ADMM-Based Distributed Linear State Estimation for Large-Scale Interconnected Power Systems

Mohammad Razeghi-Jahromi, *Member, IEEE*
ABB US Corporate Research Center (USCRC), Raleigh, NC, USA
mohammad.razeghi-jahromi@us.abb.com

***Abstract*—State estimation (SE) is a fundamental task in power systems, enabling accurate monitoring of system states to support security-constrained dispatch and control. This paper presents a distributed linear state estimation (D-LSE) approach based on the alternating direction method of multipliers (ADMM). D-LSE is particularly important for ensuring the stability, reliability, and efficiency of large-scale infrastructures such as power grids and communication networks, where traditional centralized SE methods encounter challenges related to scalability, communication bottlenecks, and computational complexity. To overcome these limitations, the proposed framework partitions the system into smaller submodels, with each computational unit estimating local states independently while exchanging only boundary primal (state) and dual (multiplier) variables along with limited system information. The iterative ADMM process refines accuracy and improves coordination across regions, leading to enhanced scalability, reduced communication overhead, and improved fault tolerance. In addition, the framework addresses practical challenges including limited measurement redundancy, dynamic network topologies, and communication delays. The effectiveness of the proposed algorithm is validated on the IEEE 14-bus system, demonstrating its potential to enable resilient, adaptive, and scalable operation of interconnected power systems.**

***Index Terms*—Distributed linear state estimation (D-LSE), alternating direction method of multipliers (ADMM), decomposition of large-scale interconnected systems, weighted least squares (WLS), mean absolute error (MAE).**

I. Introduction

State estimation (SE) plays a pivotal role in modern power system operation by providing reliable information about the underlying system states, which is essential for monitoring, control, and decision-making tasks such as security-constrained dispatch, fault detection, and stability assessment [1]. Conventional SE techniques typically rely on centralized formulations, where measurements collected across the entire network are processed at a single control center. While effective for small- to medium-scale systems, centralized approaches face significant challenges as power networks expand in size and complexity. Increasing penetration of renewable energy resources, the proliferation of distributed energy resources (DERs), and the growing interconnection of regional grids further exacerbate issues of scalability, communication bottlenecks, and computational burden [2]–[5].

To overcome these challenges, distributed state estimation (DSE) has emerged as a promising paradigm, enabling system operators to partition large-scale networks into smaller, more manageable regions [6]–[9]. Each region performs local state estimation using its own measurements and exchanges limited boundary information with neighboring regions. This distributed structure not only alleviates the computational and communication load on a central operator but also improves scalability, robustness, and fault tolerance, which are vital for modern power systems with dynamic and heterogeneous topologies [10], [11]. Additionally, DSE offers enhanced resilience to single-point failures, better compatibility with privacy-preserving requirements between different operating entities, and the flexibility to integrate emerging measurement infrastructures such as phasor measurement units (PMUs) and smart meters.

Among the various solution methodologies for DSE, the alternating direction method of multipliers (ADMM) has attracted considerable attention due to its ability to decompose large optimization problems into parallelizable subproblems while ensuring convergence to globally consistent solutions [12]–[14]. ADMM is particularly well-suited for distributed linear state estimation (D-LSE), as it allows local estimation tasks to be carried out independently, followed by iterative coordination with neighboring regions to enforce consensus on shared variables. This iterative exchange strikes a balance between computational efficiency and communication requirements, making it attractive for real-time applications [15]–[18]. Moreover, ADMM's flexibility allows incorporation of practical constraints such as communication delays, asynchronous updates, and heterogeneous measurement qualities, further extending its applicability to real-world grid operations.

This paper develops and evaluates an ADMM-based D-LSE framework designed for large-scale interconnected power systems. The proposed method enhances scalability by dividing the network into regional submodels, achieves lower communication overhead by restricting data exchange to boundary state and dual variables and a small amount of essential system information, and improves robustness through distributed processing. Furthermore, the framework considers practical challenges such as limited measurement redundancy, system partitioning, and communication delays. Its effectiveness is validated on the IEEE 14-bus system, highlighting high accuracy, convergence properties, and promising potential for extension to larger and more complex networks.

The remainder of this paper is organized as follows. Section II introduces the proposed ADMM-based distributed framework and details its algorithmic formulation. Section III presents numerical results and performance evaluation on the IEEE 14-bus system. Finally, Section IV concludes the paper.

