

Structure-inspired Parameter Estimation of Composite Load Model with Distributed Generation via Interdependency-based Parameter Grouping

Min-Seung Ko, *Member, IEEE*, Seonghan Kim, *Graduate Member, IEEE*, Jae-Kyeong Kim, *Member, IEEE*, Taesik Nam, *Member, IEEE*, Hao Zhu, *Senior Member, IEEE*, and Kyeon Hur, *Senior Member, IEEE*

Abstract—This paper proposes a structure-inspired parameter estimation method for the composite load model with distributed generation (CMPLDWG) developed by Western Electricity Coordinating Council (WECC). The high dimensionality and strong nonlinearity of this model, due to the aggregated distributed energy resource (DER_A) component, greatly complicate the reliable parameter estimation. We put forth a parameter interdependency analysis to partition the entire parameter space into smaller subsets, thereby decomposing the original high-dimensional estimation problem into multiple tractable subproblems. After applying the interdependency-based parameter grouping, the estimation for each subset is performed using both the Levenberg-Marquardt (LM) algorithm and the enhanced snake optimizer (ESO), demonstrating the solver-agnostic improvements in the convergence stability and estimation accuracy. An initialization strategy is developed to improve the robustness of subsequent optimization. Case studies in the New England 68-bus system confirm that the interdependency-based parameter grouping significantly improves the convergence speed, numerical stability, and estimation accuracy across different disturbance scenarios. The ability of the proposed estimation method is further validated through real-world measurements, and it can be broadly applicable to other modeling problems.

Index Terms—Composite load model, distributed generation, enhanced snake optimizer, Levenberg-Marquardt (LM) algorithm, parameter estimation, interdependency-based parameter grouping.

I. INTRODUCTION

MODERN electric power systems continue to witness an increasing penetration of distributed energy resources (DERs) and a growing heterogeneity of end-use load components. Thus, the accurate dynamic load modeling has become essential for reliable grid operation and planning [1]. Traditional load models fail to capture the fast and nonlinear dynamics of motors and inverter-based DERs. In this context, the composite load model with distributed generation (CMPLDWG), developed by the Western Electricity Coordinating Council (WECC), has emerged as a widely adopted and comprehensive load model [2], [3]. The CMPLDWG comprises static ZIP loads and dynamic components including four induction motors, an electronic load, and an aggregated DER (DER_A) block that represents the aggregated DER behavior. This structure enables the accurate reproduction of disturbance responses, e.g., voltage sags, frequency deviations, and DER tripping, which supports the planning and protection [4], [5].

Despite its high fidelity, the practical deployment of CMPLDWG is significantly hindered by its model complexity. It contains 183 nonlinearly and structurally coupled parameters and 25 nonlinear differential equations, many of which interact over a wide range of time scales. Due to this issue, the data-driven CMPLDWG modeling is a well-known challenge. The iterative Levenberg-Marquardt (LM) algorithm [6], as the dominant solver thanks to its efficiency in smooth and moderately scaled problems [7], [8], has substantial limitations when applied to the CMPLDWG with high dimensionality and strong nonlinearity. In particular, the wide parameter scale disparity induces significant variations in trajectory sensitivities, leading to an ill-conditioned Jacobian matrix and unstable updates. Strong nonlinearities further exacerbate this issue by introducing locally sharp curvature,

Manuscript received: July 13, 2025; revised: November 7, 2025; accepted: February 5, 2026. Date of CrossCheck: February 5, 2026. Date of online publication: April 27, 2026.

This work was supported by the International Cooperation Research Program through National Research Foundation of Korea grant funded by the Korea Government (MSIT) (No. RS-2023-00218377), Basic Science Research Program through National Research Foundation of Korea grant funded by the Korea Government (MSIT) (No. 2021R1A2C209550313), the US National Science Foundation (No. 2150571), and the US Department of Energy grant from the Advanced Manufacturing Office.

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

M.-S. Ko and H. Zhu are with the Chandra Family Department of Electrical and Computer Engineering (ECE), The University of Texas at Austin, Austin, USA (e-mail: kms4634500@utexas.edu; haozhu@utexas.edu).

S. Kim and K. Hur (corresponding author) are with the School of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea (e-mail: seonghan07@yonsei.ac.kr; khur@yonsei.ac.kr).

J.-K. Kim is with the Korea Electrotechnology Research Institute, Uiwang, 16029, Republic of Korea (e-mail: jkkim@keri.re.kr).

T. Nam is with the Korea Institute of Industrial Technology, Gwangju, Republic of Korea (e-mail: tsnam@kitech.re.kr).

DOI: 10.35833/MPCE.2025.000644



reducing effective column independence in the Jacobian matrix, and slowing or destabilizing convergence. As a result, the CMPLDWG modeling often suffers from slow convergence or divergence and frequent simulation failures [9]. To address these issues, several methods have been proposed to either reduce the effective dimensionality in the estimation process or improve the solver robustness. One class of methods involves the sensitivity-based screening, which eliminates the weakly identifiable parameters based on their effects on post-disturbance trajectories [10], [11]. By identifying the parameters that produce negligible effects, these methods retain only those deemed dynamically significant under a specific event. However, even after the sensitivity-based screening, the remaining parameter estimation problem retains most of the original dimensionality, continuing to pose significant numerical challenges.

Meanwhile, various data-driven methods have been applied to handle the non-convexity of parameter estimation problem. Deep learning-based methods such as amortized inference, deep Q -learning, and reinforcement learning aim to learn approximators from offline data to alleviate the computational burden of simulation during the online inference [9], [12]–[14]. However, they are extremely data- and computation-intensive during the offline training and very sensitive to the hyperparameter tuning. Moreover, their black-box nature makes them almost infeasible to accommodate frequent model updates due to rapidly changing renewable and demand conditions, severely limiting their adaptivity to varying operating points. The metaheuristic algorithms, represented by the genetic algorithm (GA) and particle swarm optimization (PSO), have global search capabilities with less dependence on gradient information [15], [16]. These algorithms, which require no offline dataset generation, can have less computational and memory burden compared with deep learning-based methods. In particular, the enhanced snake optimizer (ESO) has emerged as a promising metaheuristic algorithm with firm performance. Although it has not yet been applied to dynamic load models like CMPLDWG, its biologically inspired design enables a powerful combination of global exploration and adaptive local refinement, offering the potential to navigate the complex estimation landscape of model [17], [18].

Despite these advantages, the existing data-driven methods fail to fundamentally resolve the ill-conditioning in CMPLDWG modeling. The ill-conditioning is referred to as the case where the Jacobian or Hessian matrix has very large condition numbers and nearly collinear sensitivities. As noted by [19] and [20] and corroborated by our numerical tests, the ill-conditioning in CMPLDWG modeling undermines the parameter identifiability and convergence robustness. Consequently, both gradient-based and data-driven methods often exhibit instability and poor generalizability to new scenarios. These limitations highlight the need for a structural approach to address the ill-conditioning fundamentally and explicitly. Inspired by this, we partition the entire parameter space into smaller subsets, thereby decomposing the original high-dimensional estimation problem into multiple tractable subproblems. This grouping idea enables the in-

dependent estimation of each subset, and the subproblem is subject to more favorable numerical conditioning. For effective grouping, the parameters within each subset should have similar scales, while interactions across subsets should be reduced to avoid unstable parameter updates. This interdependency-based parameter grouping offers a unified solution to both requirements by forming subsets of strongly interdependent parameters with small cross-subset interactions.

Building on the interdependency-based parameter grouping, this paper presents a structure-inspired parameter estimation method of CMPLDWG, which systematically addresses the numerical, structural, and scalability challenges inherent in high-dimensional dynamic load modeling. The key contributions are summarized as follows.

- 1) We propose an interdependency-based parameter grouping that partitions the entire parameter space into smaller subsets, effectively reducing cross-subset interaction and mitigating the ill-conditioning in the estimation process.
- 2) The proposed estimation method is applied across both gradient-based methods and metaheuristic algorithms, demonstrating solver-agnostic improvement in the convergence stability and estimation accuracy.
- 3) An initialization strategy, combining pre-disturbance weighted alignment and subset-wise update, is introduced to ensure feasible starting points and improved stability during estimation.
- 4) By leveraging the physical and numerical structure of CMPLDWG, the proposed estimation method provides a scalable and interpretable solution for estimating large-scale dynamic load parameters. Case studies on the New England 68-bus system confirm the substantial improvements in convergence speed and numerical stability, which further enhance the estimation accuracy. The ability of the proposed estimation method in achieving high estimation accuracy is further validated through real-world measurements. Moreover, the proposed estimation method is broadly applicable to future extensions of load models or similar complex data-driven estimation problems in power systems or other dynamic systems.

The remainder of this paper is organized as follows. Section II formulates the problem for CMPLDWG estimation. Section III presents the interdependency-based parameter grouping and the overall estimation method. Section IV provides case studies, followed by conclusions in Section V.

II. PROBLEM FORMULATION FOR CMPLDWG ESTIMATION

The CMPLDWG captures the aggregate dynamic behaviors of loads and DERs at the transmission-level substations [2], [3]. As illustrated in Fig. 1, it combines static ZIP loads and dynamic components. The dynamic components adopt the DER_A model, a simplified yet flexible representation of aggregated DERs that includes active/reactive power control loops, frequency and voltage ride-through logics, and power limits [21], [22]. This model includes 183 nonlinearly and structurally coupled parameters and 25 nonlinear differential equations, exhibiting complex interactions across diverse time scales and control layers.

