying mathematically robust theories to practical challenges in distribution system state estimation, optimization, and control, the aim is to overcome these hurdles effectively.

The modern electric grid is transitioning from a centralized to a more distributed architecture, presenting new operational challenges due to the anticipated rapid growth of diverse distributed energy resources (DER), including rooftop photovoltaics, electric vehicles, and storage systems at the edge of the grid [2]. This shift also introduces multiple decision-makers who could manipulate the system under decentralized electricity market designs.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. Email: {jero6,kslattery7,xqin68,talkington,molzahn}@gatech.edu.

The authors gratefully acknowledge funding from PSERC project T-67: Smart Meter- Driven Distribution Grid Visibility and Control. The work of S. Talkington is supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650044. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

State estimation errors occur when there are discrepancies between the actual state of a distribution network and the estimated values derived from measurements. Accurate estimation is critical for maintaining grid stability and optimizing operations, especially as the adoption of renewable energy sources introduces uncertainty.

The inherent randomness of renewable energy resources causes it to be challenging to generate realistic power flow solutions. While classical approaches to solving the power flow equations with a large suite of possible renewable scenarios can generate useful synthetic data, this approach is challenged by numerous limitations. This includes the large amount of data processing and computational resources required and the limitation of sample-based approaches to generalized to unseen data.

II. Background

This research aims to propagate the uncertainty of renewable generation scenarios through physics-informed power flow models, specifically utilizing the LinDistFlow approximation. In this context, linear matrices are used to map n-dimensional power injection vectors to corresponding voltages, where the power injections are modeled as random variables drawn from a given distribution. The goal is to demonstrate that, for a random vector of power injections, the expected value of the maximum voltage across the grid is bounded—a key result in ensuring grid stability.

Additionally, this study will differentiate between distribution networks and transmission networks, focusing on the implications of randomness in the context of distribution grids, where renewable power is injected at various points. By leveraging the operator norm of the admittance matrix, the research will provide insights into future scenarios in which large sections of the grid randomly inject renewable power, highlighting both the opportunities and risks associated with such a system. In Fig. 1, we illustrate the high-level idea behind this research.

An important aspect of the research is a grid model-informed machine learning (ML) tool that integrates heterogeneous data streams and creates synchronous measurement snapshots for use by a hybrid robust state estimator [3]. This framework not only provides accurate state estimates but also real-time feedback for the refinement of machine learning models, leading to improved monitoring

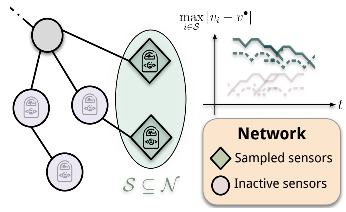

Fig. 1. Illustration of the problem: Dynamic sampling of streaming smart meters to expose extreme voltage magnitude perturbations with high probability.

performance [3]. This has been experimentally validated through simulated scenarios on an electric utility's distribution system [3]. By combining computational tools and techniques, the complex challenges presented by the integration of renewable resources into modern electric distribution grids are acknowledged.

A key contribution is the introduction of a new stochastic LinDistFlow model that is agnostic to the probability distributions of nodal power injections. This model helps to understand the probabilistic concentration of nodal voltage magnitudes, which are derived from the model and are explicitly dependent on the topology and parameters of the network. Furthermore, the LinDistFlow model allows for a linear approximation of what is originally a nonlinear, complex problem [4]. The LinDist3Flow model takes this a step further for 3-phase networks, whose applications are primarily used for high power-consuming devices [4].

A. Related work

The three-phase unbalanced LinDist3Flow approximation [4] is given as

$$v = v_0 + A_3^{-1}\,\mathrm{bdiag}(H^P)A_3^{-\top}p + A_3^{-1}\,\mathrm{bdiag}(H^Q)A_3^{-\top}q,$$

where $v \in \mathbb{R}^{3n}$ is a vector of squared single-phase voltage magnitudes for each phase (a,b,c), A_3 is the network incidence matrix, blkdiag(H^P) and blkdiag(H^Q) are the block diagonal matrices for each line(i,j), and P and Q are the active and reactive power injection vectors. The paper improves on the LinDist3Flow approximation via a mathematical optimization, extending the work of [4].