979-8-3315-8138-1

II. Distributed Linear State Estimation: Mathematical Formulation

Distributed linear state estimation approximates the relationship between system measurements and state variables as linear, enabling faster and more scalable computation than nonlinear methods. By decomposing the system into submodels with localized computations, D-LSE reduces communication overhead and allows parallel processing, making it particularly suitable for real-time monitoring, fault detection, phasor measurement unit integration, and adaptive control in large-scale power systems. While its accuracy is limited by linearity assumptions and measurement quality, D-LSE provides a practical, efficient framework for modern smart grid management, supporting timely decision-making and maintaining system stability and reliability.

The DC state estimation problem is commonly formulated as an overdetermined system of linear equations, and solved as weighted least-squares (WLS) problem. In the field of large-scale systems, the state estimation is based on a decomposition technique [10]. For small-scale systems, the model $\mathbf{z} = \mathbf{H}\mathbf{x}+\mathbf{e}$ is applied where $\mathbf{z} \in \mathbb{R}^{n_m}$ is the vector of measurements, $\mathbf{x} \in \mathbb{R}^{n_p}$ is the vector of states, $\mathbf{e} \in \mathbb{R}^{n_m}$ is the vector of random errors, and $\mathbf{H} \in \mathbb{R}^{n_m \times n_p}$ is the Jacobian matrix. As we know, the optimal solution to this (centralized) problem is given by

$$\mathbf{x}^* = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}, \tag{1}$$

where $\mathbf{R} \in \mathbb{R}^{n_m \times n_m}$ is the covariance matrix of measurement errors, i.e., $\mathbf{R} = \mathbb{E}(\mathbf{e}\mathbf{e}^T)$.

For large-scale systems (with large n_m and n_p), the model becomes computationally intensive. In an N-bus system with a measurement redundancy ratio μ (typically $\mu > 1$, around 2), the total number of parameters is $n_p = 2N-1$, consisting of N bus voltage magnitudes and $N-1$ phase angles, with the reference bus angle fixed to an arbitrary value, such as 0. Consequently, $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \in \mathbb{R}^{(2N-1)\times(2N-1)}$. The total number of measurements is given by $n_m = \mu n_p = \mu(2N-1)$. Decomposing $\mathbf{z} = \mathbf{H}\mathbf{x}+\mathbf{e}$ into L submodels yields

$$\underbrace{\begin{bmatrix}\mathbf{z}_1\\ \mathbf{z}_2\\ \vdots\\ \mathbf{z}_L\end{bmatrix}}_{\mathbf{z}\in\mathbb{R}^{n_m}} = \underbrace{\begin{bmatrix}\mathbf{H}_{11} & \mathbf{H}_{12} & \cdots & \mathbf{H}_{1L}\\ \mathbf{H}_{21} & \mathbf{H}_{22} & \cdots & \mathbf{H}_{2L}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{H}_{L1} & \mathbf{H}_{L2} & \cdots & \mathbf{H}_{LL}\end{bmatrix}}_{\mathbf{H}\in\mathbb{R}^{n_m\times n_p}} \underbrace{\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\\ \vdots\\ \mathbf{x}_L\end{bmatrix}}_{\mathbf{x}\in\mathbb{R}^{n_p}} + \underbrace{\begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\\ \vdots\\ \mathbf{e}_L\end{bmatrix}}_{\mathbf{e}\in\mathbb{R}^{n_m}}, \tag{2}$$

where

$$\mathbf{z}_i = \sum_{j=1}^{L}\mathbf{H}_{ij}\mathbf{x}_j + \mathbf{e}_i, \tag{3}$$

and $\mathbf{z}_i,\, \mathbf{e}_i \in \mathbb{R}^{n_{mi}}$, $\mathbf{x}_i \in \mathbb{R}^{n_{pi}}$, $\mathbf{H}_{ij} \in \mathbb{R}^{n_{mi}\times n_{pj}}$, $\sum_{j=1}^{L} n_{pj} = n_p$, $\sum_{j=1}^{L} n_{mj} = n_m$ for all $i = 1,\cdots,L$. The parameter vector $\mathbf{x} \in \mathbb{R}^{n_p}$ in decomposed form is $\mathbf{x}^T = [\mathbf{x}_1^T\ \mathbf{x}_2^T\ \cdots\ \mathbf{x}_L^T]$.