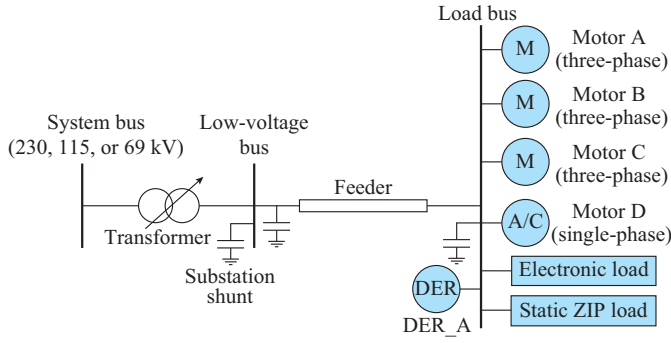


Fig. 1. Schematic diagram of CMPLDWG.

Data-driven load modeling aims to estimate the parameter vector $\theta = [\theta_i]$ that minimizes the mismatch with the measured disturbance responses. Given the measured active and reactive power trajectories, denoted as $\mathbf{P}_{\text{meas}}$ and $\mathbf{Q}_{\text{meas}}$, respectively, we form the residual error $\mathbf{r}(\theta)$ as:

$$\mathbf{r}(\theta) = \mathbf{f}(\theta) - \begin{bmatrix} \mathbf{P}_{\text{meas}} \\ \mathbf{Q}_{\text{meas}} \end{bmatrix} = \begin{bmatrix} \mathbf{P}(\theta) - \mathbf{P}_{\text{meas}} \\ \mathbf{Q}(\theta) - \mathbf{Q}_{\text{meas}} \end{bmatrix} \quad (1)$$

where $\mathbf{f}(\theta)$ denotes the mapping from θ to the simulated active and reactive power trajectories $\mathbf{P}(\theta)$ and $\mathbf{Q}(\theta)$. Following the existing work [7], [9], [20], [23], the nonlinear least-squares (NLS) problem is represented as:

$$\arg \min_{\theta} E(\theta) = \frac{1}{2} \|\mathbf{r}(\theta)\|_2^2 = \frac{1}{2} \mathbf{r}^T(\theta) \mathbf{r}(\theta) \quad (2)$$

where $E(\theta)$ is the objective function of NLS problem.

The estimation of NLS problem for CMPLDWG encounters severe ill-conditioning, affecting the convergence speed and numerical stability of iterative updates. Under such conditions, even small perturbations in initial values or measurement data can lead to significant variations in the estimated parameters, undermining the parameter identifiability and convergence robustness. This primarily arises from two factors: ① a large scaling difference across the parameters; and ② a strong nonlinearity. Although these issues may be less problematic in low-dimensional or weakly coupled models, they could significantly affect the scalability and robustness of the large and complex CMPLDWG.

First, the CMPLDWG exhibits a large scaling difference across its parameters due to substantial differences in physical units, parameter magnitudes, and parameter sensitivities. This difference causes the Jacobian matrix of residual error $\mathbf{r}(\theta)$, denoted as $\mathbf{J}'$, to span several orders of magnitude, resulting in the ill-conditioning. In the NLS problem (2), the $(i,j)^{\text{th}}$ entry of the Hessian matrix $\mathbf{H}^E$ of $E(\theta)$ is given by:

$$H_{ij}^E = \frac{\partial^2 E(\theta)}{\partial \theta_i \partial \theta_j} = \left(\frac{\partial \mathbf{r}}{\partial \theta_j} \right)^T \frac{\partial \mathbf{r}}{\partial \theta_i} + \mathbf{r}^T \frac{\partial^2 \mathbf{r}}{\partial \theta_i \partial \theta_j} \quad (3)$$

A fundamental issue of Gauss-Newton (GN)-based solvers, e. g., LM algorithm, lies in the approximation $\mathbf{H}^E \approx (\mathbf{J}')^T \mathbf{J}'$. This neglects the second-order derivatives in (3) for simplicity. However, the iterative updates become more prone to the ill-conditioning if $\mathbf{J}'$ has a severe imbalance.

Second, a strong nonlinearity in the CMPLDWG further

aggravates the ill-conditioning. $\mathbf{f}(\theta)$ is known to be highly nonlinear and state-dependent, resulting in locally sharp curvature. In regions with locally sharp curvature, the influence of the neglected second-order derivatives in (3) could be significant, and the GN-based solvers may fail to capture the accurate curvature. In practice, the iterative updates would experience poor step size choices and slow convergence. From a metaheuristic perspective, the resulting search landscape is highly nonconvex with flat regions and narrow valleys. Consequently, PSO and GA can exhibit slow convergence or premature convergence, particularly under weak parameter excitation [19].

III. INTERDEPENDENCY-BASED PARAMETER GROUPING AND PROPOSED ESTIMATION METHOD

To tackle the numerical challenges, we develop a structure-inspired parameter estimation method of CMPLDWG via interdependency-based parameter grouping. By partitioning the entire parameter space into smaller subsets, the numerical and convergence instabilities can be effectively eliminated. This partition should ensure that the resulting subsets are weakly coupled with each other.

Figure 2 illustrates the overall three-stage design of proposed estimation method.

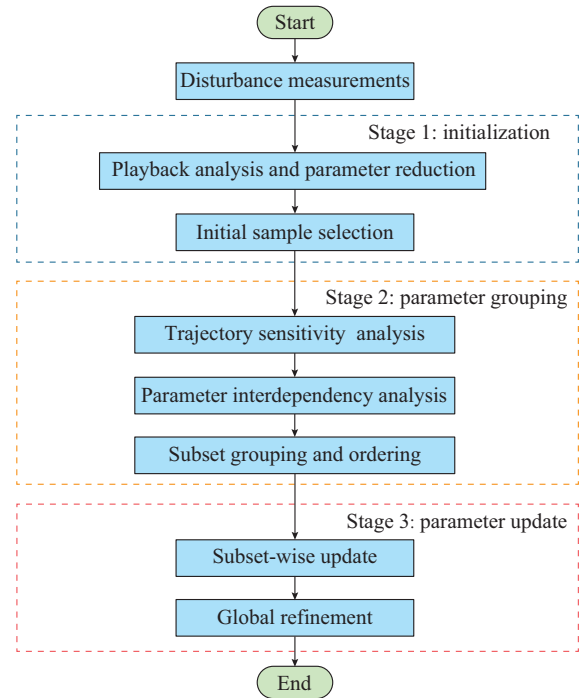


Fig. 2. Overall three-stage design of proposed estimation method.

First, an initialization stage (stage 1) is used to select the initial sample among multiple randomized parameter sets. We prioritize the pre-disturbance weighted alignment to select a physically plausible initial sample based on playback analysis provided by power system simulators [24] and parameter reduction. Next, two diagnostic steps in parameter grouping stage (stage 2) are conducted: ① trajectory sensitivity analysis, which identifies the dominant parameters affecting system dynamics; and ② parameter interdependency

analysis, which captures the pairwise couplings among all the parameters. Based on these two steps, the entire parameter space is systematically divided into smaller subsets of strongly interdependent parameters, while the cross-subset interactions are minimized. A subset ordering is further determined based on their importance. At the parameter update stage (stage 3), an iterative process for estimating each subset separately is developed, i.e., subset-wise update, along with a global refinement step. The primary benefit is that the ill-conditioning can be mitigated by suppressing cross-subset interactions, thereby improving the convergence stability across various optimization solvers.

A. Initialization Stage

Upon obtaining the disturbance measurements, we first perform the playback analysis and parameter reduction. Our numerical analysis shows that only 69 of the total 183 parameters are dynamically relevant and structurally identifiable in typical fault scenarios. The remaining parameters are insignificant for different reasons and thus can be excluded from the estimation task.

To obtain the starting point, we adopt a sampling-based strategy in which multiple parameter sets are randomly generated within their reasonable ranges. The trajectory response of each sample θ is simulated and evaluated using a weighted average of the pre- and post-disturbance residual errors as:

$$\theta_0 = \arg \min_{\theta} \left(\omega_{\text{pre}} \left\| \mathbf{r}_{\text{pre}}(\theta) \right\|_2^2 + \omega_{\text{post}} \left\| \mathbf{r}_{\text{post}}(\theta) \right\|_2^2 \right) \quad (4)$$

where $\mathbf{r}_{\text{pre}}$ and $\mathbf{r}_{\text{post}}$ are the pre- and post-disturbance residual errors, respectively; ω_{pre} and ω_{post} are the pre- and post-disturbance weight hyperparameters, respectively; and θ_0 is the selected initial sample. In this paper, $\omega_{\text{pre}} > \omega_{\text{post}}$ is set to prioritize the pre-disturbance weighted alignment, thereby improving the performance of subsequent stages. The initial sample θ_0 is selected for the subsequent analysis.

