Robust and Scalable Federated AC WLS State Estimation via Decentralised Consensus

Adnan Anwar
School of Information Technology, Deakin University, Melbourne, Australia
Email: adnan.anwar@deakin.edu.au

***Abstract*—Widely used centralised weighted least squares (WLS) estimators face scalability and privacy challenges in large, and distributed networks. This paper proposes a fully decentralised federated AC WLS framework that performs peer-to-peer state estimation without a central coordinator. Two consensus mechanisms (Metropolis and Gossip) are employed to achieve robust information aggregation under stochastic link failures. An adaptive damped Gauss–Newton update with the nearest symmetric positive-definite (SPD) projection ensures numerical stability and convergence. We conduct extensive experiments on the IEEE 30-bus, IEEE 118-bus, and IEEE 300-bus systems that show the proposed framework achieves accuracy comparable to centralised WLS while reducing communication overhead. The results demonstrate that the proposed fully decentralised federated learning offers a scalable, communication-efficient, data-local and privacy-aware solution for next-generation AC state estimation in large interconnected power grids.**

***Index Terms*—federated learning, WLS, state estimation**

I. Introduction

Accurate estimation of system states is crucial for the secure and economic operation of modern electric power grids. The alternating current (AC) weighted least squares (WLS) method has been the industry standard for handling nonlinear measurement models and network losses for operational decision-making [1]. With the integration of renewable generation, distributed energy resources, and cross-regional interconnections, the traditional centralised state estimation (SE) paradigm has experienced increasing challenges in computation, privacy, and communication [2]. Besides, centralised architectures require extensive data transfer to a central control centre, exposing critical infrastructure to latency and cybersecurity risks.

Recent advances in distributed and hierarchical SE frameworks have partially addressed these issues by partitioning networks into regions and performing coordinated estimation. However, such approaches typically rely on supervisory layers or synchronous exchanges, limiting their resilience under communication failures. Emerging federated learning concepts offer a new direction by enabling collaborative estimation across regions while preserving local data privacy [3]. For example, the authors' earlier work demonstrated a Byzantine-resilient federated learning framework for energy system load forecasting, which safeguarded client data and model integrity through gradient quantisation and secure aggregation [4]. However, this framework relied on a centralised FL setup where a coordinating server aggregates model updates, which poses the possibility of a single point of failure and potential information bottleneck. In contrast, a fully decentralised federated learning paradigm eliminates this dependency, allowing regions to communicate directly and reach consensus autonomously. The measurement information and system Jacobian also remain private and confidential to each region.

For the state estimation problem, a fundamental gap remains in realising a fully decentralised, communication-efficient, and numerically stable SE framework for nonlinear AC systems using a federated setup. Current methods often struggle with ill-conditioned Jacobians, loss of positive definiteness in normal matrices, and degraded convergence when inter-regional coupling or packet losses occur. Furthermore, empirical validations on large networks are limited, leaving the computational–communication trade-offs largely unexplored. This motivates a federated AC WLS formulation that couples detailed system modelling with adaptive numerical stabilisation and high-performance consensus in a decentralised setup.

1) Related Work: Distributed state estimation (SE) methods have been widely studied to address the scalability limitations of centralised WLS estimation in large power systems [7]–[9]. Classical multi-area SE frameworks perform local Gauss–Newton updates and exchange boundary variables or shared states via supervisory coordination, which introduces area ownership, synchronisation requirements, and potential single points of failure under communication uncertainty.

ADMM-based decentralised SE formulations enforce consensus through augmented Lagrangian and primal–dual updates [8]. While effective, such approaches require explicit exchange of primal and dual variables and are sensitive to synchronisation and unreliable communication. Federated learning (FL) has recently been adopted in power and energy applications to enable collaborative learning while keeping raw data local, typically using a central aggregator [4]. Although communication-efficient, server-based FL architectures introduce coordination bottlenecks and single points of failure [10].

In contrast, the proposed framework enables fully decentralised federated state estimation through peer-to-peer consensus on sufficient statistics. Regions exchange only local Gauss–Newton normal-equation components, without sharing raw measurements, boundary states, or Lagrange multipliers. Each region independently reconstructs the global update and estimates the full system state, eliminating area ownership and

This work is supported by the Deakin Mini ARC Analog Program. ChatGPT-5 is used to assist with language editing. All conceptual content and technical contributions are solely of the author.