The least squares estimation is obtained by minimizing

$$J = \sum_{i=1}^{L}\mathbf{e}_i^T\mathbf{R}_i^{-1}\mathbf{e}_i, \tag{4}$$

with respect to $\mathbf{x}_i$ where $\mathbf{R}_i \in \mathbb{R}^{n_{m_i}\times n_{m_i}}$ is the covariance matrix of submodel i measurement errors, i.e., $\mathbf{R}_i = \mathbb{E}(\mathbf{e}_i\mathbf{e}_i^T)$. Furthermore, try to find $\mathbf{x}_i$ in such a way that the interaction between the L submodels is minimized. Thus, an interaction variable $\mathbf{w}_i \in \mathbb{R}^{n_{mi}}$

$$\mathbf{w}_i := \sum_{j\neq i}^{L}\mathbf{H}_{ij}\mathbf{x}_j, \tag{5}$$

is defined, providing the influence from all subsystems $j \neq i$ to the subsystem i. Thus, we can rewrite equation (3) as

$$\mathbf{z}_i = \mathbf{H}_{ii}\mathbf{x}_i + \mathbf{w}_i + \mathbf{e}_i. \tag{6}$$

Following equation (4)

$$J = \sum_{i=1}^{L}(\mathbf{z}_i - \mathbf{H}_{ii}\mathbf{x}_i - \mathbf{w}_i)^T\mathbf{R}_i^{-1}(\mathbf{z}_i - \mathbf{H}_{ii}\mathbf{x}_i - \mathbf{w}_i), \tag{7}$$

has to be minimized with respect to $\mathbf{x}_i$ and $\mathbf{w}_i$ within the constraint of equation (5), i.e.,

$$\begin{aligned}&\{\hat{\mathbf{x}}_i, \hat{\mathbf{w}}_i\} :=\\ &\arg\min_{\mathbf{x}_i,\mathbf{w}_i}\ \sum_{i=1}^{L}\left(\mathbf{z}_i - \mathbf{H}_{ii}\mathbf{x}_i - \mathbf{w}_i\right)^T\mathbf{R}_i^{-1}\left(\mathbf{z}_i - \mathbf{H}_{ii}\mathbf{x}_i - \mathbf{w}_i\right),\\ &\text{subject to}\quad \mathbf{w}_i = \sum_{j\neq i}^{L}\mathbf{H}_{ij}\mathbf{x}_j.\end{aligned} \tag{8}$$

ADMM is a distributed optimization method that combines dual decomposition and augmented Lagrangian principles to efficiently solve large-scale problems by decomposing them into smaller subproblems [12]. ADMM is especially useful in scenarios where the problem can be divided among multiple processors or agents, each solving part of the overall problem.

In ADMM, the augmented Lagrangian is formed using a penalty parameter $\rho > 0$, which weights the constraint violation in the optimization, and a Lagrange multiplier (dual variable) $\boldsymbol{\lambda}_i \in \mathbb{R}^{n_{mi}}$, which ensures that the constraints are satisfied.

$$\begin{aligned}&\mathcal{L}_\rho(\mathbf{x}_i, \mathbf{w}_i, \boldsymbol{\lambda}_i) =\\ &\sum_{i=1}^{L}\Bigg(\left(\mathbf{z}_i - \mathbf{H}_{ii}\mathbf{x}_i - \mathbf{w}_i\right)^T\mathbf{R}_i^{-1}\left(\mathbf{z}_i - \mathbf{H}_{ii}\mathbf{x}_i - \mathbf{w}_i\right)\\ &\qquad + \boldsymbol{\lambda}_i^T\Big(\mathbf{w}_i - \sum_{j\neq i}^{L}\mathbf{H}_{ij}\mathbf{x}_j\Big) + \frac{\rho}{2}\Big\|\mathbf{w}_i - \sum_{j\neq i}^{L}\mathbf{H}_{ij}\mathbf{x}_j\Big\|^2\Bigg).\end{aligned} \tag{9}$$

