Unified Bayesian State Estimation for Distribution Networks with DERs under Low Observability and Heterogeneous Sampling

Sipeng Luo, Tianyang Zhao, Gengfeng Li, Zhaohong Bie
School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, 710049, China
SipengLuo@stu.xjtu.edu.cn

***Abstract*—In medium- and low-voltage distribution systems, dynamic state estimation is substantially impeded by the inability to capture rapid fluctuations of distributed energy resources due to heterogeneous sensor sampling rates, a challenge further exacerbated by incomplete spatial coverage. To address these issues, a novel distribution system state estimation framework is proposed. A multi-task Gaussian process is employed to capture these DER fluctuation characteristics and to impute missing measurements arising from sampling inconsistencies. Furthermore, unobserved data resulting from limited sensor coverage are recovered through a sparse Bayesian matrix completion approach. The robustness of the proposed method is demonstrated through simulations on the IEEE 33-bus and 123-bus systems under conditions of high data missing, low observability, and measurement noise. High accuracy in both data interpolation and voltage state recovery is achieved.**

***Index Terms*—Medium- and Low-voltage distribution network, State estimation, Sparse Bayesian matrix completion, Multi-task Gaussian process.**

I. Introduction

DISTRIBUTION system state estimation (DSSE) is defined as the inference of system state variables from limited measurements, extended by dynamic state estimation (DSE) to track time-varying system evolution and capture transient dynamics in modern power grids [1]. Such tracking is essential for real-time monitoring, optimal scheduling, protection coordination, and resilience planning in distribution networks. Because field measurements are typically insufficient to ensure observability, pseudo-measurements are leveraged by DSSE to augment data adequacy and provide actionable situational awareness [2]. However, reliable estimation remains difficult due to two compounding factors: (1) low metering density relative to system scale, and (2) inadequate capture of the complex and rapid fluctuations of distributed energy resources (DERs) due to heterogeneous sampling rates of smart metering infrastructure.

Extensive work has been conducted for DSSE. Weighted least squares (WLS) static state estimation is one widely adopted DSSE method, by which bad data can be detected and identified [3]. Another approach has been developed in which conventional supervisory control and data acquisition (SCADA) data are combined with phasor measurements using robust linear weighted least squares (RLWLS) for state estimation [4]. Furthermore, dynamic state estimation methods, such as extended Kalman filter (EKF) and unscented Kalman filter (UKF), have been applied to integrate temporal information and improve estimation accuracy in distribution networks with high measurement variability [2]. However, high system observability is required by these methods, which is challenging to achieve in medium- and low-voltage distribution networks.

Several works have addressed DSSE under low observability conditions. A three-phase least squares approach is presented in [5] to accommodate diverse measurements, while the normal equation method is applied in [6] to determine real-time states with explicit observability analysis. Furthermore, an unscented particle filter (UPF) is proposed in [7] to align particle distributions with the true posterior, and Quantum Brain Storm Optimization (QBSO) is employed in [8] to resolve the nonlinear DSSE formulation. However, homogeneity of input data is assumed by these methods, whereas in practice, heterogeneous sampling rates are exhibited by different smart meters. Consequently, the dynamic characteristics of DERs cannot be adequately captured, and the accuracy of DSSE is thereby undermined.

Little research has been conducted on the heterogeneous sampling rate characteristics of smart meters. Machine learning techniques, including MLP-based imputation frameworks with time encoding [9], have been applied to reconstruct missing smart meter data with high accuracy. Linear imputation, k-nearest neighbors, and generative adversarial imputation networks (GAIN) are compared in [10] for missing data, evaluating their impact on forecasting accuracy. However, these imputation methods are predominantly deterministic and fail to quantify reconstruction uncertainty. Moreover, spatial correlations in distribution systems with DERs are ignored by simple temporal interpolation. Therefore, a method that exploits spatio-temporal correlations and provides confidence intervals for imputed measurements is required.

Distinct advantages in uncertainty quantification and sparse data recovery are demonstrated by Bayesian methods. A Bayesian DSSE algorithm is proposed in [11] to handle non-Gaussian measurement uncertainties and correlations between pseudomeasurements, providing accurate uncertainty intervals for estimated states. Another Bayesian framework based on physics-informed sparse matrix completion is proposed in [12], to accurately estimate the states in low-observability conditions. Furthermore, A forecasting aided-DSSE (FA-DSSE) method using VB-ACIF delivers improved estimation accuracy and is suitable for practical distribution system operation [13].