979-8-3315-8138-1

ADMM-style coordination. Here, federated refers to the decentralised architecture with data locality and peer autonomy rather than stochastic gradient-based model training.

2) Contribution: This paper proposes a *Decentralised Federated AC Weighted Least Squares-based State Estimation (Fed-DSE)* framework that eliminates central coordination and enables peer-to-peer cooperation among regional estimators. The method integrates two powerful and widely accepted consensus mechanisms (Metropolis and Gossip) for decentralised information fusion. For numerical stability, the proposed method also employs an adaptive Levenberg–Marquardt damping along with a robust symmetric positive-definite (SPD) projection, which ensures stability of Gauss–Newton updates under stochastic link failures. Comprehensive experiments on IEEE 30-bus, IEEE 118-bus, and IEEE 300-bus systems show that the proposed approach attains precision while improving scalability, robustness, and resilience to communication uncertainty in large-scale power grids. The contributions are:

- We formulate a *federated AC WLS* estimator that achieves fully decentralised, data-local and privacy-aware state estimation without any central coordinator.
- We develop a *fully decentralised framework* employing high-performing consensus mechanisms (Metropolis and Gossip) [11] with adaptive damping and SPD-projected Gauss–Newton updates for numerical robustness.
- We conduct *extensive experimental evaluation* on IEEE 30, IEEE 118, and IEEE 300-bus networks, demonstrating scalability, comparable accuracy to the central aggregator, and superior communication efficiency under stochastic link losses.

II. System and Measurement Model

Modern power grids are large, geographically dispersed, and operated by multiple utilities. Consider an n_b-bus AC system with complex bus voltages $V_i = V_{m,i}e^{j\theta_i}$, where $V_{m,i}$ and θ_i denote the magnitude and phase angle at bus i, respectively. θ_{nonref} is the phase angle of slack bus. The system state is

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\theta}_{\text{nonref}} \\ \mathbf{V}_m \end{bmatrix} \in \mathbb{R}^n, \qquad n = (n_b - 1) + n_b. \tag{1}$$

The nonlinear AC measurement model is [1]

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e}, \qquad \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \tag{2}$$

where $\mathbf{z}$ stacks active/reactive power injections, branch flows, and bus voltages (V_m). The noise covariance $\mathbf{R}$ defines the weighting $\mathbf{W} = \mathbf{R}^{-1}$.

For decentralised operation, the network is partitioned into K regions (agents), each with its own SCADA system and local measurements $\mathbf{z}_k$, noise covariance $\mathbf{R}_k$, and nonlinear function $\mathbf{h}_k(\mathbf{x})$. The global model is thus

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_K \end{bmatrix}, \quad \mathbf{R} = \operatorname{blkdiag}(\mathbf{R}_1, \ldots, \mathbf{R}_K), \quad \mathbf{h}(\mathbf{x}) = \begin{bmatrix} \mathbf{h}_1(\mathbf{x}) \\ \vdots \\ \mathbf{h}_K(\mathbf{x}) \end{bmatrix}. \tag{3}$$

A. Centralised Weighted Least Squares SE

A central control centre estimates the most probable state by minimizing the weighted least-squares cost [3]:

$$\min_{\mathbf{x}} J(\mathbf{x}) = \tfrac{1}{2}(\mathbf{z} - \mathbf{h}(\mathbf{x}))^\top \mathbf{W}(\mathbf{z} - \mathbf{h}(\mathbf{x})) = \sum_{k=1}^{K} J_k(\mathbf{x}), \tag{4}$$

with each regional contribution

$$J_k(\mathbf{x}) = \tfrac{1}{2}(\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}))^\top \mathbf{W}_k(\mathbf{z}_k - \mathbf{h}_k(\mathbf{x})). \tag{5}$$

Each region can thus compute its gradient and Jacobian locally, reflecting the modular structure of multi-area grids.