Alternating direction method of multipliers consists of the iterations

$$\begin{aligned}\mathbf{x}_i^{(k+1)} &:= \arg\min_{\mathbf{x}_i}\ \mathcal{L}_\rho\left(\mathbf{x}_i, \mathbf{w}_i^{(k)}, \boldsymbol{\lambda}_i^{(k)}\right),\\ \mathbf{w}_i^{(k+1)} &:= \arg\min_{\mathbf{w}_i}\ \mathcal{L}_\rho\left(\mathbf{x}_i^{(k+1)}, \mathbf{w}_i, \boldsymbol{\lambda}_i^{(k)}\right),\\ \boldsymbol{\lambda}_i^{(k+1)} &:= \boldsymbol{\lambda}_i^{(k)} + \rho\Bigg(\mathbf{w}_i^{(k+1)} - \sum_{j\neq i}^{L}\mathbf{H}_{ij}\mathbf{x}_j^{(k+1)}\Bigg).\end{aligned} \tag{10}$$

979-8-3315-8138-1

Differentiating $\mathcal{L}_\rho(\cdot)$ with respect to $\mathbf{x}_i$ and $\mathbf{w}_i$ for all $i = 1, \cdots, L$ yields

$$\hat{\mathbf{x}}_i^{(k+1)} = \left(2\mathbf{H}_{ii}^T\mathbf{R}_i^{-1}\mathbf{H}_{ii} + \rho\sum_{k\neq i}^{L}\mathbf{H}_{ki}^T\mathbf{H}_{ki}\right)^{-1} \times \left(2\mathbf{H}_{ii}^T\mathbf{R}_i^{-1}\left(\mathbf{z}_i - \hat{\mathbf{w}}_i^{(k)}\right) + \sum_{j\neq i}^{L}\mathbf{H}_{ji}^T\left(\boldsymbol{\lambda}_j^{(k)} + \rho\hat{\mathbf{w}}_j^{(k)}\right) - \frac{\rho}{2}\sum_{k\neq i\neq l}^{L}\mathbf{H}_{ki}^T\mathbf{H}_{kl}\hat{\mathbf{x}}_l^{(k)}\right), \tag{11}$$

$$\hat{\mathbf{w}}_i^{(k+1)} = \left(2\mathbf{R}_i^{-1} + \rho\mathbf{I}_{n_{m_i}}\right)^{-1}\left(2\mathbf{R}_i^{-1}\left(\mathbf{z}_i - \mathbf{H}_{ii}\hat{\mathbf{x}}_i^{(k+1)}\right) - \boldsymbol{\lambda}_i^{(k)} + \rho\sum_{j\neq i}^{L}\mathbf{H}_{ij}\hat{\mathbf{x}}_j^{(k+1)}\right), \tag{12}$$

$$\boldsymbol{\lambda}_i^{(k+1)} = \boldsymbol{\lambda}_i^{(k)} + \rho\left(\hat{\mathbf{w}}_i^{(k+1)} - \sum_{j\neq i}^{L}\mathbf{H}_{ij}\hat{\mathbf{x}}_j^{(k+1)}\right), \tag{13}$$

respectively, where $\mathbf{I}_{n_{m_i}} \in \mathbb{R}^{n_{m_i}\times n_{m_i}}$ is an identity matrix. A suitable initial point for $\hat{\mathbf{x}}_i^{(0)}$, $\hat{\mathbf{w}}_i^{(0)}$, and $\boldsymbol{\lambda}_i^{(0)}$ is a flat start, where all bus magnitudes are set to one, and all bus angles are set to zero. Let N_i represent the number of buses in submodel i. Therefore, $n_{pi} = 2N_i - 1$. Using this flat start approach, we initialize all $\mathbf{x}_i = [\theta_2\ \theta_3\ \cdots\ \theta_{N_i}\ V_1\ V_2\ \cdots\ V_{N_i}]^T$ for $i = 1, \cdots, L$ as

$$\hat{\mathbf{x}}_i^{(0)} = \begin{bmatrix}\mathbf{0}_{N_i-1}\\ \mathbf{1}_{N_i}\end{bmatrix},$$
$$\hat{\mathbf{w}}_i^{(0)} = \sum_{j\neq i}^{L}\mathbf{H}_{ij}\hat{\mathbf{x}}_j^{(0)},$$
$$\boldsymbol{\lambda}_i^{(0)} = \rho\left(\hat{\mathbf{w}}_i^{(0)} - \sum_{j\neq i}^{L}\mathbf{H}_{ij}\hat{\mathbf{x}}_j^{(0)}\right) = \mathbf{0}, \tag{14}$$

where $\mathbf{0}_{N_i-1}$ and $\mathbf{1}_{N_i}$ are vectors consisting entirely of zeros and ones of size $N_i - 1$ and N_i respectively.