III. Problem formulation

The LinDist3Flow approximation is the linearization of the complex and non-linear AC power flow equations for 3-phase radial distribution networks [4] [5] . The LinDistFlow model linearizes the DistFlow equations by assuming that the active and reactive line losses are much smaller than the active and reactive line flows. This linearization makes the model computationally efficient and suitable for optimization problems in distribution systems [4].

A. The LinDistFlow Model

The LinDistFlow model approximates the nodal voltage vector as

$$v = v_0 + R\,p + X\,q,$$

where p and q are the active and reactive power injection vectors, and the nodal matrices R and X are given by

$$R = A_3^{-T}\,\mathrm{bldiag}(H^P)\,A_3^T, \quad X = A_3^{-1}\,\mathrm{bldiag}(H^Q)\,A_3^{-T}.$$

Here, for each branch $\ell := (i,j) \in \mathcal{E}$, the local impedance matrices are defined as

$$H_\ell^P = \begin{bmatrix} -2r_{ij}^{aa} & r_{ij}^{ab} - \sqrt{3}x_{ij}^{ab} & r_{ij}^{ac} + \sqrt{3}x_{ij}^{ac} \\ r_{ij}^{ba} + \sqrt{3}x_{ij}^{ba} & -2r_{ij}^{bb} & r_{ij}^{bc} - \sqrt{3}x_{ij}^{bc} \\ r_{ij}^{ca} - \sqrt{3}x_{ij}^{ca} & r_{ij}^{cb} + \sqrt{3}x_{ij}^{cb} & -2r_{ij}^{cc} \end{bmatrix} \quad (1a)$$

$$H_\ell^Q = \begin{bmatrix} -2x_{ij}^{aa} & x_{ij}^{ab} + \sqrt{3}r_{ij}^{ab} & x_{ij}^{ac} - \sqrt{3}r_{ij}^{ac} \\ x_{ij}^{ba} - \sqrt{3}r_{ij}^{ba} & -2x_{ij}^{bb} & x_{ij}^{bc} + \sqrt{3}r_{ij}^{bc} \\ x_{ij}^{ca} + \sqrt{3}r_{ij}^{ca} & x_{ij}^{cb} - \sqrt{3}r_{ij}^{cb} & -2x_{ij}^{cc} \end{bmatrix} \quad (1b)$$

and bldiag(H^P) and bldiag(H^Q) are block diagonal matrices (of dimensions $3m \times 3m$, where m is the number of branches) that aggregate these blocks. In symbols,

$$\mathrm{bldiag}(H^P) = \begin{bmatrix} H_1^P & \dots & 0_{3\times3} \\ \vdots & \ddots & \vdots \\ 0_{3\times3} & \dots & H_m^P \end{bmatrix},$$

$$\mathrm{bldiag}(H^Q) = \begin{bmatrix} H_1^Q & \dots & 0_{3\times3} \\ \vdots & \ddots & \vdots \\ 0_{3\times3} & \dots & H_m^Q \end{bmatrix}.$$

B. Error Minimization in a parameterized LinDist3Flow model

Let y denote the measured nodal voltage vector. The prediction error (or residual) is then

$$e = y - v = y - (R\,p + X\,q)\,.$$

We quantify this error using a least-squares loss function:

$$\mathcal{L}(r,x) = \frac{1}{2}\|e\|^2 = \frac{1}{2}\,\|y - (R(r)\,p + X(x)\,q)\|^2\,.$$

Our objective is to minimize this loss, i.e.,

$$\min_{r,x}\ \mathcal{L}(r,x) = \min_{r,x}\ \frac{1}{2}\,\|y - (R(r)\,p + X(x)\,q)\|^2\,.$$

Minimizing $\mathcal{L}(r,x)$ adjusts the branch parameters r and x so that the model's predicted voltage v closely matches the measurements y. This process not only enhances the accuracy of the state estimation but also provides sensitivity information about how changes in r and x affect the overall error, thereby calibrating the model to better reflect the physical system.This optimization process estimates the best-fitting values of the resistance and reactance vectors, **r** and **x**, to ensure that the predicted voltage profile aligns as closely as possible with the observed (noisy) measurements.