B. Federated SE

Linearising $\mathbf{h}_k(\mathbf{x})$ around the current estimate $\mathbf{x}^{(t)}$ gives

$$\mathbf{r}_k = \mathbf{z}_k - \mathbf{h}_k(\mathbf{x}^{(t)}), \qquad \mathbf{H}_k = \frac{\partial \mathbf{h}_k}{\partial \mathbf{x}}\Big|_{\mathbf{x}^{(t)}}, \tag{6}$$

Therefore, the local normal equations become

$$\mathbf{A}_k = \mathbf{H}_k^\top \mathbf{W}_k \mathbf{H}_k, \qquad \mathbf{b}_k = \mathbf{H}_k^\top \mathbf{W}_k \mathbf{r}_k. \tag{7}$$

A centralised federated SE aggregates these summaries to perform the global Gauss–Newton update:

$$\mathbf{A}\,\Delta\mathbf{x} = \mathbf{b}, \quad \mathbf{A} = \sum_{k=1}^{K} \mathbf{A}_k, \ \mathbf{b} = \sum_{k=1}^{K} \mathbf{b}_k. \tag{8}$$

Although this development reduces raw data exchange, the central aggregator still forms a single point of failure and raises privacy concerns. These challenges motivate us to develop a peer-to-peer, consensus-based, decentralised SE formulation.

III. Federated Consensus Learning for Decentralised SE (DSE)

In the decentralised approach, regions exchange only minimal information with their neighbours over a communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the set of regions and $\mathcal{E}$ represents bidirectional communication links, such as secure inter-utility data channels following IEC 61850 message exchanges. Each region k constructs a local information vector:

$$\mathbf{s}_k = \begin{bmatrix} \operatorname{vecu}(\mathbf{A}_k) \\ \mathbf{b}_k \end{bmatrix} \in \mathbb{R}^d, \qquad d = \tfrac{n(n+1)}{2} + n, \tag{9}$$

where $\operatorname{vecu}(\cdot)$ represents the upper-triangular elements of a symmetric matrix. d is the total dimension of the local information vector each region shares in consensus. The goal of consensus is for all regions to compute the global average:

$$\bar{\mathbf{s}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{s}_k, \tag{10}$$

It corresponds to the averaged normal equation components $\bar{\mathbf{A}} = \frac{1}{K}\sum_k \mathbf{A}_k$ and $\bar{\mathbf{b}} = \frac{1}{K}\sum_k \mathbf{b}_k$. Each region reconstructs

$$\mathbf{A} \approx K\,\bar{\mathbf{A}}, \qquad \mathbf{b} \approx K\,\bar{\mathbf{b}}, \tag{11}$$

and independently solves for $\Delta\mathbf{x}$ without central coordinator.

979-8-3315-8138-1

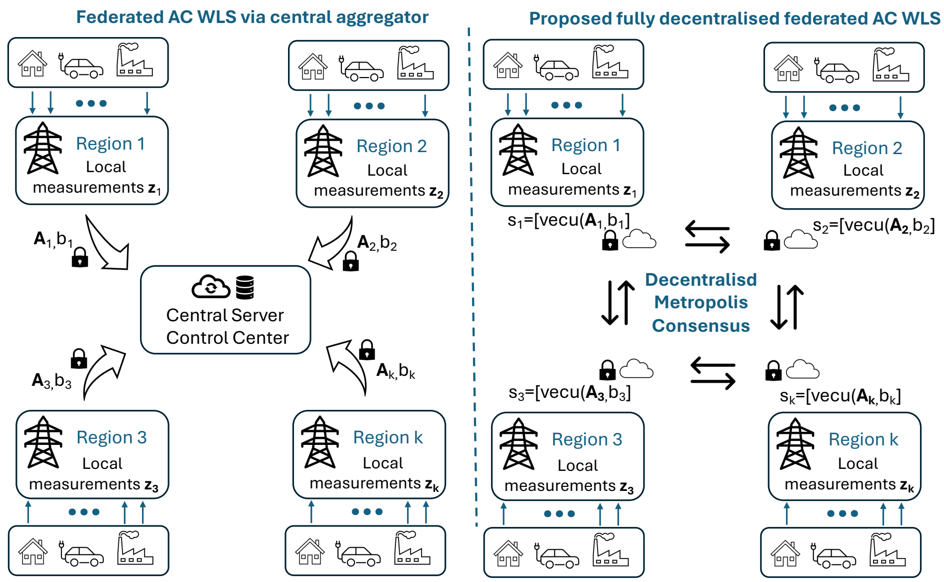

Fig. 1: Comparison between centralised and decentralised federated AC WLS frameworks.