The following algorithm outlines the iterative process for obtaining the solution to distributed linear state estimation, i.e., $\hat{\mathbf{x}}_i$ for all $i = 1, \cdots, L$. The state and dual variables are first initialized using (14), after which the algorithm proceeds by iteratively updating the state and dual variables according to (11)-(13) until the stopping criteria are satisfied.

Algorithm 1 Distributed linear state estimation (D-LSE)

initialize
Set $k = 0$ and initialize $\hat{\mathbf{x}}_i^{(0)}$, $\hat{\mathbf{w}}_i^{(0)}$, and $\boldsymbol{\lambda}_i^{(0)}$ as (14)
repeat
Update $\hat{\mathbf{x}}_i^{(k+1)}$ as (11): $\mathbf{x}$-minimization
Update $\hat{\mathbf{w}}_i^{(k+1)}$ as (12): $\mathbf{w}$-minimization
Update $\boldsymbol{\lambda}_i^{(k+1)}$ as (13): Dual update
until
The stopping criteria $\frac{\left\|\hat{\mathbf{x}}_i^{(k+1)} - \hat{\mathbf{x}}_i^{(k-1)}\right\|}{\left\|\hat{\mathbf{x}}_i^{(k)}\right\|} \leq \varepsilon \ll 1$ is met

ADMM is globally convergent for convex problems and can achieve linear convergence rates in strongly convex settings, while for nonconvex problems it may converge to a stationary point under certain conditions, though guarantees are weaker [12]. Its practical performance depends on the penalty parameter ρ and the problem decomposition. ADMM's simplicity, scalability, and flexibility have made it popular in applications such as machine learning, image processing, and distributed optimization. In D-LSE for large-scale power systems, ADMM enables each region to solve a smaller localized problem, supports parallel computation for faster state estimation, and minimizes communication by sharing only boundary state and multiplier variables with minimal system information. The distributed structure improves robustness, allowing regions to operate independently during failures, and its iterative nature provides frequent updates for real-time monitoring. Challenges include convergence speed sensitivity to ρ, communication delays, careful parameter tuning, and the reliance on accurate boundary measurements for overall estimation quality.

III. Numerical Results

A. Simulation Setup

The simulation set-up for numerically validating the D-LSE algorithm follows a two-step process [19]. First, a detailed single-phase circuit model incorporating the shunt admittances of the cables is used in MATLAB to perform load flow studies. Second, historical smart meter data spanning one year with a 5-minute time step, was used as input for the load flow analysis. We utilized historical load data from the New York ISO (https://www.nyiso.com/) and applied a long short-term memory (LSTM) network for load forecasting, which was then input into the state estimation algorithm. To simulate real-world conditions, Gaussian noise components corresponding to a 10% error were added to the forecasted values. Using the fact that in a Gaussian distribution $\pm 3\sigma$ deviation about the mean covers more than 99.7% area of the Gaussian curve, the standard deviation of forecasted measurement was computed as

$$\sigma_{z_i} = \frac{z_{t_i} \times \%error}{3 \times 100} = \frac{z_{t_i}}{30}, \tag{15}$$

where z_{t_i} is the forecasted value of the pseudo-measurement z_i. In practice, the standard deviation of real measurements is derived from sensor characteristics, and is typically small, whereas pseudo-measurements correspond to higher-variance load estimates.

B. Simulation Results

This section demonstrates the effectiveness of the proposed D-LSE method on the IEEE 14-bus test system, which is partitioned into two submodels. Submodel 1 estimates 13 phase angles, i.e., $\mathbf{x}_1 \in \mathbb{R}^{13}$, while submodel 2 estimates 14 voltage magnitudes, i.e., $\mathbf{x}_2 \in \mathbb{R}^{14}$. A Monte Carlo study with 300 runs is conducted, where in each run a subset of 82 measurements is randomly selected. These include a total of 51 measurements, with $\mathbf{z}_1 \in \mathbb{R}^{26}$ and $\mathbf{z}_2 \in \mathbb{R}^{25}$, consisting of one actual voltage magnitude measurement ($|V_n|$)

and 50 pseudo-measurements (P_n, Q_n, P_{mn}, and Q_{mn}). As an illustration, Simulation 2 out of 300 uses the following selected measurements