B. Parameter Grouping Stage

Parameter grouping stage is designed to achieve effective dimensionality reduction while preserving critical parameter interdependency that affects the numerical conditioning. The first-order trajectory sensitivity analysis provides the evaluation on the importance of individual parameters. Based on the simulated trajectory with a fixed model input, we employ the finite-difference approach to approximate the trajectory sensitivity with respect to parameter θ_i as [20]:

$$\mathbf{J}^f(\theta_i) = \frac{\partial \mathbf{f}(\theta)}{\partial \theta_i} \approx \frac{\mathbf{f}(\theta_0 + \epsilon_i \mathbf{e}_i) - \mathbf{f}(\theta_0 - \epsilon_i \mathbf{e}_i)}{2\epsilon_i} \quad (5)$$

where $\mathbf{J}^f(\theta_i)$ is the trajectory sensitivity with respect to parameter θ_i ; $\mathbf{e}_i$ is the Kronecker vector with 1 at the i^{th} entry and 0 elsewhere; and ϵ_i is the perturbation step size. Note that $\mathbf{J}^f$ is equal to $\mathbf{J}^r$. If $\left\| \mathbf{J}^f(\theta_i) \right\| \rightarrow 0$ and can be negligible, parameter θ_i can be treated as unidentifiable and excluded to improve the parameter identifiability of the estimation problem.

After eliminating unidentifiable parameters, we further partition the remaining parameters into subsets to ensure that

the numerical conditioning within each subset is significantly improved. It is crucial to account for parameter interdependencies. When strongly interdependent parameters are assigned to different subsets, separately optimizing these subsets may significantly alter the descent direction and lead to an unstable trajectory, as detailed in Supplementary Material A. The example in Supplementary Material A indicates that, as the numerical conditioning worsens, the parameter interdependency can largely affect the estimation accuracy. Following [25] and [26], we evaluate the interdependency between any two parameters θ_i and θ_j based on the second-order derivative of $\mathbf{f}(\theta)$, as approximated by:

$$H_{i,j}^f = \frac{\partial^2 \mathbf{f}(\theta)}{\partial \theta_i \partial \theta_j} \approx \frac{1}{4\epsilon_i \epsilon_j} (\mathbf{f}_{++} - \mathbf{f}_{+-} - \mathbf{f}_{-+} + \mathbf{f}_{--}) \quad (6)$$

where $H_{i,j}^f$ is the $(i,j)^{\text{th}}$ entry of the Hessian matrix $\mathbf{H}^f$ of $\mathbf{f}(\theta)$; and $\mathbf{f}_{++} = \mathbf{f}(\theta_0 + \epsilon_i \mathbf{e}_i + \epsilon_j \mathbf{e}_j)$, $\mathbf{f}_{+-} = \mathbf{f}(\theta_0 + \epsilon_i \mathbf{e}_i - \epsilon_j \mathbf{e}_j)$, $\mathbf{f}_{-+} = \mathbf{f}(\theta_0 - \epsilon_i \mathbf{e}_i + \epsilon_j \mathbf{e}_j)$, and $\mathbf{f}_{--} = \mathbf{f}(\theta_0 - \epsilon_i \mathbf{e}_i - \epsilon_j \mathbf{e}_j)$ are computed using the finite-difference rule.

By normalizing the element-wise L1 norm of the vector-valued interaction metric for each parameter pair (θ_i, θ_j) , i.e., $\mathbf{H}_{\theta_i, \theta_j}^f$, the normalized second-order interdependency (nSI) can be obtained as:

$$nSI(\theta_i, \theta_j) = \frac{\left\| \mathbf{H}_{\theta_i, \theta_j}^f \right\|_{1, \text{sum}}}{\max_{i \neq j} \left\| \mathbf{H}_{\theta_i, \theta_i}^f \right\|_{1, \text{sum}}} \quad (7)$$

where $\left\| \cdot \right\|_{1, \text{sum}}$ denotes the element-wise L1 norm, which is the sum of absolute values in a vector; and $nSI(\theta_i, \theta_j)$ is the nSI value for parameter pair (θ_i, θ_j) .

The nSI value characterizes not only the parameter interdependency but also the parameter scaling. This value is related to the sensitivity to the joint changes of two parameters. Thus, the parameter pairs with similar scaling and consistent dynamic influences have higher nSI values, thereby ensuring the scale-consistent grouping. Additionally, the nSI value can approximate the second-order derivatives in (3), which is ignored by GN-based solvers. Therefore, the interdependency-based parameter grouping inherently incorporates the second-order curvature to improve the convergence stability by addressing the model nonlinearity.

By using the nSI value, the interdependency-based parameter grouping could structurally mitigate the ill-conditioning. The core idea is to update subsets of strongly interdependent parameters. Formally, we partition all the parameters into disjoint subsets, with any two subsets S_m and S_n satisfying:

$$nSI(\theta_i, \theta_j) < \zeta \quad \forall \theta_i \in S_m, \forall \theta_j \in S_n \quad (8)$$

where ζ is a pre-selected threshold.

By properly selecting ζ , we could partition all the parameters into K subsets, each corresponding to a lower-dimensional and better-conditioned subproblem. In practice, this is implemented through a hierarchical clustering process. The independent parameters satisfying the threshold are first extracted into a subset and then removed from the pool. This procedure is repeated until all parameters are assigned. Each

subset will be sequentially updated, starting from the one containing the most influential parameters according the trajectory sensitivity analysis. As demonstrated by numerical tests later, the interdependency-based parameter grouping is the key to improving the numerical conditioning. It is worth noting that the parameters can be grouped solely based on the parameter interdependency analysis, even without the domain knowledge of model. Accordingly, the interdependency-based parameter grouping can be applied to any other dynamic model.

C. Parameter Update Stage

The interdependency-based parameter grouping decouples the original high-dimensional estimation problem into K sub-problems. Within each subset, the strongly interdependent parameters are jointly updated while fixing the parameters in other subsets. For this subset-wise update, a variety of data-driven methods can be utilized, among which we select two representative ones, the LM algorithm and ESO, with a trade-off in the computational complexity and accuracy.

The LM algorithm combines the gradient and GN directions by using the diagonal of the approximate Hessian matrix for improved stability. The LM update is expressed as:

$$\begin{cases} \theta_{S_k}^{n+1} = \theta_{S_k}^n - \mathbf{D}^n \mathbf{r}(\theta_{S_k}^n) \\ \mathbf{D}^n = \left((\mathbf{J}^r)^T \mathbf{J}^r + \lambda^n \text{diag}((\mathbf{J}^r)^T \mathbf{J}^r) \right)^{-1} (\mathbf{J}^r)^T \end{cases} \quad (9)$$

where the superscript n denotes the index of iteration; θ_{S_k} is the vector of parameters in subset S_k ; $\mathbf{D}^n$ is the LM update direction matrix; and λ is the damping factor. For simplicity, the matrix $\mathbf{J}^r$ here stands for the Jacobian matrix evaluated at $\theta_{S_k}^n$. Intuitively, the LM update direction is proportional to

the gradient and inversely proportional to the curvature. To further enhance the numerical stability, we can adopt a reduced step size by introducing a scalar coefficient $\alpha \in (0, 1]$:

$$\mathbf{D}^n = \alpha \left((\mathbf{J}^r)^T \mathbf{J}^r + \lambda^n \text{diag}((\mathbf{J}^r)^T \mathbf{J}^r) \right)^{-1} (\mathbf{J}^r)^T \quad (10)$$

This can adaptively control the updated step size while maintaining the descent direction. The LM update is efficient in terms of local convergence at a modest computational cost. However, it is less effective at exploring highly nonlinear or ill-conditioned landscapes because it relies solely on local curvature.

The ESO is a population-based metaheuristic and alternative algorithm when a global parameter search is required. The ESO extends the original snake optimizer (SO) [17], which is inspired by biological behaviors, through several enhancements. Specifically, the ESO incorporates the opposition-based learning to enhance the population diversity, adopts the lens-imaging projection to explore distant search regions, and balances the adaptive exploration-exploitation. In addition, the advanced perturbation operators such as sine-cosine updates, Tent-chaos mapping, and Cauchy mutations are applied to enhance the global search performance [18]. Based on these operators, the ESO randomly partitions the parameters in S_k into two vectors $\theta_{S_k}^f$ and $\theta_{S_k}^m$, which are updated separately. In a nutshell, the ESO is particularly suitable for addressing the highly nonlinear search space in the parameter estimation for CMPLDWG, albeit at the expense of much higher computational complexity than LM updates.

To support the stable and effective optimization within each subset, we apply a targeted re-initialization procedure prior to the subset-wise update. The schematic diagram of subset estimation process is shown in Fig. 3.