IV. Results

This experiment evaluates the performance of the optimization procedure in recovering the best-fitting parameters $\theta = [\mathbf{r}; \mathbf{x}]$ that describe the physical characteristics of the network lines.

To implement the optimization, we follow the methodology proposed by Babak in Section IV of his paper. Specifically, we construct the reduced 3-phase incidence matrix A_3 where each line-phase pair is treated independently. This design accounts for multigraph structures (i.e., multiple connections between the same buses) and unbalanced systems, where some lines may not be present on all three phases. In Babak's approach, each bus-phase and line-phase combination is mapped to a unique row and column, respectively, and entries in A_3 are assigned values of $+1$, -1, or 0, depending on whether the bus is the sending or receiving end and whether the phase is active. Using this matrix, we build the impedance matrices R and X based on estimated resistance and reactance values, encoded in the block diagonal matrices H_p and H_q. These matrices are constructed by scaling identity blocks according to the parameter vector θ, and concatenating them using a block diagonal operation. This structure allows the optimization process to be phase-aware and topology-consistent, even in the presence of missing or unbalanced phases.

A. Analytical results

To verify the solution of the new LinDist3Flow model, we compare our result against the result of an AC Power Flow Solver. Utilizing the PowerModelsDistribution package in Julia, we can utilize the built-in functions to find the voltage magnitudes and phase angles that comprise the solution. We simply call this function and print the voltage magnitudes and phase angles at each bus, which can be seen here: Furthermore, we can write our own implementation of the

```
EXIT: Optimal Solution Found.
Bus primary: Voltage Magnitude = [0.22653762676502015, 0.2284
828000276323, 0.22794679222855124], Phase Angle = [-0.2242550
50717533, -120.11324476955176, 120.12281836393022] degrees
Bus sourcebus: Voltage Magnitude = [0.2299932494710589, 0.229
9932501528472, 0.22999325019670214], Phase Angle = [-1.422073
286572225e-6, -120.00000090575658, 119.99999910593104] degree
s
Bus loadbus: Voltage Magnitude = [0.2225213393694953, 0.22672
706378734653, 0.22557738938786964], Phase Angle = [-0.4842383
2883928797, -120.2425346718837, 120.27383544648818] degrees
```

Fig. 2. Sample AC Power Flow Result: Shows the voltage magnitudes and phase angles associated with each phase per line

LinDist3Flow model, utilizing our block diagonal matrix generation in the process. The result of this calculation is then compared against the solver's solution, and the error between the two is calculated to show that the approximation's output is within the acceptable bounds for error (i.e. less than or equal to 2% error).

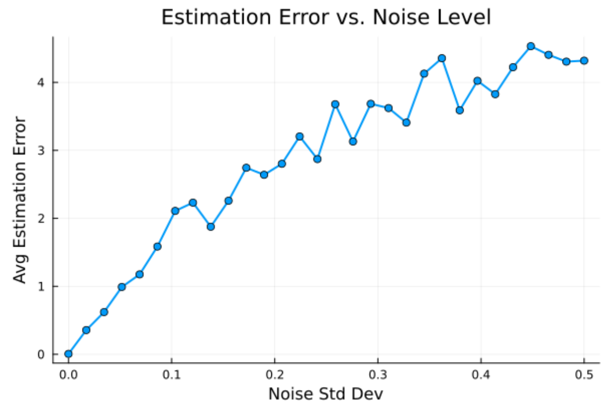

Fig. 3. Error vs number of iterations for the Adam solver [6].