- P_n: 10 out of 14 real power injection at bus 1, 2, 3, 4, 5, 6, 8, 10, 11, 12;
- Q_n: 10 out of 14 reactive power injection at bus 1, 2, 3, 4, 5, 6, 7, 8, 10, 13;
- P_{mn}: 15 out of 20 real power flow from bus 1, 6, 6, 9, 10, 4, 4, 13, 9, 6, 7, 1, 2, 12, 2 to bus 2, 12, 11, 14, 11, 7, 9, 14, 10, 13, 9, 5, 5, 13, 3 respectively
- Q_{mn}: 15 out of 20 reactive power flow from bus 1, 2, 9, 6, 1, 2, 4, 3, 7, 4, 5, 13, 6, 7, 4 to bus 2, 4, 14, 11, 5, 5, 5, 4, 9, 7, 6, 14, 12, 8, 9 respectively; and
- $|V_n|$: 1 out of 14 voltage magnitude at bus 7.

This selection results in a redundancy ratio of approximately $51/27 \approx 1.89$. Measurement variances were set as $\sigma_v^2 = 9 \times 10^{-4}$, $\sigma_p^2 = \sigma_q^2 = 10^{-4}$, and $\sigma_{p_f}^2 = \sigma_{q_f}^2 = 64 \times 10^{-6}$.

A 10% error is introduced by adding an independent, standard normally distributed random variable to the output of the power flow, treating all p_i, q_i, p_{f_i}, and q_{f_i} as pseudo-measurements. In a worst-case scenario, only one real measurement v_i is used, resulting in a pseudo-measurement usage of $(10+10+15+15)/(1+10+10+15+15) = 50/51 \approx 98\%$. Additionally, the variance of all pseudo-measurements is increased by a factor of 30, as discussed in (15). The following performance metrics are then evaluated:

- Voltage mean absolute error: $\mathbb{E}(|\mathbf{v}^* - \hat{\mathbf{v}}|)$ p.u.
- Angle mean absolute error: $\mathbb{E}(|\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}|)$ degree

where $\mathbf{x}^{*T} = [\boldsymbol{\theta}^{*T} \ \mathbf{v}^{*T}]$ denote the results of the centralized LSE in (1), and $\hat{\mathbf{x}}^T = [\hat{\boldsymbol{\theta}}^T \ \hat{\mathbf{v}}^T]$ represent the outcomes of the proposed D-LSE described in Algorithm 1. The numerically computed Jacobian matrix $\mathbf{H} \in \mathbb{R}^{51\times27}$ is partitioned as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix} = \begin{bmatrix} 26\times13 & 26\times14 \\ 25\times13 & 25\times14 \end{bmatrix}. \quad (16)$$

The results of the proposed distributed linear state estimator are compared against the centralized linear state estimator given in (1). Figure 1 summarizes the Monte Carlo simulation outcomes, demonstrating that the mean absolute errors for both phase angles and voltage magnitudes remain consistently low across all scenarios. Each marker represents the result of an individual simulation run, while the red line denotes the average performance over the 300 trials, highlighting the overall robustness of the proposed approach. On average, the mean absolute error is 0.1631° for phase angles and 0.0057 p.u. for voltage magnitudes over all trials. For a representative case (Simulation 2, using the measurement set described earlier), Figure 2 compares the estimated phase angles and voltage magnitudes at all buses obtained from the proposed distributed method (black markers) and the centralized solution (red markers). The close agreement between the two confirms that the distributed estimator converges to the optimal centralized solution. Figure 3 illustrates the convergence behavior of the algorithm based on the stopping criterion defined in Algorithm 1, for a tolerance of $\varepsilon = 0.005$. This criterion measures the change in the state variables between successive iterations. The convergence trends for both submodels indicate that the stopping conditions are satisfied within 300 iterations. Complementarily, Figure 4 depicts the evolution of the dual variables $\boldsymbol{\lambda}_1 \in \mathbb{R}^{26}$ and $\boldsymbol{\lambda}_2 \in \mathbb{R}^{25}$ over the same iteration horizon with $\rho = 25$. After approximately 250 iterations, the updates of the dual variables become negligible, which is consistent with the stabilization observed in Figure 3.

All algorithmic parameters, including $\varepsilon = 0.005$ and $\rho = 25$, are kept fixed across all simulation runs to ensure a fair and consistent evaluation.