A. Metropolis Consensus-based Federated DSE

The Metropolis consensus algorithm enables synchronous peer-to-peer averaging with topology-dependent weights and guarantees deterministic convergence over connected graphs [5]. Each node assigns weights based on local node degrees, forming a symmetric, row-stochastic matrix that ensures stable information aggregation without central coordination.

At each consensus round τ, node k updates its local information vector by averaging its own value with those received from active neighbours,

$$\mathbf{s}_k^{(\tau+1)} = w_{kk}\mathbf{s}_k^{(\tau)} + \sum_{j\in\mathcal{N}_k\cap\mathcal{A}_k^{(\tau)}} w_{kj}\mathbf{s}_j^{(\tau)}, \tag{12}$$

where $\mathcal{N}_k$ denotes the neighbour set and $\mathcal{A}_k^{(\tau)}$ the set of active communication links, each subject to independent packet drops with probability p_{drop}. Consensus iterations terminate once neighbouring values agree within a prescribed tolerance or a maximum number of rounds is reached. Upon convergence, all regions obtain a common estimate of the aggregated sufficient statistics, enabling identical local Gauss–Newton updates.

B. Gossip Consensus-based Federated DSE

The Gossip consensus algorithm relies on random pairwise information exchange instead of synchronous weighted averaging, making it lightweight and asynchronous [6]. At each local iteration g, an edge (i, j) is randomly selected from $\mathcal{E}$. If the communication link is active (with probability $1-p_{\mathrm{drop}}$), the two participating regions perform pairwise averaging:

$$\mathbf{s}_i^{(g+1)} = \tfrac{1}{2}\big(\mathbf{s}_i^{(g)} + \mathbf{s}_j^{(g)}\big), \tag{13}$$

$$\mathbf{s}_j^{(g+1)} = \tfrac{1}{2}\big(\mathbf{s}_i^{(g)} + \mathbf{s}_j^{(g)}\big). \tag{14}$$

All other regions retain their previous values. Through repeated random interactions, all regions gradually converge to the global average $\bar{\mathbf{s}}$.

Both consensus mechanisms achieve convergence under a connected communication graph, though with distinct dynamics. Metropolis consensus offers deterministic and topology-dependent convergence rates. On the other hand, gossip consensus converges with speed proportional to the frequency of pairwise interactions. Fig. 1 illustrates the working philosophies between centralised and decentralised federated AC WLS frameworks. In the proposed Metropolis approach, regions exchange compact summaries $\mathbf{s}_k = [\mathrm{vecu}(\mathbf{A}_k); \mathbf{b}_k]$ for consensus. It eliminates the need for a central server and enhances both scalability and robustness.

C. Privacy and data protection

Privacy and data protection in this framework are provided at an architectural level through data locality and decentralised coordination. Each region retains its raw measurements and noise realisations locally and exchanges only compact normal-equation summaries for consensus. The framework assumes an honest-but-curious, non-colluding threat model and does not provide cryptographic or differential privacy guarantees. While raw measurement data are protected, certain structural information (e.g., approximate topology or measurement sensitivity) may be inferable from exchanged summaries.

IV. Robust Gauss-Newtonate and Numerical Stability Mechanism

After consensus, each region reconstructs global matrices $\mathbf{A}$ and $\mathbf{b}$ from $(\bar{\mathbf{A}}, \bar{\mathbf{b}})$ and performs a robust Gauss-Newton (GN) update. Noisy or incomplete data can make $\mathbf{A}$ ill-conditioned or non-SPD in weakly coupled grids. Hence, a regularised SPD-projected solution is adopted for numerical stability.

*1) **Adaptive Damped Gauss–Newton Step:** :* At iteration t, the update follows $(\mathbf{A}+\lambda\mathbf{I}), \Delta\mathbf{x} = \mathbf{b}$, where $\lambda \geq 0$ adaptively balances the GN and gradient-descent behaviours. Following the Levenberg–Marquardt principle, λ is increased whenever