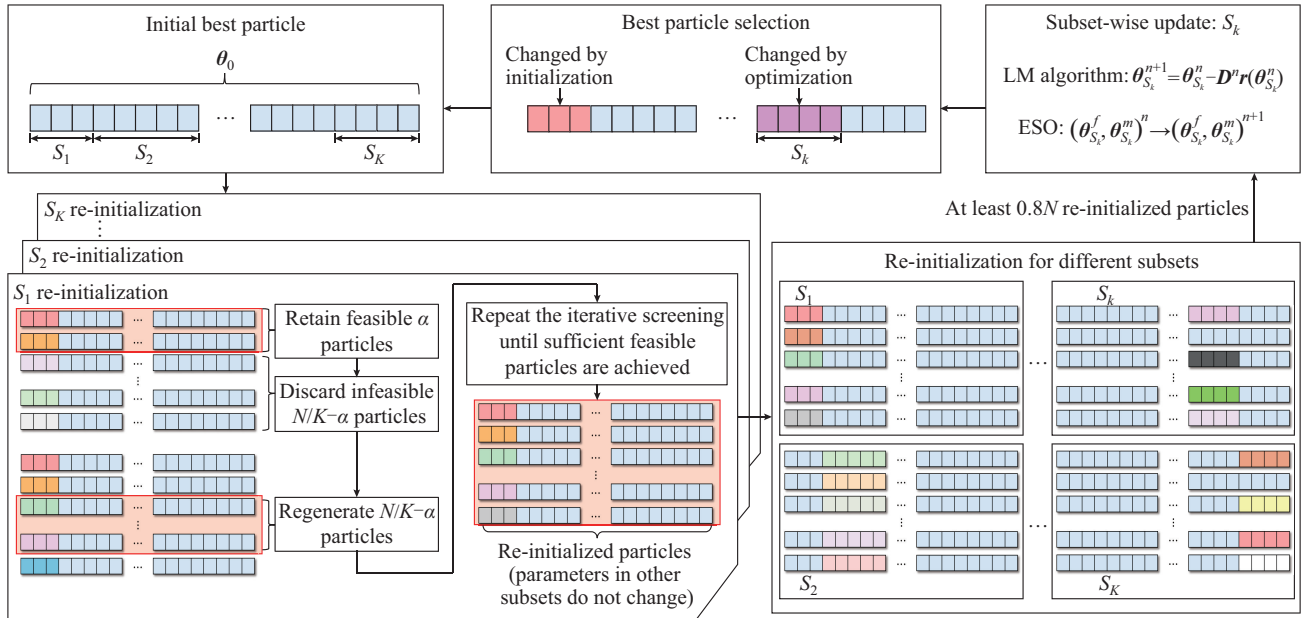


Fig. 3. Schematic diagram of subset estimation process.

By perturbing θ_0 within predefined parameter bounds, N/K particles are randomly generated for each subset S_k . Particles that produce unstable or infeasible simulation trajectories are

discarded, and the remaining α particles are retained. If the number of feasible particles falls below the threshold $0.8N/K$, the sampling space is adaptively tightened around

previously retained samples. This iterative screening continues until sufficient feasible candidates are obtained, providing both reliable starting points for local solvers and sufficient diversity for heuristic search. Upon completing the subset-wise updates, the best particle is selected to serve as the initial reference for the next iteration.

A global refinement step is then implemented to fine-tune the entire parameter space. It essentially performs updates for all parameters together within their allowable ranges since the subset-wise updates have obtained sufficiently accurate estimation. For the LM-based refinement, the allowable ranges are determined based on the iterative updates. For the ESO-based refinement, the search domain is given by the final population dispersions across all subsets. The global refinement step ensures the consistency and global alignment across the subsets.

IV. CASE STUDIES

A. Simulation Environment and Setup

To validate the proposed estimation method, we conduct case studies using the New England 68-bus system with a 138 kV base voltage and a total load of approximately 17 GW. In CMPLDWG, we remove unidentifiable parameters via an initial screening process, leaving 69 parameters as unknown ones. A complete list of estimation parameters along with representative nominal values is provided in Supplementary Material B. The DER_A fraction is assumed to be known, following the method in [27]. Except for physical bounds defined in the WECC guidelines [28], [29], no prior knowledge about the parameters is assumed. The simulation is performed via a Python interface connected to the PSS/E simulator, running on a computer system equipped with an AMD Ryzen-7 3700X 3.60-GHz CPU and 32 GB RAM.

We perform the parameter estimation in five simulation scenarios and real-world measurements. Scenarios 1-4 focus on estimating the composite load at Bus 41, while Scenario 5 focuses on Bus 29.

1) In Scenario 1, a three-phase short circuit is applied on Bus 14 at $t=1$ s and cleared at $t=1.1$ s to induce system-wide disturbances. The composite load at Bus 41 comprises load demand of $(1000+j300)$ MVA and DER generation of $(200+j25)$ MVA. Scenario 1 represents a simplified case, where only a single parameter, i. e., the motor A fraction (FmA), is treated as unknown during estimation.

2) Scenario 2 performs estimation over all 69 parameters, with all other settings identical to those in Scenario 1.

3) In Scenario 3, a line fault is set on the branch between Buses 41 and 42, resulting in a more localized excitation pattern. In addition to increasing the DER generation at Bus 41 to $(400+j40)$ MVA, additional DERs are assumed to be installed at Buses 29 and 42, all represented by the CMPLDWG, and the total DER generation is $(650+j65)$ MVA. Scenario 3 is designed to demonstrate the generalizability of the proposed estimation method.

4) Based on the configuration of Scenario 3, the DER generation at Buses 29, 39, and 42 is gradually increased by $(50+j5)$ MVA, yielding an additional total DER generation of

$(150+j15)$ MVA at each step. Simulation is carried out when the total DER generation increases from $(800+j80)$ MVA to $(1400+j140)$ MVA, and the resulting deviations in active and reactive power at Bus 41 are illustrated by the solid lines in Fig. 4. In general, the power deviations become progressively larger as the total DER generation increases, but the overall influence remains relatively limited. When the increased DER generation is concentrated at Bus 42 (the adjacent bus of Bus 41), significantly larger deviations in active and reactive power at Bus 41 are observed, as illustrated by the dotted lines in Fig. 4. This indicates stronger dynamic interactions with the composite load at Bus 41. Accordingly, to verify the robustness of the proposed estimation method, we select the case with concentrated DER generation increments at Bus 42, with a total DER generation of $(1400+j400)$ MVA, as shown in Scenario 4.

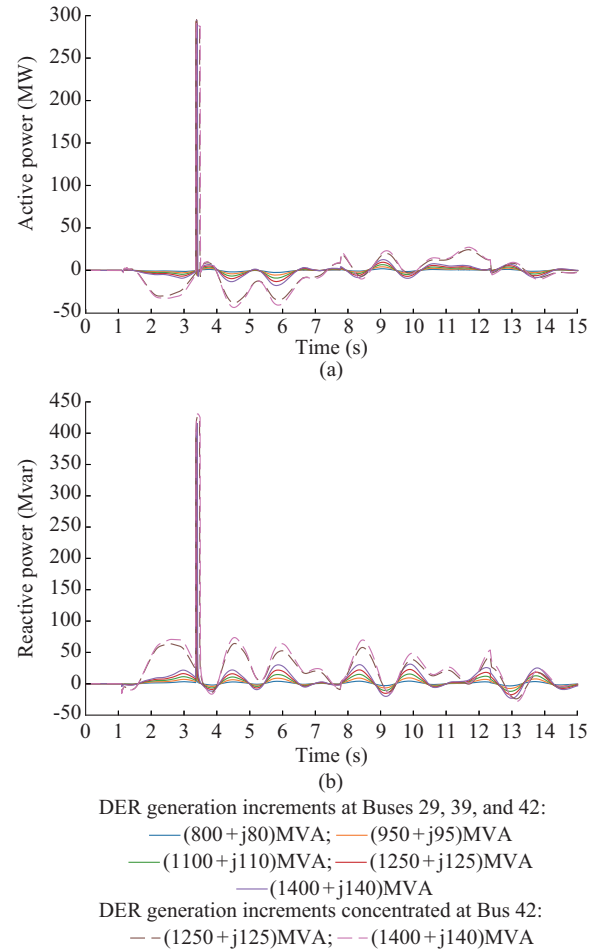


Fig. 4. Deviations in active and reactive power at Bus 41 with different total DER generation. (a) Active power. (b) Reactive power.

5) In Scenario 5, the total DER generation is $(800+j80)$ MVA, while the estimation target is changed to the composite load at Bus 29. This scenario is designed to evaluate the ability of the proposed estimation method to achieve the accurate parameter estimation under realistic operating conditions. To emulate the realistic measurement conditions that often involve significant noise, white Gaussian noise is added to the voltage, active power, and reactive power ob-

tained from the simulation before performing estimation. Furthermore, to reflect the practical situations where a good initial point may not be available, the estimation is intentionally initialized using the lower bound for the allowable range of each parameter, as shown in Supplementary Material C. In addition, we also incorporate real-world measurements from PQube devices to assess the proposed estimation method.

B. Effect of Interdependency-based Parameter Grouping on Estimation Stability

To evaluate the role of interdependency-based parameter grouping, we first consider Scenario 1, where only one parameter, i.e., FmA, is estimated while all other parameters are initialized at the ground-truth values. This isolates the local convergence behavior without the confounding effect of high dimensionality. Figure 5 compares the estimation results using LM algorithm with and without interdependency-based parameter grouping in Scenario 1, with FmA initialized at 0.9.