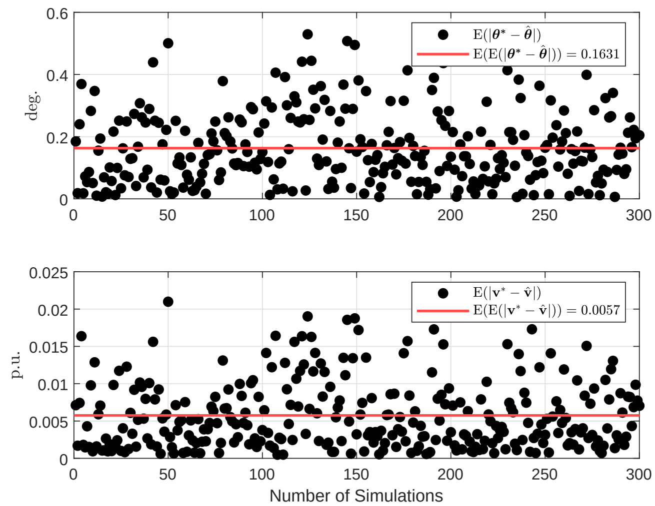

Fig. 1: Mean absolute errors for phase angles and voltage magnitudes between the proposed D-LSE and centralized LSE in the IEEE 14-bus system.

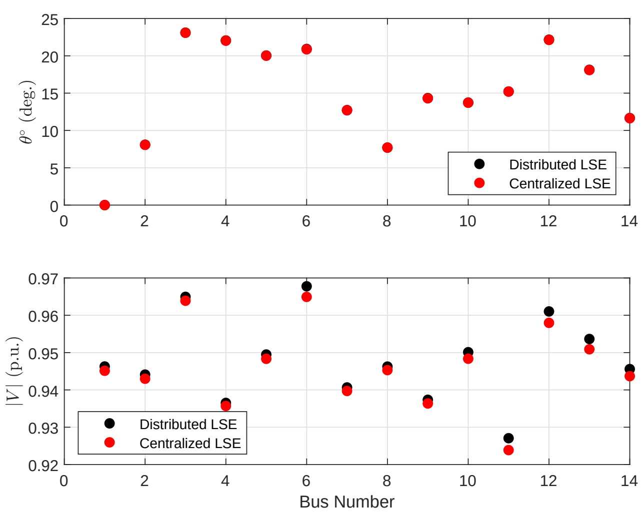

Fig. 2: A comparison of phase angle and voltage magnitude estimations between the proposed D-LSE and the centralized LSE for Simulation 2.

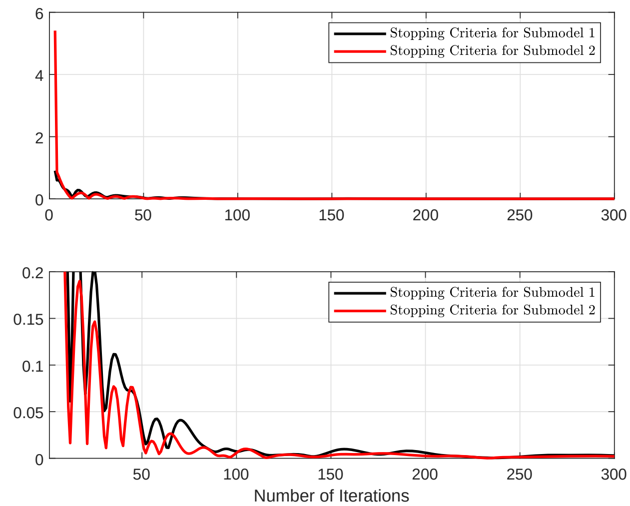

Fig. 3: Stopping criteria for submodels 1 and 2 using the proposed D-LSE method in Simulation 2.

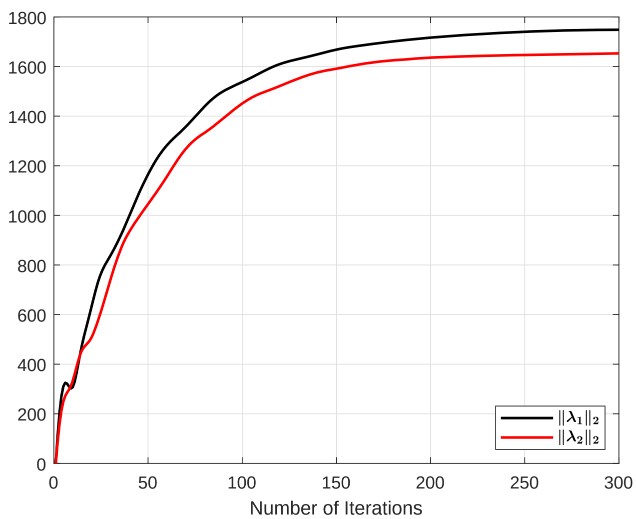

Fig. 4: Convergence behavior of the dual variables in the D-LSE estimation of phase angles and voltage magnitudes in Simulation 2.