Algorithm 1: Federated Decentralized SE Framework

Input: Partitioned grid with K regions; initial state $\mathbf{x}^{(0)}$; consensus type $\mathcal{C} \in \{\text{Metropolis}, \text{Gossip}\}$; tolerances $\varepsilon_x, \varepsilon_{\text{cons}}$; max iteration $T_{\max}$.
Output: Decentralised steady-state estimate $\mathbf{x}^*$.
Initialise: $\mathbf{x}_k^{(0)} = \mathbf{x}^{(0)}, \forall k$.
for $t = 0{:}T_{\max}$ *(GN iter.)* **do**
 Compute local residuals and Jacobians:
 $\mathbf{r}_k = \mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k^{(t)})$, $\mathbf{H}_k = \partial \mathbf{h}_k / \partial \mathbf{x}|_{\mathbf{x}_k^{(t)}}$.
 Form local normal equations $\mathbf{A}_k = \mathbf{H}_k^\top \mathbf{W}_k \mathbf{H}_k$, $\mathbf{b}_k = \mathbf{H}_k^\top \mathbf{W}_k \mathbf{r}_k$, $\mathbf{s}_k = [\text{vecu}(\mathbf{A}_k); \mathbf{b}_k]$.
 if $\mathcal{C} =$ *Metropolis* **then**
  Perform synchronous averaging $\mathbf{s}_k^{(\tau+1)} = w_{kk}\mathbf{s}_k^{(\tau)} + \sum_{j \in \mathcal{N}_k \cap \mathcal{A}_k^{(\tau)}} w_{kj}\mathbf{s}_j^{(\tau)}$ until consensus gap $\leq \varepsilon_{\text{cons}}$.
 else if $\mathcal{C} =$ *Gossip* **then**
  Repeat N_g gossip exchanges per GN iter. $(i,j) \sim \mathcal{E}$ with active link prob. $(1 - p_{\text{drop}})$.
 Reconstruct $\mathbf{A} \approx K\bar{\mathbf{A}}$, $\mathbf{b} \approx K\bar{\mathbf{b}}$.
 Project $\mathbf{A}$ to SPD and solve $(\mathbf{A}_{\text{SPD}} + \lambda\mathbf{I})\Delta\mathbf{x} = \mathbf{b}$.
 Update $\mathbf{x}_k^{(t+1)} = \mathbf{x}_k^{(t)} + \Delta\mathbf{x}$.
 if $\|\Delta\mathbf{x}\|_2 < \varepsilon_x$ **then**
  break
 end
end
return $\mathbf{x}^* = \mathbf{x}_k^{(t+1)}$.

Cholesky factorisation fails, ensuring well-conditioned updates across heterogeneous regions [1].

*2) **Nearest SPD Projection and Stable Update**:* Numerical errors or missing data may cause the normal matrix $\mathbf{A}$ to lose symmetry or positive definiteness, which can make the GN update unstable. To restore these properties, $\mathbf{A}$ is first symmetrised as $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top)$ and then projected onto the nearest symmetric positive-definite (SPD) matrix. This projection is obtained from the eigenvalue decomposition $\mathbf{B} = \mathbf{V}\mathbf{S}\mathbf{V}^\top$, where any small negative eigenvalues in $\mathbf{S}$ are replaced by a minimal positive value ϵ to form $\mathbf{A}_{\text{SPD}} = \mathbf{V}\mathbf{S}_+\mathbf{V}^\top$. The corrected matrix is guaranteed to be well-conditioned and suitable for Cholesky factorisation. The final damped system $(\mathbf{A}_{\text{SPD}} + \lambda\mathbf{I})\Delta\mathbf{x} = \mathbf{b}$ is then solved iteratively, increasing λ only if instability occurs. The resulting update $\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)} + \Delta\mathbf{x}$ keeps voltage magnitudes within $0.95 \leq V_{m,i} \leq 1.05$ p.u. This procedure ensures numerical stability of decentralised GN iterations. A step-by-step procedure of the proposed Fed-DSE is summarised in Alg. 1.

The SPD projection and adaptive damping are invoked only when numerical instability is detected, such as loss of positive definiteness or Cholesky factorisation failure, and serve as stabilisation safeguards rather than default modifications of the WLS formulation. Here, the joint use of adaptive damping and SPD projection ensures convergence even when $\mathbf{A}$ is distorted by communication losses or asynchronous updates.

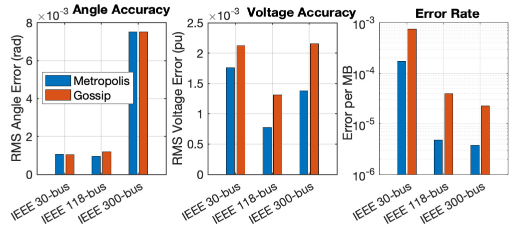

Fig. 2: Accuracy and communication efficiency comparison between Metropolis and Gossip consensus algorithms for IEEE 30-, 118-, and 300-bus systems under $p_{\text{drop}} = 0.3$.