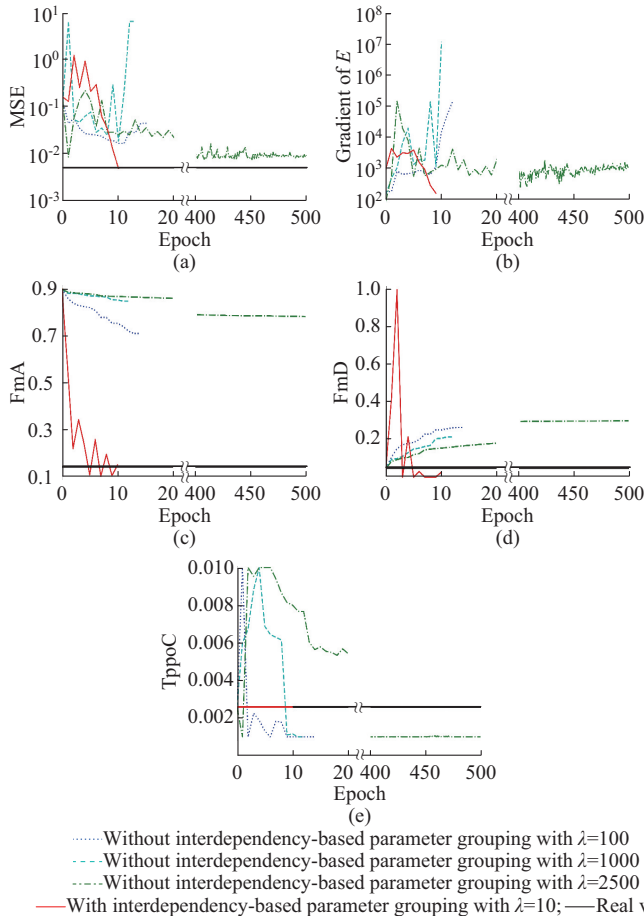


Fig. 5. Comparison of estimation results using LM algorithm with and without interdependency-based parameter grouping in Scenario 1. (a) Mean squared error (MSE). (b) Gradient of E . (c) Change of FmA. (d) Change of FmD. (e) Change of TppoC.

As illustrated in Fig. 5, the estimation without interdependency-based parameter grouping exhibits poor robustness. The convergence is highly sensitive to the damping factor λ , and the LM algorithm frequently diverges or oscillates when

$\lambda < 2500$. The convergence instability is attributed to the residual cross-coupling of FmA with other fixed yet sensitive parameters in the same subset, e.g., motor D fraction (FmD) and open-circuit transient time constant of motor C (TppoC), which distort the error surface. In several runs, FmA is driven away from its real value as it compensates for mismatches introduced by these fixed yet sensitive parameters.

In contrast, the estimation with interdependency-based parameter grouping isolates FmA with its most directly coupled parameters, enabling the localized optimization within a better-conditioned subspace. This enables the successful recovery of the real value of FmA, smoother convergence, and significantly lower estimation error. However, the instability persists due to abrupt changes in parameters such as FmD in the same subset, highlighting the limitations of gradient-based updates in nonlinear regions. Overall, the interdependency-based parameter grouping improves the local stability but does not entirely eliminate the convergence instability under the strong nonlinearity.

C. Initialization Evaluation

Figure 6 compares the initialization results under varying pre-disturbance weights r_{pre} . The weighted MSE over multiple random samples with 69 parameters unknown is computed to assess the effect of pre-disturbance weight. As illustrated in Fig. 6(a), increasing the pre-disturbance weight r_{pre} improves the initialization quality. Notably, the case with $r_{\text{pre}} = 100$ yields the lowest median weighted MSE, the narrowest interquartile range, and the fewest extreme outliers. The proportion of samples with a weighted MSE below 100, which indicates almost zero error in the steady state, increases from 14% with $r_{\text{pre}} = 1$ to 28% with $r_{\text{pre}} = 100$, and then decreases to 20% with $r_{\text{pre}} = 1000$. This trend confirms that emphasizing the pre-disturbance weighted alignment improves the probability of obtaining physically plausible and dynamically stable parameter sets.

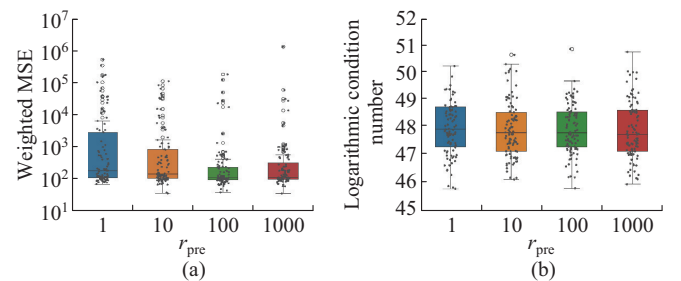


Fig. 6. Comparison of initialization results under varying pre-disturbance weights. (a) Weighted MSE. (b) Logarithmic condition number.

Additionally, higher r_{pre} values alleviate the ill-conditioning by anchoring the initial parameter set closer to the feasible operating region. Logarithmic condition numbers of the Jacobian matrices show consistent reductions in both median values and dispersion as r_{pre} increases, as shown in Fig. 6(b). This indicates that a better pre-disturbance weighted alignment improves the sensitivity scaling and positions the system near a well-conditioned region of the parameter space. Accordingly, $r_{\text{pre}} = 100$ is adopted in subsequent studies for its favorable balance between the accuracy and robustness.

D. Results of Interdependency-based Parameter Grouping

Figure 7 shows the heat maps of nSI values within the entire parameter space in Scenarios 2 and 3. The nSI values in Fig. 7 reveal structurally consistent but disturbance-dependent interaction patterns in both scenarios. Based on the interdependency analysis, the parameters are grouped into six subsets (as shown by the red boxes) for Scenarios 2 and 3, using hierarchical clustering with a threshold of $\zeta=0.2$, as shown in Table I. Note that the parameters in each subset from bottom to top and from left to right in Fig. 7 are consistently ordered with those from left to right in Table I.

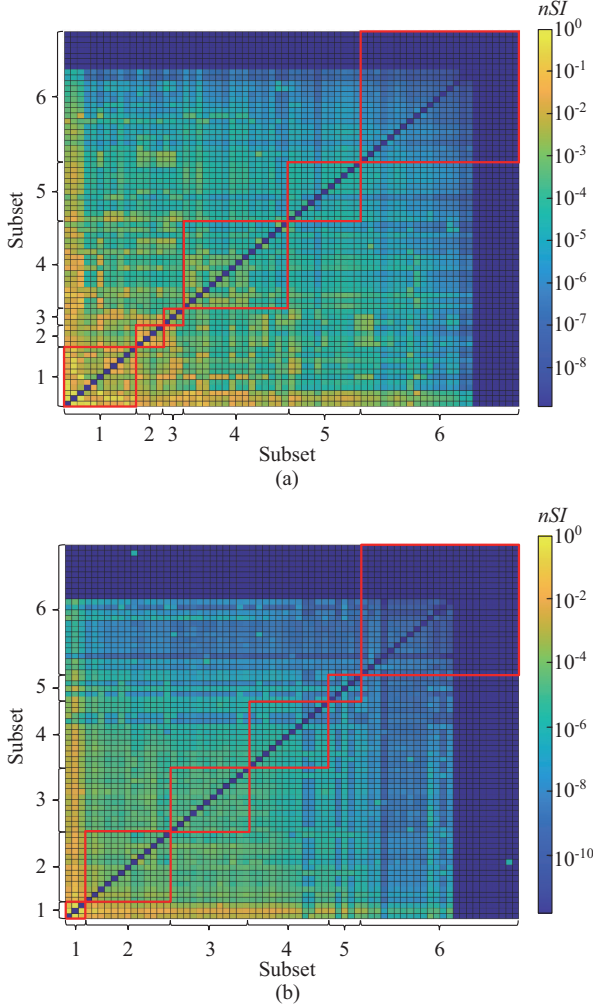


Fig. 7. Heat maps of nSI values within entire parameter space in Scenarios 2 and 3. (a) Scenario 2. (b) Scenario 3.

In both scenarios, the DER-related parameters exhibit low interdependency, while motor-related and static load-related parameters form denser interaction clusters. These results confirm that the motor-related dynamics consistently dominate the interdependency, while the DER-related interdependency remains weak across different fault types in Scenarios 2 and 3. In addition, Scenario 2 displays more extensive and stronger interdependencies. In contrast, Scenario 3 presents more localized interactions due to the limited propagation of the line fault.

TABLE I
PARAMETER SUBSETS FOR SCENARIOS 2 AND 3 BASED ON
INTERDEPENDENCY ANALYSIS

Scenario	Subset	Parameter
2	Subset 1	TppoA, TppoC, TppoB, LpA, LppA, FmB, LppC, RaC, LC, RaA, LpC
	Subset 2	Fel, Pfel, frecel, FmA
	Subset 3	FmD, FmC, TpoC
	Subset 4	LB, TpoB, LA, TpoA, LpB, LFmB, RaB, HA, LppB, HB, LFmA, PFs, Q1c, Tiq, Np2, Nq2
	Subset 5	LFmC, HC, Trv, Tg, LsC, Kp1, Np1, etrqA, Trf, Kpg, P2c
	Subset 6	P1c, Q2c, Tp, Kq1, Nq1, Tf, CmpKpf, LsA, LsB, etrqB, etrqC, CmpKqf, Tpod, Kq2, Kp2, Ddn, Kig, Tv, Kqv, Dup, Vfrac, rrpwr, Tv, Xe
3	Subset 1	TppoA, TppoB, TppoC
	Subset 2	LA, LpA, LppA, LC, HA, LB, LFmB, Fel, Pfel, LppB, FmD, RaA, RaC
	Subset 3	LpB, LFmA, TpoA, FmC, LFmC, RaB, TpoB, LppC, TpoC, frecel, LpC, HC
	Subset 4	FmA, FmB, LsB, HB, etrqA, Tp, Q1c, etrqB, P1c, Trv, Q2c, LsC, PFs
	Subset 5	P2c, Trf, Kpg, Nq1, Tpod
	Subset 6	etrqC, Tg, Ddn, CmpKqf, Kp1, Tf, Tiq, Np1, CmpKpf, LsA, Kig, Kq1, Dup, Tv, Kp2, Np2, Kq2, Nq2, Kqv, Vfrac, rrpwr, Tv, Xe

Note: Trv, Trf, Kqv, Tp, Tiq, Ddn, Dup, Tpod, Kpg, Kig, Vfrac, Tg, rrpwr, Tv, and Xe are the DER-related parameters; LfmA, RaA, LA, LsA, LpA, TpoA, TppoA, HA, etrqA, LfmB, RaB, LB, LsB, LpB, TpoB, TppoB, HB, etrqB, LfmC, RaC, LC, LsC, LpC, TpoC, TppoC, HC, and etrqC are the motor A-, B-, and C-related parameters; Tv, Tf, Kp1, Np1, Kq1, Nq1, Kp2, Np2, Kq2, Nq2, CmpKpf, and CmpKqf are the motor D-related parameters; FmA, FmB, FmC, FmD, and Fel are the fraction-related parameters; and Pfel, frecel, PFs, P1c, P2c, Q1c, and Q2c are the static or electronic load-related parameters.