IV. Conclusion

This paper presents a distributed linear state estimation framework for large-scale interconnected power systems, utilizing the alternating direction method of multipliers. The proposed approach partitions the network into smaller submodels, enabling localized state estimation through the exchange of only boundary primal (state) and dual (multiplier) variables along with limited system information. This design effectively addresses the scalability and communication bottlenecks of centralized methods. The iterative coordination facilitated by ADMM improves estimation accuracy, robustness, and fault tolerance. Validation on the IEEE 14-bus system demonstrates the effectiveness of the method and its capability to manage practical challenges such as measurement redundancy, communication delays, and evolving system topologies. Overall, the results highlight the potential of ADMM-based D-LSE as a resilient, adaptive, and scalable solution for next-generation power system monitoring and control.

References

[1] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.

[2] J. Vijaychandra, B. R. V. Prasad, V. K. Darapureddi, B. V. Rao, and Ł. Knypiński, "A review of distribution system state estimation methods and their applications in power systems," *Electronics*, vol. 12, no. 3, p. 603, 2023.

[3] K. Nainar and F. Iov, "Smart meter measurement-based state estimation for monitoring of low-voltage distribution grids," *Energies*, vol. 13, no. 20, p. 5367, 2020.

[4] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875–3883, 2016.

[5] F. Ahmad, A. Rasool, E. Ozsoy, R. Sekar, A. Sabanovic, and M. Elitaş, "Distribution system state estimation-a step towards smart grid," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 2659–2671, 2018.

[6] D. M. Falcao, F. F. Wu, and L. Murphy, "Parallel and distributed state estimation," *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 724–730, 2002.

[7] L. Xie, D.-H. Choi, S. Kar, and H. V. Poor, "Fully distributed state estimation for wide-area monitoring systems," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1154–1169, 2012.

[8] S. Kar, G. Hug, J. Mohammadi, and J. M. Moura, "Distributed state estimation and energy management in smart grids: A consensus innovations approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1022–1038, 2014.

[9] G. N. Korres, "A distributed multiarea state estimation," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 73–84, 2010.

[10] H. Karimipour and V. Dinavahi, "Parallel domain decomposition based distributed state estimation for large-scale power systems," in *2015 IEEE/IAS 51st Industrial & Commercial Power Systems Technical Conference (I&CPS)*. IEEE, 2015, pp. 1–5.

[11] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6137–6147, 2019.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[13] Y. Wang, W. Yin, and J. Zeng, "Global convergence of admm in nonconvex nonsmooth optimization," *Journal of Scientific Computing*, vol. 78, no. 1, pp. 29–63, 2019.

[14] J. Bai, J. Li, F. Xu, and H. Zhang, "Generalized symmetric admm for separable convex optimization," *Computational Optimization and Applications*, vol. 70, no. 1, pp. 129–170, 2018.

[15] Y. Du, W. Zhang, and T. Zhang, "Admm-based distributed state estimation for integrated energy system," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 2, pp. 275–283, 2019.

[16] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, 2012.

[17] S. Parsegov, S. Kubentayeva, E. Gryazina, A. Gasnikov, and F. Ibáñez, "Admm-based distributed state estimation for power systems: Evaluation of performance," *IFAC-PapersOnLine*, vol. 53, no. 5, pp. 182–188, 2020.

[18] N. Aljohani, T. Zou, A. S. Bretas, and N. G. Bretas, "Multi-area state estimation: A distributed quasi-static innovation-based model with an alternative direction method of multipliers," *Applied Sciences*, vol. 11, no. 10, p. 4419, 2021.

[19] M. Razeghi-Jahromi, M. Choobineh, R. Rodrigues, and A. Melese, "State estimation for ac distribution systems using an lstm network for pseudo-measurement modeling," in *2025 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2025.

979-8-3315-8138-1