In summary, the combined effects of low observability and heterogeneous sampling rates in medium and low voltage distribution systems with DERs penetration are inadequately addressed, particularly regarding uncertainty quantification of imputed measurements and exploitation of spatio-temporal correlations. A unified Bayesian DSSE framework is proposed to address these issues. Multi-task Gaussian Process (MTGP) is employed to compensate for the low measurement redundancy resulting from heterogeneous sampling rates of smart meters, where an intrinsic coregionalization model (ICM) kernel function is utilized to capture DER characteristics. For the state estimation process, sparse Bayesian matrix completion (SBMC) is applied with linearized three-phase power flow equations imposed as physical constraints. The main contributions of this paper are summarized as follows:

1) Unlike existing deterministic approaches, estimation uncertainty is quantified by the proposed method, thereby improving the credibility of DSSE results for risk-aware

Supported by Smart Grid-National Science and Technology Major Project (2030) Key Technologies and Demonstration Project for the Security of Distribution Networks Integrated with Multi-Type Emerging Entities (2025ZD0805900). Corresponding author: Tianyang Zhao.

979-8-3315-8138-1

operational decisions.

2) A SBMC-based DSSE method is proposed, by which accurate state estimates are obtained under low observability through the incorporation of linearized physical constraints, whereas WLS methods fail.
3) By utilizing an ICM kernel of MTGP, the correlations between measurement tasks are captured, enabling accurate reconstruction of rapidly fluctuating renewable energy dynamics from heterogeneous sampling rates.

II. Assumptions

A distribution system is defined as the portion of an electric power system through which electric energy is delivered from transmission systems or generating sources to end-use customers. Distribution substations, primary and secondary distribution lines, distribution transformers, and service connections are included. The schematic diagram of such a system, operating under heterogeneous measurements and structural uncertainty, is depicted in Fig. 1. In this paper, the bus voltage of medium and low voltage distribution systems is estimated through DSSE.

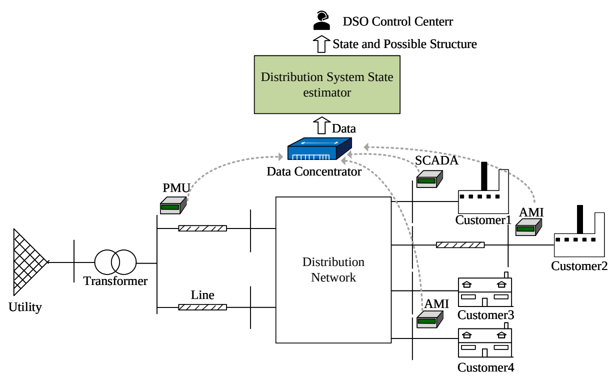

Fig. 1. Illustration of The Scenario Characterized by Heterogeneous Sampling Rates and Low Observability.

Several assumptions are made to solve the problem: (1) Three principal classes of data acquisition sensors are employed [12]: Phasor measurement units (PMUs), by which voltage magnitude and the real and imaginary components of bus voltage are sampled at 0.1-second intervals; SCADA devices, by which bus voltage magnitudes are recorded at 1-minute intervals; and advanced metering infrastructure (AMI), by which active and reactive power measurements are collected at 15-minute intervals. (2) Historical measurement data is assumed to be available to capture the spatio-temporal correlations of the system states. (3) The spatio-temporal data matrix of the distribution system is assumed to be low-rank.

III. Methodology

A two-stage pipeline is outlined in this section: (1) heterogeneous-rate measurements are aligned by MTGP with an ICM kernel, and time-synchronized probabilistic estimates are output; (2) unobserved states are recovered by SBMC under linearized three-phase power-flow constraints. The flowchart of the proposed method is presented in Fig. 2.