Despite differences in the fault types and system responses, Scenarios 2 and 3 yield consistent patterns aligned with the physical hierarchy of CMPLDWG. Subset 2 consistently isolates the electronic load-related parameters, indicating their weak interactions with other parameters. Subset 6 also shows stable clustering of DER-related parameters, reflecting their modular roles and low sensitivity to fault types. Fraction-related parameters exhibit partial consistency in Scenarios 2 and 3, highlighting their common influence on the composite load behavior. However, the scenario-dependent variations are also observed. In Scenario 3, the sub-transient time constants TppoA, TppoB, and TppoC are tightly grouped, reflecting a localized dynamic response. In Scenario 2, these parameters are dispersed due to the widespread excitation. This fragmentation observed in Scenario 2 contrasts with the spatially confined subsets in Scenario 3.

To quantitatively assess the benefit of interdependency-based parameter grouping, Table II reports the interdependency-to-sensitivity ratio (ISR), defined as k_2/S , in Scenarios 2 and 3 under the initial parameter set and subsets 1-6, where k_2 is the interdependency metric that captures the off-diagonal dominance in the Hessian matrix $\mathbf{H}^E$, computed via eigenvalue distribution; and S is the aggregated sensitivity from the diagonal entries of $\mathbf{H}^E$, as detailed in [30]. The ISR

thus characterizes the numerical conditioning, where values below the threshold 100 generally indicate the well-conditioned estimation problems. In Scenario 2, the initial parameter set yields an ISR of 2.34×10^5 , reflecting the severe parameter collinearity and ill-conditioning. The interdependency-based parameter grouping substantially improves the numerical properties: subsets 1-3 exhibit ISR values of 1.93×10^{-1} , 1.56×10^{-1} , and 5.79×10^{-3} , respectively, demonstrating reductions of up to five orders. Subset 4 also exhibits acceptable numerical conditioning, while subsets 5 and 6 display higher ISR values, primarily due to their extremely low sensitivity. Nevertheless, their limited impact on the objective function mitigates the risk of numerical instability. In Scenario 3, a similar trend can be observed. Without the interdependency-based parameter grouping, the initial parameter set yields an ISR of 3.56×10^4 , again indicating ill-conditioning. However, the interdependency-based parameter grouping yields significant improvements across most subsets. Only subset 6 exceeds the threshold 100, with an ISR of 8.08×10^3 . Like Scenario 2, this is attributed to very low sensitivity rather than ill-conditioning of the original problem. These results collectively validate that the interdependency-based parameter grouping consistently enhances the numerical stability across different scenarios, primarily by isolating locally well-conditioned and influential parameters.

TABLE II
ISR IN SCENARIOS 2 AND 3 UNDER INITIAL PARAMETER SET
AND SUBSETS 1-6

Scenario	Set	k_2	S	k_2/S
2	Initial parameter set	2.85×10^{11}	1.22×10^6	2.34×10^5
	Subset 1	8.62×10^4	4.47×10^5	1.93×10^{-1}
	Subset 2	1.15×10^5	7.38×10^5	1.56×10^{-1}
	Subset 3	2.30×10^2	3.99×10^4	5.79×10^{-3}
	Subset 4	2.38×10^4	1.17×10^3	2.03×10^1
	Subset 5	6.77×10^4	5.06×10^2	1.34×10^2
	Subset 6	4.11×10^8	5.47×10^2	7.53×10^5
3	Initial parameter set	7.31×10^{14}	2.05×10^{10}	3.56×10^4
	Subset 1	1.15×10^3	8.25×10^9	1.39×10^{-7}
	Subset 2	4.63×10^5	2.37×10^8	1.95×10^{-3}
	Subset 3	2.09×10^6	4.92×10^7	4.26×10^{-2}
	Subset 4	3.43×10^5	4.31×10^5	7.96×10^{-1}
	Subset 5	3.18×10^2	6.56×10^1	4.85×10^0
	Subset 6	2.60×10^7	3.22×10^3	8.08×10^3

The trajectory sensitivity analysis is conducted to quantify the influence of each parameter and determine the order of subsets for parameter estimation. Figure 8 presents the normalized sensitivities of all parameters.

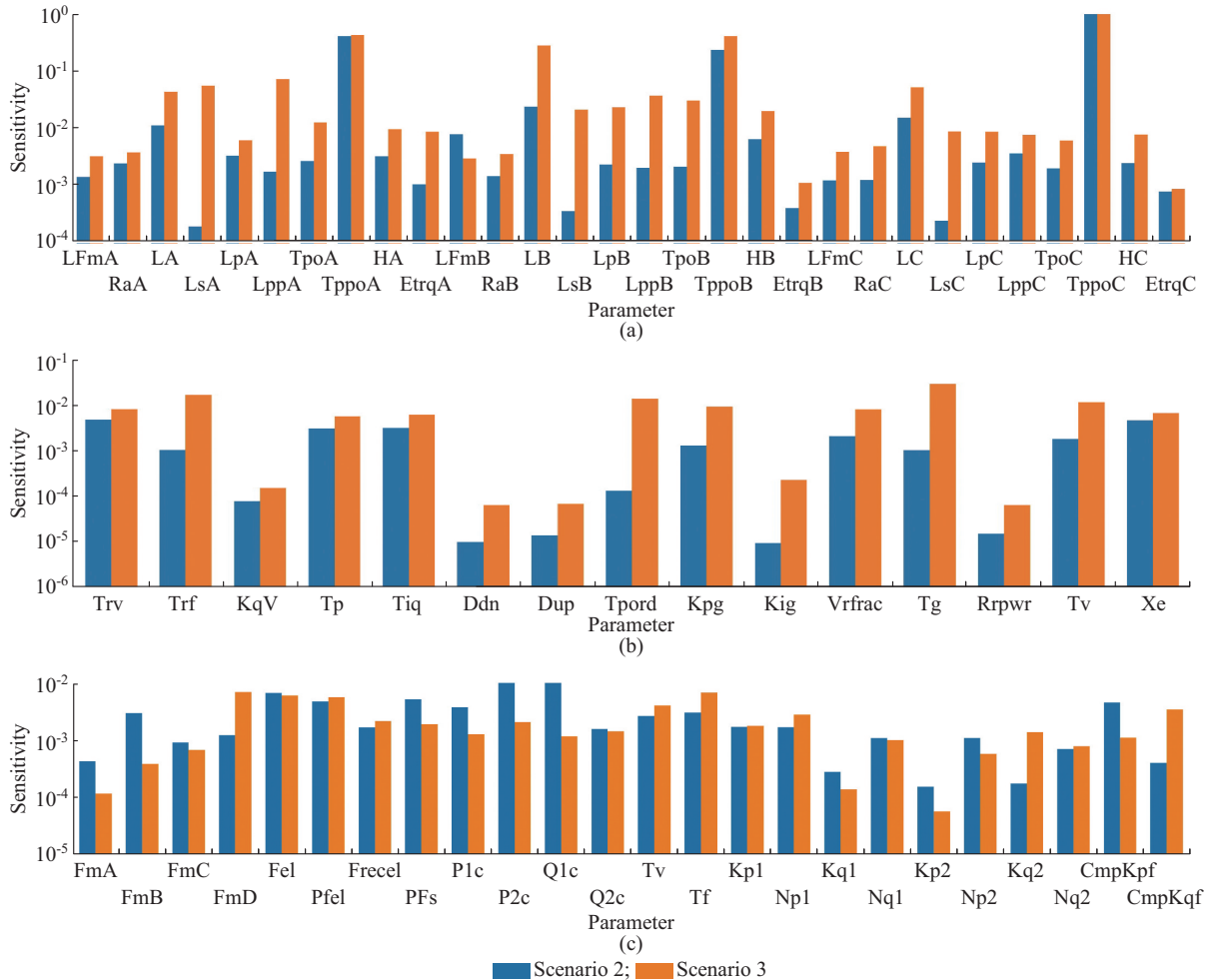


Fig. 8. Normalized sensitivities of all parameters. (a) Motor A-, B-, and C-related parameters. (b) DER-related parameters. (c) Other parameters.