B. Computational results

a) Experiment Description: In this experiment, we evaluated the robustness of the Optimized LinDist3Flow parameter estimation algorithm under varying noise conditions. A synthetic 3-phase, 30-bus radial distribution network was constructed with known resistance and reactance parameters, denoted by θ_{true}. Using these parameters, we generated clean voltage measurements via the LinDist3Flow model. To simulate measurement noise, zero-mean Gaussian noise with 30 increasing standard deviation levels ranging from 0.0 to 0.5 was added to the voltage data. For each noise level, we performed 30 independent trials of parameter estimation using the ADAM optimization algorithm and recorded the ℓ_2-norm error. The average estimation error across trials is plotted as a function of the noise level (see Fig. 3), illustrating the sensitivity and resilience of the optimization framework.

V. Discussion

A spectral sampler algorithm that strategically accounts for electrical distance without assuming specific probability distributions is called the "Shine the Flashlight" approach [7] . This algorithm selects sensors with the highest confidence bounds to sample in the next time step. It is designed to establish concentration inequalities for maximal voltage perturbations under uncertain power injections.

The algorithm employs a closed-form solution, ensuring fixed computational costs at each time step and eliminating the need for an optimization program. It relies on the LinDistFlow approximation and necessitates computable uncertainty bounds for power injections. The methodology provides robust theoretical guarantees and operates with high probability [3]. Notably, it does not presume specific probability distributions, making it applicable to various scenarios in real-time sensor sampling for voltage violation detection [8].

While the proposed framework for grid state estimation shows promise, it is essential to recognize that practical implementation may face challenges. One limitation is the need for extensive real-world data to validate the efficacy of the stochastic LinDistFlow model and the spectral bandit algorithm. Additionally, the computational requirements and potential scalability issues in large-scale distribution grids should be considered as limitations. Addressing these challenges will be crucial in ensuring the successful implementation of the proposed framework in real-world electric distribution systems.

VI. Conclusion

In conclusion, the integration of renewable energy resources into electric distribution grids requires innovative approaches to manage the uncertainties they introduce. The research presents a comprehensive framework for grid state estimation that combines advanced statistical modeling with machine learning techniques. The introduction of a stochastic LinDistFlow model, along with a spectral bandit algorithm, enhances the ability to monitor voltage stability effectively.

Future work will focus on optimizing convergence guarantees for specific algorithm components and improving computational efficiency while continuing to quantify uncertainties within the objective function. It will also focus on developing distributed algorithms that can handle large-scale networks while maintaining computational efficiency. Additionally, advancements in data integration techniques will address challenges posed by asynchronous smart meter readings. Through these advances, our goal is to contribute to more resilient and efficient electric distribution systems capable of adapting to the dynamics of modern energy resources.

References

[1] A. Rangarajan, R. K. Gupta, D. K. Molzahn, and L. A. Roald, "Forecast-Aided State Estimation in Unbalanced Distribution Networks Using Smart Meter Data under Limited Communication Bandwidth," *submitted*, 2024.
[2] R. E. Helou, K. Lee, D. Wu, L. Xie, S. Shakkottai, and V. Subramanian, "OpenGridGym: An Open-Source AI-Friendly Toolkit for Distribution Market Simulation," Mar. 2022. arXiv:2203.04410 [cs, eess].
[3] H. Yue, W. Zhang, U. C. Yilmaz, T. Yildiz, H. Huang, H. Liu, Y. Lin, and A. Abur, "Graph-learning-assisted state estimation using sparse heterogeneous measurements," *Electric Power Systems Research*, vol. 235, p. 110644, Oct. 2024.
[4] B. Taheri, R. K. Gupta, and D. K. Molzahn, "Optimizing parameters of the lindistflow power flow approximation for distribution systems," August 2024.
[5] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. V. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.
[6] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
[7] V. Kekatos, G. B. Giannakis, and B. Wollenberg, "Optimal Placement of Phasor Measurement Units via Convex Relaxation," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1521–1530, 2012.
[8] C. Huang *et al.*, "Smart Meter Pinging and Reading Through AMI Two-Way Communication Networks to Monitor Grid Edge Devices and DERs," *IEEE Transactions on Smart Grid*, pp. 4144–4153, 2022.