A. Multi-task Gaussian process

Given the physical coupling among adjacent nodes, voltage and power at one node are typically correlated with those of electrically connected nodes. MTGP is employed to exploit

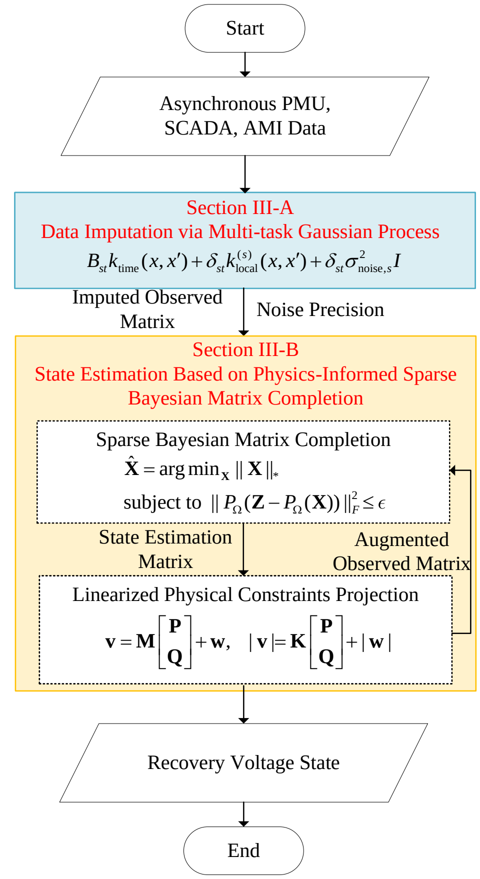

Fig. 2. Flowchart of DSSE Method for medium and low voltage Distribution System

the correlations between measurements. Specifically, each measurement is treated as a task, and the MTGP method is utilized to capture both intra-task and inter-task correlations. The GP prior for the MTGP is defined as follows:

$$f(\cdot)=\mathcal{GP}(m_\phi(\cdot),k_\theta(\cdot,\cdot)), \tag{1}$$

where $m_\phi(\cdot)$ is the mean function and $k_\theta(\cdot,\cdot)$is the covariance function parameterized by ϕ and θ, respectively. The mean function can be implemented by a multilayer perceptron with parameters ϕ globally shared across all time series. Temporal correlations within each time-series are captured by the kernel $k_\theta(\cdot,\cdot)$, with θ denoting kernel hyperparameters. High imputation performance using only limited measurements is enabled by effective encoding of prior knowledge into the GP model.

To ensure that the superposition of physically coupled global trends and rapid, node-specific fluctuations induced by DERs is accurately modeled, a multi-task covariance structure based on the ICM is proposed. Within this framework, the total covariance is decomposed into a shared global component, by which spatial correlations across tasks are captured via a task correlation matrix, and an independent local component, through which node-specific characteristics are preserved. The Radial Basis Function (RBF) is employed as the fundamental temporal kernel for both components, as the smooth continuity of physical power states and the non-linear temporal patterns of renewable generation are effectively represented by its properties. Consequently, the multi-task kernel is defined as the sum of these shared and local contributions:

$$\begin{aligned}\mathbf{K}_{st}&=\mathbf{K}((s,x),(t,x'))=B_{st}k_{\text{time}}(x,x')\\&+\delta_{st}k^{(s)}_{\text{local}}(x,x')+\delta_{st}\sigma^2_{\text{noise},s}I,\end{aligned} \tag{2}$$

979-8-3315-8138-1

$$k_{\text{time}}(x,x') = \sigma^2_{\text{time}} \exp(-\frac{(x-x')^2}{2\ell^2_{\text{time}}}), \tag{3}$$

$$k^{(s)}_{\text{local}}(x,x') = \sigma^2_{\text{local},s} \exp(-\frac{(x-x')^2}{2\ell^2_{\text{local},s}}), \tag{4}$$

where $B_{st} = LL^\top \in \mathbb{R}^{S\times S}$ is a positive semidefinite task-covariance matrix (learned via Cholesky factor L) capturing inter-task correlations, $k_{\text{time}}(x,x')$ is a shared temporal RBF kernel, $k^{(s)}_{\text{local}}(x,x')$ is a task-specific RBF kernel, $\sigma^2_{\text{noise},s}$ is task-specific observation noise variance, and δ_{st} is the Kronecker delta. Intuitively, $B_{st}k_{\text{time}}$ models synchronous, common-source temporal co-variations across tasks, whereas $k^{(s)}_{\text{local}}$ captures task-specific dynamics, and $\sigma^2_{\text{noise},s}$ accounts for heterogeneous sensor noise. Hyperparameters ℓ_{time} and $\ell_{\text{local},s}$ are shared and task-specific signal variance.