In Scenarios 2 and 3, the motor A-, B-, and C-related parameters, particularly the sub-transient time constants T_{ppoA} , T_{ppoB} , and T_{ppoC} , consistently dominate the sensitivity distribution, confirming their critical roles during the faults. The DER-related parameters exhibit higher sensitivities in Scenario 3 than that in Scenario 2. This difference likely reflects the larger DER generation adopted in Scenario 3, which amplifies the impact of DER control parameters on system-wide voltage regulation. Based on these normalized sensitivities, the subset ordering for sequential estimation is determined to prioritize the most influential subsets. In Scenario 2, the optimal order is subsets 1, 4, 5, 2, 3, and 6, reflecting the dominant motor dynamics followed by load composition and DER contributions. In Scenario 3, the optimal order is adjusted to subsets 1, 2, 3, 6, 4, and 5, ensuring that the DER-related parameters are estimated earlier.

E. Parameter Estimation Results

The performance of the proposed estimation method is validated by jointly analyzing the convergence stability and estimation accuracy. Figure 9 shows the convergence of residual errors with and without interdependency-based parameter grouping in Scenarios 2 and 3 across various optimization solvers including LM algorithm, GA, PSO, and ESO. Note that Scenario 2 is initialized with a parameter set that has a higher initial error to better evaluate the solver robustness under greater parameter uncertainties.

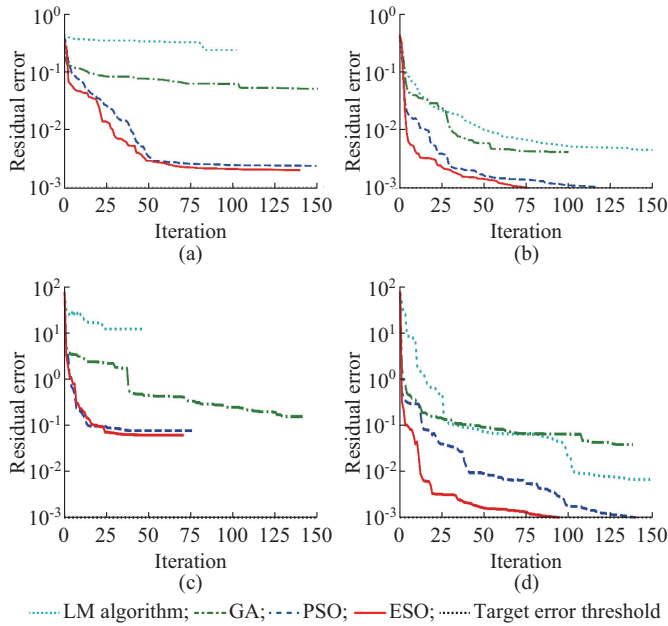


Fig. 9. Convergence of residual errors with and without interdependency-based parameter grouping in Scenarios 2 and 3 across various optimization solvers. (a) Without interdependency-based parameter grouping in Scenario 2. (b) With interdependency-based parameter grouping in Scenario 2. (c) Without interdependency-based parameter grouping in Scenario 3. (d) With interdependency-based parameter grouping in Scenario 3.

As shown in Fig. 9, without the interdependency-based parameter grouping, all optimization solvers exhibit poor convergence in both scenarios. The LM algorithm converges quickly due to the severe parameter collinearity and ill-condi-

tioning. GA becomes trapped in flat regions within the search space, often converging to solutions with relatively large residual errors. PSO and ESO exhibit significant improvements, with ESO achieving slightly higher estimation accuracy and faster convergence speed. However, both ESO and PSO still fail to reach the target error threshold.

With the interdependency-based parameter grouping, the convergence is dramatically stabilized across all optimization solvers in both scenarios. The LM algorithm shifts from convergence stagnation to achieving stable and monotonic error reduction. GA exhibits smoother convergence, but remains less accurate in Scenario 2. PSO and ESO demonstrate significant improvements, achieving highly accurate solutions, with ESO consistently outperforming PSO in the estimation accuracy and convergence speed. This can also be validated through Fig. 10, which shows the final estimation results of active and reactive power trajectories with interdependency-based parameter grouping in Scenarios 2 and 3 across various optimization solvers.

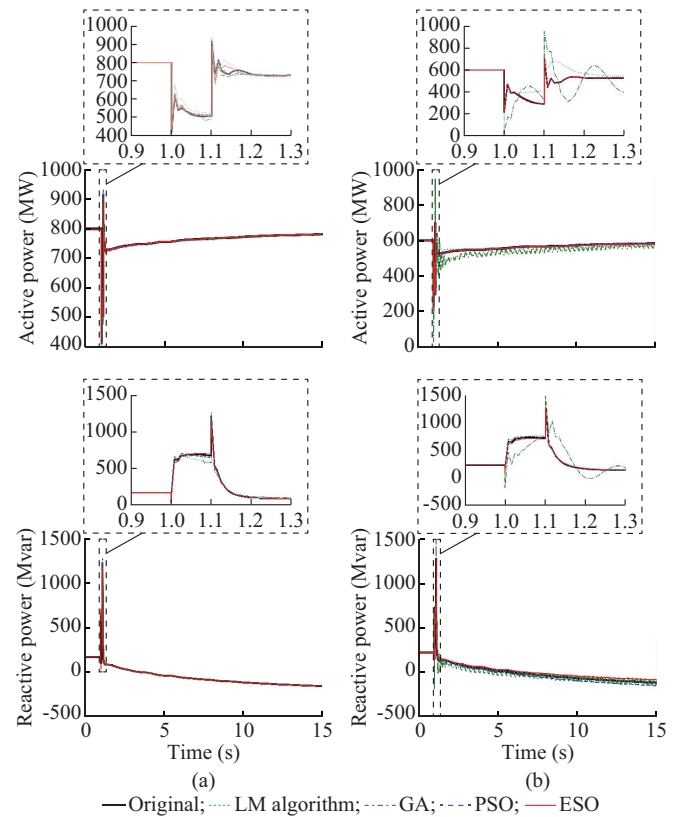


Fig. 10. Final estimation results of active and reactive power trajectories with interdependency-based parameter grouping in Scenarios 2 and 3 across various optimization solvers. (a) Scenario 2. (b) Scenario 3.

The significant improvements with interdependency-based parameter grouping become even more pronounced in Scenario 2, where the initial condition is more challenging. In this scenario, the ability of interdependency-based parameter grouping in suppressing parameter interactions and reducing the search complexity plays a more critical role in stabilizing convergence and guiding optimization solvers toward feasible regions of the solution space. As illustrated in Fig. 10, with interdependency-based parameter grouping, the LM al-

gorithm achieves practically acceptable estimation results, and PSO and ESO consistently deliver near-perfect reconstructions of key dynamic trajectories in both scenarios; while GA remains less accurate in Scenario 2. These results highlight that the interdependency-based parameter grouping can effectively enhance the parameter identifiability while suppressing the parameter interactions.

Figure 11 shows the final estimation results of active and reactive power trajectories with interdependency-based parameter grouping in Scenarios 4 and 5 with ESO. Compared with Fig. 10(b), Fig. 11(a) shows that, although the composite load at Bus 41 remains unchanged, the system exhibits a substantially different dynamic response during both the disturbance and recovery periods in Scenario 4. The resulting MSE is approximately 0.0019 in Scenario 4, which is slightly higher than those in Scenarios 2 and 3. This can be attributed to the more pronounced variation of the dynamic response in Scenario 4. Nevertheless, the proposed estimation method still achieves highly accurate estimation despite such distinct system behaviors.

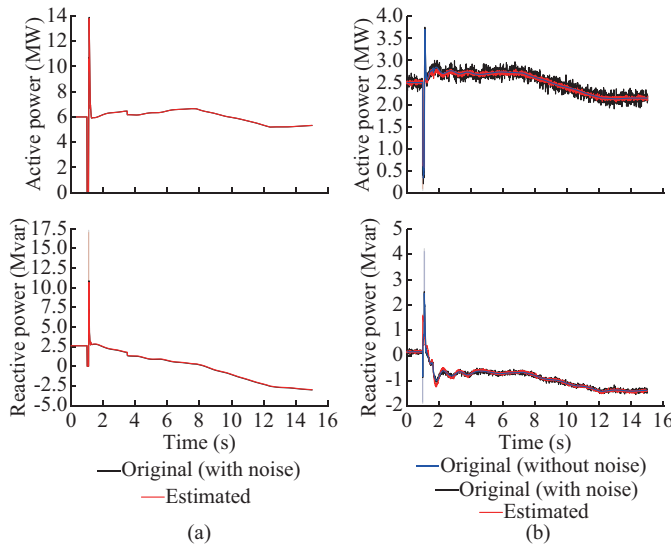


Fig. 11. Final estimation results of active and reactive power trajectories with interdependency-based parameter grouping in Scenarios 4 and 5 with ESO. (a) Scenario 4. (b) Scenario 5.

In addition, the influence of ζ on the estimation accuracy and computational efficiency is analyzed in Scenario 4. When $\zeta = 0.2, 0.5, 0.6$, and 0.8 , the numbers of subsets are 5, 6, 7, and 8, respectively. All four ζ values exhibit stable convergence toward nearly the same final value. However, as the number of subsets increases, the elapsed time per epoch shows an increasing trend from approximately 130 s to 178 s, while the total number of epochs to convergence decreases from 82 to 65. Consequently, the total elapsed time remains nearly constant at around 11000 s, indicating that the overall computational efficiency is not significantly compromised. Hence, in Scenario 4, ζ between 0.2 and 0.8 can be flexibly utilized without significant degradation in estimation accuracy and computational efficiency. This robustness can be explained by the relatively weak sensitivity of interdependency-based parameter grouping within this range. How-

ever, extremely large or small ζ values may induce extreme parameter grouping, which will affect both the estimation accuracy and computation time.