V. Results and Discussion

A. Experimental Setup

The proposed Fed-DSE framework is evaluated on three IEEE test systems: IEEE 30, 118, and 300-bus systems. Each system is partitioned into four regions to represent geographically distributed control areas communicating over unreliable communication links. The Fed-DSE-based estimator is implemented using MATLAB and MATPOWER, with identical parameters for all test cases: GN iterations set to 50, link-drop probability $p_{\text{drop}} = 0.3$, and consensus tolerances $\varepsilon_{\text{cons}} = 10^{-6}$. Each region executes local WLS updates and participates in decentralised consensus via either the Metropolis or the Gossip algorithm to achieve consensus.

The IEEE 30-bus system contains 41 branches, while the 118- and 300-bus systems contain 186 and 411 branches, respectively. MATPOWER is used to generate full AC measurement sets of bus injections, line flows, and voltage magnitudes. Gaussian noise with standard deviations of 1% for active/reactive powers and 0.5% for voltage magnitudes was added to represent realistic measurement uncertainty. The algorithms are benchmarked in terms of residual norm $\|r\|_2$, state update norm $\|\Delta x\|_2$, root-mean-square (RMS) estimation errors, and total data volume transmission loss, discussed in the sections below.

B. Accuracy and Communication Efficiency

The overall accuracy and communication efficiency of the proposed framework are illustrated in Fig. 2. For all test systems, both consensus algorithms achieved high estimation accuracy in the 10^{-3} range, which is comparable to a centralised AC WLS benchmark. The RMS voltage error remained below 5×10^{-3} p.u., while the RMS angle error did not exceed 0.01 rad even in the largest IEEE 300-bus system. These results confirm that decentralisation introduced no significant degradation in state estimation performance compared to the central SE model. While the Gossip-consensus exhibits slightly lower communication costs, the error rate per MB communication and time requirement are higher for this scheme compared to the Metropolis method.

C. Convergence Analysis

Figures 3 and 4 show the convergence behaviour of the proposed framework in terms of residual norm and state-update norm across Gauss–Newton (GN) iterations. For all IEEE test

979-8-3315-8138-1

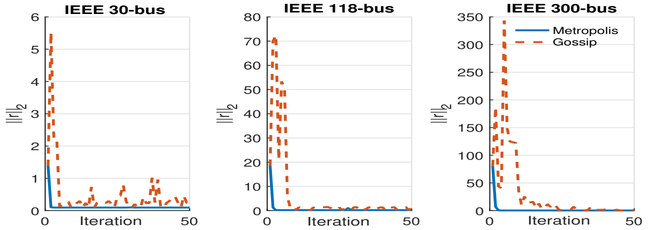

Fig. 3: Residual norm convergence trajectories ($\|r\|_2$ versus iteration) for the IEEE 30, 118, and 300-bus systems under Metropolis and Gossip consensus.

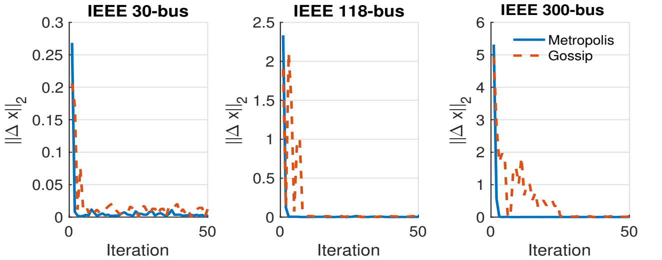

Fig. 4: State update norm convergence trajectories ($\|\Delta x\|_2$ vs iter) for 30, 118, and 300-bus systems for Metropolis and Gossip consensus.

systems, both Metropolis- and Gossip-based variants exhibit stable convergence towards the centralised AC WLS solution, with the residual norm $\|\mathbf{r}\|_2$ decreasing to near-zero values within approximately 5–10 iterations. This behaviour aligns with standard local convergence properties of Gauss–Newton methods and known consensus dynamics.