The MTGP prior function f can be written as:

$$\begin{bmatrix} f_1 \\ \vdots \\ f_S \end{bmatrix} \sim \mathcal{GP}\left(\begin{bmatrix} m_\phi(\tilde{\boldsymbol{x}}_1) \\ \vdots \\ m_\phi(\tilde{\boldsymbol{x}}_S) \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{11} & \cdots & \mathbf{K}_{1S} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_{S1} & \cdots & \mathbf{K}_{SS} \end{bmatrix} \right). \tag{5}$$

The GP prior learnable parameters $\psi = \{\phi, \theta, L\}$ are optimized by maximizing the log marginal likelihood (LML) given in [12]. Once the GP prior parameters are optimized and encoded, the posterior distribution is calculated using Gaussian identities based on the observed data, from which the target tasks are predicted.

The predictive distribution of $\tilde{y}^*_s$ is

$$p(\tilde{y}^*_s|\tilde{x}^*_s, \tilde{y}_s, \tilde{x}_s) = \mathcal{N}(\mu^*, C^*), \tag{6}$$

with

$$\mu^* = m_\phi(\tilde{x}^*_s) + \mathbf{K}_{*x}(\mathbf{K}_{xx})^{-1}(\tilde{y}_s - m_\phi(\tilde{x}_s)), \tag{7}$$

$$\mathbf{C}^* = \mathbf{K}_{**} - \mathbf{K}_{*x}(\mathbf{K}_{xx})^{-1}\mathbf{K}_{x*}. \tag{8}$$

where $\mathbf{K}_{xx}$ is the full block covariance on training inputs assembled by the ICM kernel, $\mathbf{K}_{*x}$ and $\mathbf{K}_{**}$ are the corresponding cross- and self-covariance blocks between targets and training inputs.

Remark III.1. With bounded ICM kernels and Gaussian noise, the MTGP posterior mean is equivalent to vector-valued kernel ridge regression, ensuring a well-posed and convex estimator for aligning heterogeneous-rate measurements into time-synchronized probabilistic predictions [14].

B. Physically-Constrained Sparse Bayesian Matrix Completion

The imputed time-aligned measurements produced by the MTGP are subsequently utilized as inputs for state estimation. To recover the complete matrix $\mathbf{X}$ from incomplete observations, the SBMC method is applied. Unlike deterministic Matrix Completion (MC), by which only point estimates are provided, or standard Bayesian Matrix Completion (BMC), in which dense latent factors are typically assumed, SBMC is distinguished by the incorporation of sparsity-inducing priors. Through this formulation, Automatic Rank Determination (ARD) is enabled, whereby model complexity is inferred without manual tuning. Furthermore, uncertainties quantified by the MTGP are explicitly incorporated as priors, ensuring that error propagation is mitigated. Consequently, the underlying low-rank structure is effectively exploited by leveraging: i) spatial correlations between measurement locations, and ii) physical constraints from linearized power flow equations [15].

Based on the imputation result of MTGP, the noisy matrix can be defined as:

$$\mathbf{Z} = [P_b, Q_b, \Re(\mathbf{v}_b), \Im(\mathbf{v}_b), |\mathbf{v}_b|]. \tag{9}$$

Let m denote the set of all non-slack buses. Each row of the observation matrix $\mathbf{Z} \in \mathbb{R}^{mn}$, $n = 5$, represents measurements for bus $b \in m$: active power P_b, reactive power Q_b, and real and imaginary parts of voltage phasors $\Re(\mathbf{v}_b)$ and $\Im(\mathbf{v}_b)$. Let $\Omega_{train} \subseteq \{1,\dots,m\} \times \{1,\dots,n\}$ denote the known entries in $\mathbf{Z}$. The observation matrix $P_\Omega(\mathbf{Z})$ is:

$$[P_\Omega(\mathbf{Z})]_{mn} = \begin{cases} \mathbf{Z}_{mn}, & \textit{if} (m,n) \in \Omega_{train} \\ 0, & \textit{otherwise} \end{cases}. \tag{10}$$

The complete low-rank matrix is recoverd by the deterministic matrix completion formulation, as

$$\begin{aligned} &\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} ||\mathbf{X}||_* \\ &\text{subject to } ||P_\Omega(\mathbf{Z} - P_\Omega(\mathbf{X}))||^2_F \le \epsilon, \end{aligned} \tag{11}$$

$$\mathbf{v} = \mathbf{M}\begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix} + \mathbf{w}, \tag{12}$$

$$|\mathbf{v}| = \mathbf{K}\begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix} + |\mathbf{w}|, \tag{13}$$

where, the nuclear norm $||\mathbf{X}||_*$ is defined as the sum of singular values of matrix $\mathbf{X}$. (12), (13) The linearized power flow constraints relating voltage phasors $\mathbf{v}$ and voltage magnitudes $|\mathbf{v}|$ to the power measurements are captured by (12) and (13) [12]. Here:

$$\mathbf{M} = \left(\mathbf{Y}^{-1}_{\mathbf{LL}}\text{diag}(\overline{\mathbf{w}})^{-1}, -j\mathbf{Y}^{-1}_{\mathbf{LL}}\text{diag}(\overline{\mathbf{w}})^{-1}\right), \tag{14}$$

$$\mathbf{K} = \left(|\text{diag}(\overline{\mathbf{w}})|^{-1}\Re(\text{diag}(\overline{\mathbf{w}})\mathbf{M})\right), \tag{15}$$

where, $\mathbf{w} = -\mathbf{Y}^{-1}_{\mathbf{LL}}\mathbf{Y}_{\mathbf{L0}}\mathbf{v_0}$ is the zero-load voltage. $\mathbf{v_0} \in \mathbb{C}^3$ denotes the complex vectors collecting the three-phase nodal voltage at the slack bus. Here, $\mathbf{Y_{LL}} \in \mathbb{C}^{m\times m}$ and $\mathbf{Y_{L0}} \in \mathbb{C}^{m\times 3}$ are the sub-matrices of the three-phase admittance matrix.

Under incoherence and sufficient samples, stable recovery is achieved by nuclear-norm-regularized matrix completion [16]. When physical constraints are additionally enforced, the feasible set is further restricted, thereby tightening the recovery and improving robustness to noise and missingness.

The formulation in (11) can be recast as a semidefinite program. However, it is rendered computationally inefficient for large matrices, and prior knowledge of the matrix rank is generally required. Furthermore, extensive parameter tuning and data-dependent supervision are necessitated. To overcome these shortcomings and effectively incorporate the MTGP output, an SBMC framework is proposed. A significant advantage regarding scalability is offered by this approach. In contrast to computationally intensive alternatives, the complexity of the proposed Variational Bayesian Matrix Factorization is restricted to $\mathcal{O}(N \times R^2)$ per iteration, where N denotes the number of buses and R represents the low rank (typically $R \ll N$). Consequently, a linear relationship between computational cost and the system size is established, ensuring suitability for practical implementation. Within this framework, equations (12) and (13) are employed as physical constraints, with the detailed derivation process presented in [17].

It should be noted that the output of MTGP provides the uncertainty in imputations, which can be used as an estimate of the noise variances for the imputed measurements.

IV. Case Study

The proposed Bayesian framework is simulated on the IEEE 33-bus test system [18] and IEEE 123-bus test system [19] with the Julia language in AMD® Epyc 9754 128-core processor. Load and irradiance profiles are derived from field data. The voltage profiles are obtained via the three power flow

979-8-3315-8138-1

TABLE I
ROOT MEAN SQUARE ERROR OF IMPUTED TIME-SERIES DATA (MEAN VALUES)

System	Gaussian Noise Measurement	Measurement	40% Measurement Redundancy		60% Measurement Redundancy		80% Measurement Redundancy		90% Measurement Redundancy	
			Linear Interpolation	GP	Linear Interpolation	GP	Linear Interpolation	GP	Linear Interpolation	GP
IEEE 33 System	1%	Active Power	0.0070%	0.0011%	0.0027%	0.0003%	0.0017%	0.0003%	0.0019%	0.0003%
		Reactive Power	0.0150%	0.0020%	0.0068%	0.0009%	0.0066%	0.0009%	0.0050%	0.0009%
	5%	Active Power	0.0038%	0.0008%	0.0037%	0.0005%	0.0017%	0.0005%	0.0012%	0.0004%
		Reactive Power	0.0223%	0.0042%	0.0110%	0.0016%	0.0050%	0.0013%	0.0052%	0.0011%
IEEE 123 System	1%	Active Power	0.0039%	0.0004%	0.0011%	0.0003%	0.0008%	0.0002%	0.0011%	0.0003%
		Reactive Power	0.0019%	0.0003%	0.0010%	0.0002%	0.0007%	0.0002%	0.0004%	0.0002%
	5%	Active Power	0.0035%	0.0017%	0.0020%	0.0013%	0.0020%	0.0012%	0.0018%	0.0013%
		Reactive Power	0.0027%	0.0015%	0.0020%	0.0011%	0.0014%	0.0009%	0.0018%	0.0010%

analyses in OpenDSS [20]. Photovoltaics systems are installed at bus 18 in the IEEE 33-bus test system and at bus 66 in the IEEE 123-bus test system.