Convergence is maintained under stochastic link failures, consistent with the experimental setup in Section V-A, which is empirically validated across the IEEE 30-, 118-, and 300-bus systems. The adaptive damping and symmetric positive-definite (SPD) projection further enhance numerical stability by preserving conditioning of the normal equations, as reflected by the bounded and decaying state-update norms.

The proposed method is empirically validated and locally convergent under standard assumptions commonly adopted in distributed Gauss–Newton and consensus-based optimization.

D. Performance comparison

Table I shows the performance of centralised and decentralised federated state estimation (Fed-DSE) for the IEEE 118-bus network. The centralised AC WLS and centralised federated SE configurations produce identical results, which confirms that the federated aggregation reproduces the optimal centralised estimate. Proposed decentralised consensus-based models maintain this high accuracy while removing the need for a single-point aggregator. Among them, Metropolis variant shows the lowest RMSE in both voltage angle and magnitude.

Both schemes have similar per-iteration computational cost dominated by Gauss–Newton updates, but differ in communication overhead. Metropolis converges in fewer synchronous rounds, yielding faster wall-clock convergence in well-connected networks. Gossip uses lighter pairwise exchanges but typically requires more iterations, offering greater robustness under asynchronous and lossy communication.

TABLE I: Performance Comparison for IEEE 118-Bus System

Method	RMSEVa rad	RMSE-Vm pu	Time (s)	Err/MB
Central WLS	9.58e–04	1.05e–03	0.19	3.45e–01
Central Fed [4]	9.58e–04	1.05e–03	0.06	1.13e–03
Decen. Fed (Gossip) [6]	1.18e–03	1.31e–03	0.77	3.98e–05
Decen. Fed (Metropolis)	**9.62e–04**	**7.72e–04**	**0.65**	**4.74e–06**

The results based on the Metropolis and gossip-based Fed-DSE frameworks ensure a benchmark centralised AC WLS-level precision while improving scalability, robustness, and resilience to communication uncertainty in IEEE test systems.

VI. Conclusion

This work proposed a peer-to-peer consensus-driven federated state estimation (Fed-DSE) framework that enables fully decentralised operation without a central coordinator. By leveraging local weighted least-squares updates and consensus exchange among neighbouring areas, the proposed approach achieves very similar estimation accuracy as the centralised AC WLS and server-based federated methods. Future research will extend this paradigm towards a dynamic SE mechanism and extend its robustness against cyberattack scenarios.

References

[1] A. Abur and A. Gómez Expósito, *Power System State Estimation: Theory and Implementation*. New York, NY, USA: Taylor & Francis, 2004.

[2] A. J. Conejo, S. de la Torre, and M. Canas, "An Optimization Approach to Multiarea State Estimation," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 213–221, Feb. 2007,

[3] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, Fort Lauderdale, FL, USA, 2017, pp. 1273–1282.

[4] M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "A Secure Federated Learning Framework for Residential Short-Term Load Forecasting," *IEEE Transactions on Smart Grid*, vol. 15, no. 2, pp. 2044–2055, Mar. 2024, doi: 10.1109/TSG.2023.3292382.

[5] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Symp. Information Processing in Sensor Networks (IPSN)*, Los Angeles, CA, USA, Apr. 2005, pp. 9–16.

[6] M. A. Husnoo, A. Anwar, M. E. Haque, and A. N. Mahmood, "Decentralized Federated Anomaly Detection in Smart Grids: A P2P Gossip Approach," *arXiv preprint* arXiv:2407.15879 [cs.CR], Jul. 2024.

[7] G. Cheng, Y. Lin, A. Abur, A. Gómez-Expósito, and W. Wu, "A survey of power system state estimation using multiple data sources: PMUs, SCADA, AMI, and beyond," *IEEE Transactions on Smart Grid*, vol. 15, no. 1, pp. 1129–1151, Jan. 2024.

[8] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, May 2013, doi: 10.1109/TPWRS.2012.2218275.

[9] J. Ban, J. Im, Y.-J. Kim, and J. Zhao, "Decentralization of phasor-aided state estimation using local state vector extension," *IEEE Transactions on Power Systems*, vol. 36, no. 5, pp. 4645–4659, Sept. 2021.

[10] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated Learning for Internet of Things: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1622–1658, 2021, doi: 10.1109/COMST.2021.3075439.

[11] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007, doi: 10.1109/JPROC.2006.887293.

979-8-3315-8138-1