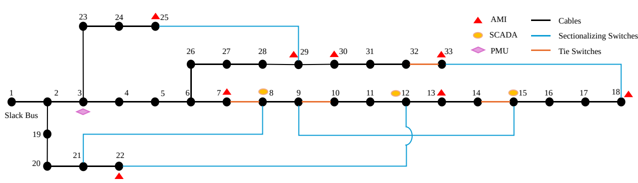

Fig. 3. Measurement configuration at 10% FAD for IEEE 33 test system

The locations of sensors in IEEE 33 system are presented in Fig. 3. Data from 15 aggregated AMI sensors, 9 SCADA sensors, and 1 PMU are considered, providing only 10% system observability (FAD = 10%). The sensor locations are specified as follows:

1) 15 time-series data from the aggregated AMI sensors are from the following load buses:
   a) Phase A - Bus 13, 18, 25, 29, 33.
   b) Phase B - Bus 7, 18, 22, 29, 30.
   c) Phase C - Bus 7, 13, 25, 30, 33.
2) The 9 time-series data of SCADA sensors are obtained at the three-phase buses 8, 12, and 15.
3) A PMU sensor is placed at bus 3, which measures the voltages at three phases at 0.1 s interval.

Unobserved entries resulting from limited measurement redundancy are not incorporated into the training process, ensuring that the actual sampling conditions encountered in practice are accurately reflected.

Training is performed via Adam optimization (learning rate: 0.01, max epochs: 200) by maximizing the LML, through which mean network weights and kernel hyperparameters are jointly updated with intrinsic Bayesian regularization. Early stopping (patience: 30 epochs) is employed to ensure convergence. A dataset spanning 200 time snapshots per sensor is utilized, where the partitioning into training and test sets is determined by the measurement redundancy level. Finally, the MTGP-based imputation is benchmarked against linear interpolation.

The superior performance of the MTGP-based approach under high missing data rates and noisy conditions is demonstrated in Table I. Noise levels of 1% and 5% are selected to reflect realistic uncertainties from high-precision smart meters (Class 0.2S [21]) and lower-grade sensors, respectively. Significant RMSE reductions are consistently exhibited across both test systems; notably, under a measurement redundancy level of 40%, the error is decreased from 0.0070% to 0.0011% in the IEEE 33 system, and from 0.0039% to 0.0004% in the IEEE 123 system. This precision is attributed to the MTGP's capability to model complex temporal correlations via its covariance kernel. Unlike linear interpolation, the rapid fluctuations characteristic of renewable energy generation are effectively captured by the MTGP even under severe data sparsity. Consequently, high-fidelity data reconstruction is ensured for downstream state estimation tasks in networks of varying scales. As illustrated in Fig. 4, a non-linear increase in training time is observed, rising from 2.75s for 5 states to approximately 1980s for 150 states. This trend is attributed to the cubic complexity ($\mathcal{O}(N^3)$) required for covariance matrix inversion during hyperparameter optimization. Consequently, an "offline training, online prediction" strategy is implemented, wherein the computationally intensive learning is restricted to the offline phase to ensure rapid online state estimation.

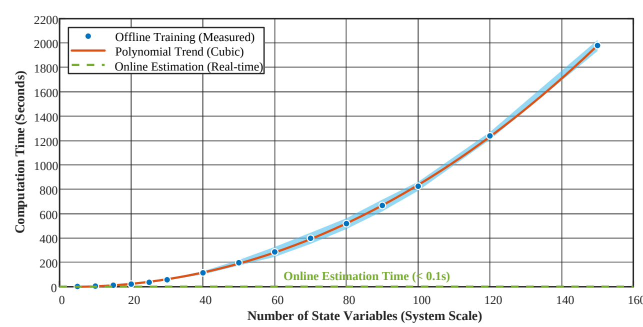

Fig. 4. Relationship of the State number and Computation Time.

The predicted mean and variance from consistent time series are used in SBMC for state estimation. At any instant, an incomplete observation matrix with structure illustrated in (9) is available, with known elements randomly sampled to reflect different FADs. The ratio of available measurements to system states is quantified by global measurement redundancy, as defined in [22]. Initial factorized matrices are obtained via Singular Value Decomposition (SVD) of the observation matrix. Approximate posterior distributions are updated using mean field variational Bayes. For linearly interpolated time series, the precisions are also updated. Iterations proceed until the convergence criterion $\frac{||\hat{\mathbf{X}}^{iter}-\hat{\mathbf{X}}^{iter-1}||_F}{||\hat{\mathbf{X}}^{iter-1}||_F} \leq 10^{-8}$ is satisfied.

In order to quantify the performance of the SBMC-based approaches, MAPE (Mean Absolute Percentage Error) and MIAE (Mean Integrated Absolute Error) metrics are used for voltage magnitude and voltage angle, respectively. The IEEE 33 test system and 123 test system are considered, with the FAD 10% and 20%. Measurement locations and quantities are

979-8-3315-8138-1

TABLE II
PERFORMANCE COMPARISON OF STATE ESTIMATION UNDER DIFFERENT FAD SCENARIOS

FAD Scenario	Test System	Noise Level	MTGP+SBMC (Proposed)						LN+SBMC (Baseline)	
			Voltage Magnitude			Voltage Angle			Voltage Mag.	Voltage Angle
			MAPE	PICP	MPIW	MIAE	PICP	MPIW	MAPE	MIAE
FAD=10%	IEEE 33 System	1%	4.5079%	40.63%	0.0114%	1.7840%	51.04%	0.0126%	5.2879%	1.9810%
		5%	3.6304%	45.83%	0.0164%	1.5553%	79.17%	0.0182%	4.4492%	1.7797%
	IEEE 123 System	1%	1.2915%	45.82%	0.0365%	3.3641%	13.45%	0.0258%	1.6681%	2.5932%
		5%	1.1538%	74.18%	0.0424%	3.3870%	50.18%	0.0323%	1.6678%	2.5936%
FAD=20%	IEEE 33 System	1%	4.0951%	46.88%	0.0098%	1.7609%	53.13%	0.0126%	4.9729%	1.8116%
		5%	4.0574%	47.92%	0.0141%	1.8009%	68.75%	0.0184%	4.8421%	1.6531%
	IEEE 123 System	1%	0.9893%	76.00%	0.0631%	3.9147%	14.18%	0.0410%	1.6037%	2.7936%
		5%	0.9693%	76.00%	0.0637%	3.6289%	25.45%	0.0411%	1.6033%	2.7913%

randomly varied for any given FAD. Complete availability of all elements in matrix $\mathbf{Z}$ is provided at an FAD of 100%.

As detailed in Table II, estimation performance is evaluated under varying FAD. Superior voltage recovery is consistently exhibited by the proposed MTGP+SBMC framework compared to the LN+SBMC baseline, with the IEEE 123 system's MAPE reduced to below 1.0% at 20% FAD. This accuracy is attributed to the effective utilization of spatial correlations for state reconstruction. Consequently, robust estimation is maintained even under low observability conditions.

Uncertainty is quantified via Prediction Interval Coverage Probability (PICP) and Mean Prediction Interval Width (MPIW). With increased FAD, PICP is significantly improved to 76.00% in the IEEE 123 system, indicating better encapsulation of true states. This enhancement is driven by the reduction of epistemic uncertainty through greater measurement availability. Consequently, higher reliability is achieved with improved grid visibility.

V. Conclusion

A Bayesian framework is proposed for DSSE in medium and low voltage distribution networks, with heterogeneous sensor sampling rates and incomplete spatial coverage being addressed. High-quality data imputation for heterogeneous measurements is achieved via MTGP, capturing fast fluctuation of DERs, and providing time-synchronized probabilistic estimates. Subsequently, robust recovery of unobserved voltage and power states is realized through a physics-constrained SBMC approach. The efficacy of the proposed methodology is demonstrated on the IEEE 33-bus system and the IEEE 123-bus system, where high-fidelity voltage recovery is substantiated. Future work is directed toward mitigating the computational burden of the MTGP method through algorithmic optimizations for high-dimensional state spaces. Additionally, self-adaptive tuning strategies will be investigated to address the sensitivity of SBMC hyperparameters, thereby further enhancing the framework's adaptability.

References

[1] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A survey on state estimation techniques and challenges in smart distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312–2322, 2019.

[2] 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, 2017.

[3] T. P. Vishnu, V. Viswan, and A. M. Vipin, "Power system state estimation and bad data analysis using weighted least square method," in *2015 International Conference on Power, Instrumentation, Control and Computing (PICC)*, 2015, pp. 1–5.

[4] S. K. Kotha, B. Rajpathak, M. Mallareddy, and R. Bhuvanagiri, "Wide area measurement systems based power system state estimation using a robust linear-weighted least square method," *Energy Reports*, vol. 9, pp. 23–32, 2023, 2022 The 3rd International Conference on Power and Electrical Engineering. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352484723007898

[5] M. Baran and A. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Transactions on Power Systems*, vol. 9, no. 3, pp. 1601–1609, 1994.

[6] C. Lu, J. Teng, and W.-H. Liu, "Distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 10, no. 1, pp. 229–240, 1995.

[7] B. Chen, H. Li, and X. Su, "Dynamic state estimation of distribution network base on pseudo measurement modeling and upf," in *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2019, pp. 54–59.

[8] A. Ito and H. Mori, "Application of quantum brain storm optimization to robust state estimation in distribution systems," in *2023 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2023, pp. 1–5.

[9] H. Wang, F. Zhang, Y. Cai, and Y. Chen, "A novel imputation approach for power load time series data based on tsdatawig," *Scientific Reports*, vol. 15, no. 1, p. 20863, jul 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-05481-4

[10] D.-S. Lee and S.-Y. Son, "Pv forecasting model development and impact assessment via imputation of missing pv power data," *IEEE Access*, vol. 12, pp. 12 843–12 852, 2024.

[11] P. A. Pegoraro, A. Angioni, M. Pau, A. Monti, C. Muscas, F. Ponci, and S. Sulis, "Bayesian approach for distribution system state estimation with non-gaussian uncertainty models," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 11, pp. 2957–2966, 2017.

[12] S. Dahale and B. Natarajan, "Bayesian framework for multi-timescale state estimation in low-observable distribution systems," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4340–4351, 2022.

[13] S. Chatterjee, S. Ghosh, S. Chakrabarti, and A. Mohapatra, "Data-driven bayesian framework to mitigate smart meter asynchrony in forecasting aided distribution system state estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–12, 2025.

[14] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 11 2005. [Online]. Available: https://doi.org/10.7551/mitpress/3206.001.0001

[15] S. Luo, T. Zhao, and Z. Bie, "Hydistflow.jl: A unified dynamic hybrid ac/dc power flow package for der-rich distribution systems," *SoftwareX*, vol. 33, p. 102486, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352711025004522

[16] S. N. Negahban, P. Ravikumar, M. J. Wainwright, and B. Yu, "A unified framework for high-dimensional analysis of \$m\$-estimators with decomposable regularizers," *ArXiv*, vol. abs/1010.2731, 2009. [Online]. Available: https://api.semanticscholar.org/CorpusID:656534

[17] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2520–2530, 2020.

[18] S. H. Dolatabadi, M. Ghorbanian, P. Siano, and N. D. Hatziargyriou, "An enhanced ieee 33 bus benchmark test system for distribution system studies," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 2565–2572, 2021.

[19] K. P. Schneider, B. A. Mather, B. C. Pal, C. W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. R. Pereira, L. F. Ochoa, L. R. de Araujo, R. C. Dugan, S. Matthias, S. Paudyal, T. E. McDermott, and W. Kersting, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3181–3188, 2018.

[20] D. Montenegro, M. Hernandez, and G. A. Ramos, "Real time opendss framework for distribution systems simulation and analysis," in *2012 Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (TD-LA)*, 2012, pp. 1–5.

[21] International Electrotechnical Commission, *Electricity metering equipment – Particular requirements – Part 22: Static meters for AC active energy (classes 0.1 S, 0.2 S and 0.5 S)*, International Electrotechnical Commission (IEC), Geneva, Switzerland, 2020, iEC 62053-22:2020.

[22] A. S. Bretas, A. Rossoni, R. D. Trevizan, and N. G. Bretas, "Distribution networks nontechnical power loss estimation: A hybrid data-driven physics model-based framework," *Electric Power Systems Research*, vol. 186, p. 106397, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779620302030