To emulate more realistic operating conditions, Scenario 5 introduces significant measurement noise, and the original active and reactive power trajectories with noise are illustrated in Fig. 11(b). Despite the presence of noise, the proposed estimation method successfully reconstructs the active and reactive power trajectories that closely follow the original ones. Furthermore, despite starting from a poor initial point, the MSE stabilizes rapidly, which would be attributed to the effectiveness of the initialization strategy as well as the strong convergence capability of ESO. The final estimation achieves an MSE of approximately 0.035, which is higher than that in Scenarios 2-4 but still demonstrates strong estimation accuracy under the noisy environment.

The effect of the population size N of ESO on the convergence speed and computational efficiency is further investigated in Supplementary Material C. The results indicate that $N = 120$ provides a favorable balance between the convergence speed and computational efficiency in Scenario 5.

The effectiveness of the proposed estimation method is further validated using real-world measurements from PQube devices. The original voltage, active power, and reactive power trajectories are obtained from field measurements. The playback analysis is conducted such that the measured voltage is applied as the grid-side input, and the estimation process seeks the parameter set that minimizes the error between the measured active and reactive power and the estimated one. In addition to PQube devices, phasor measurement unit (PMU) data can also be utilized for this purpose. When higher-resolution measurements are available, the data can be down-sampled and effectively employed in the parameter estimation process without loss of generality.

The final estimation results of active and reactive power trajectories under datasets 1 and 2 with ESO are shown in Fig. 12, where dataset 1 exhibits two consecutive voltage drops, while dataset 2 shows an abnormal recovery process due to delayed voltage recovery. As reported in [20], directly using raw measurements in parameter estimation can lead to unreasonable results. In our experiment, this issue arises in dataset 2; accordingly, the preprocessed data obtained after smoothing have been used for estimation. The MSEs for datasets 1 and 2 are 0.173 and 3.15, respectively, with the latter exhibiting higher estimation error.

Compared with the simulation results, the results using real-world measurements show lower numerical stability, particularly under rapid dynamic changes, where larger deviations are observed. These discrepancies may arise from the limitations of the composite load model itself, which has a limited freedom to fully express the load dynamics. Nevertheless, the proposed estimation method successfully provides estimation results that follow the overall measured trend well.

In summary, the above findings confirm that the interdependency-based parameter grouping substantially improves both the convergence stability and estimation accuracy across various optimization solvers in different scenarios.

This grouping transforms the original high-dimensional estimation problem into multiple tractable subproblems, allowing optimization solvers to operate more reliably and efficiently even under severe parameter uncertainties. Moreover, its applicability to real-world measurements underscores its robustness and practical utility for high-dimensional parameter estimation.

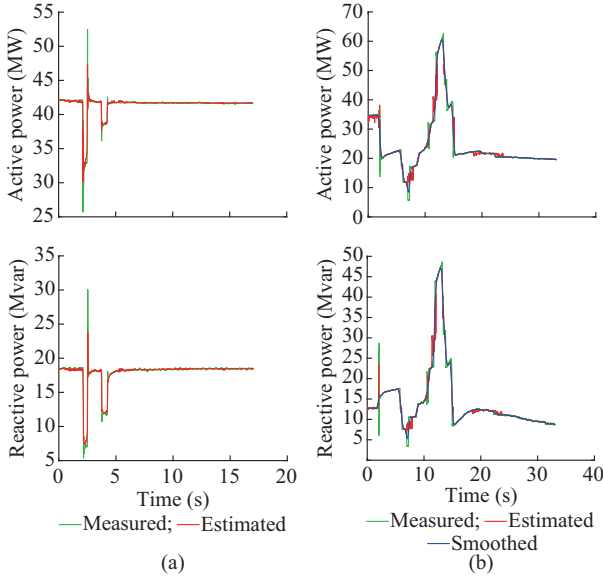


Fig. 12. Final estimation results of active and reactive power trajectories under datasets 1 and 2 with ESO. (a) Dataset 1. (b) Dataset 2.

F. Key Limitations and Practical Considerations

Despite the demonstrated benefits of the proposed estimation method, several limitations should be noted.

First, the interdependency-based parameter grouping is inherently dependent on the specific disturbance used for trajectory sensitivity analysis and parameter interdependency analysis. Since both analyses are conducted based on parameters in particular scenarios, the resulting subsets may not fully capture the parameter interactions under different operating conditions or fault types that activate different dynamic mechanisms. Consequently, multiple disturbance scenarios may be necessary to ensure comprehensive parameter identifiability and robust grouping applicable to a broader range of operating conditions.

Second, the interdependency-based parameter grouping reduces the problem dimensionality within each subset and alleviates the memory limitations that may arise when estimating the complete parameter set simultaneously. However, the additional computation burden required for interdependency analysis and multiple subset-wise update rounds increase the total estimation time. As the number of subsets grows, the cumulative solver executions may offset the efficiency gains in solving individual subproblems. Balancing the grouping granularity, parameter interdependency analysis, and total computational time remains a key consideration for scaling the proposed estimation method to even larger or more complex load models.

Finally, it is important to note that the composite load model inherently exhibits time-varying characteristics. This

implies that, unlike conventional estimation methods that rely on a fixed set of parameters estimated from a single disturbance, continuous or probabilistically informed estimation methods are required. One possible direction is to improve the estimation speed, enabling continuous parameter updates. With sufficient data and accumulated estimation results, the feasible boundaries or probabilistic distributions of individual parameters can be characterized. Once these narrowed parameter boundaries are established, subsequent estimations can be performed within the confined regions, thereby achieving significantly faster convergence. Furthermore, by leveraging these outcomes, neural network-based models such as those presented in [9] could be trained to perform quasi-real-time parameter estimation, thereby enhancing the practicality of the proposed estimation method for field deployment.

V. CONCLUSION

This paper presents a structure-inspired parameter estimation method of CMPLDWG via interdependency-based parameter grouping. By dividing the entire parameter space into smaller subsets, the original high-dimensional and tightly coupled estimation problem is transformed into multiple tractable subproblems. The interdependency-based parameter grouping improves the numerical conditioning and enables the stable and accurate parameter estimation across various optimization solvers including LM algorithm, GA, PSO, and ESO. Case studies of the New England 68-bus system in various scenarios demonstrate that the interdependency-based parameter grouping significantly improves both convergence stability and estimation accuracy. Validation using real-world measurements from PQube devices further demonstrates that the proposed estimation method yields accurate estimation, even under atypical dynamic events. The proposed estimation method not only enables the accurate estimation for CMPLDWG but is also extensible to other composite or measurement-driven load models, offering a scalable solution for dynamic models in increasingly complex power systems.

REFERENCES

- [1] Y. Lin and J. Wang, "A generic dynamic load model framework," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 2, pp. 399-406, Mar. 2022.
- [2] WECC, "WECC dynamic composite load model specifications," Western Electricity Coordinating Council (WECC), Tech. Rep., 2015.
- [3] Z. Ma, Z. Wang, Y. Wang *et al.*, "Mathematical representation of WECC composite load model," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 5, pp. 1015-1023, Sept. 2020.
- [4] P. Pourbeik, J. Weber, D. Ramasubramanian *et al.*, "An aggregate dynamic model of distributed energy resources or power system stability studies," *CIGRE Science & Engineering*, vol. 14, pp. 38-48, Jun. 2019.
- [5] J.-K. Kim, S. Lee, J. S. Kim *et al.*, "The need for modeling the impact of behind-the-meter generation trip on primary frequency response through operational experiences in Korea power system," *IEEE Transactions on Power Systems*, vol. 37, no. 2, pp. 1661-1664, Mar. 2022.
- [6] N. Rao and S. Tripathy, "Power system static state estimation by the Levenberg-Marquardt algorithm," *IEEE Transactions on Power Apparatus and Systems*, no. 2, pp. 695-702, Mar. 1980.
- [7] J. -K. Kim, K. An, J. Ma *et al.*, "Fast and reliable estimation of composite load model parameters using analytical similarity of parameter sensitivity," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp.

- 663-671, Mar. 2015.
- [8] S. Son, S. H. Lee, H. Choi *et al.*, "Improvement of composite load modeling based on parameter sensitivity and dependency analyses," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 242-250, Sept. 2013.
 - [9] B. Tan, J. Zhao, and N. Duan, "Amortized Bayesian parameter estimation approach for WECC composite load model," *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 1517-1529, Feb. 2023.
 - [10] J. Ma, D. Han, R. He *et al.*, "Reducing identified parameters of measurement-based composite load model," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 76-83, Feb. 2008.
 - [11] J.-K. Kim and K. Hur, "On the perturbation size of the finite difference method for trajectory sensitivity-based assessment of power system dynamics with non-smooth behavior," *International Journal of Electrical Power & Energy Systems*, vol. 155, p. 109647, Jan. 2024.
 - [12] S. R. Khazeinyasab, J. Zhao, and N. Duan, "WECC composite load model parameter identification using deep learning approach," in *Proceedings of 2022 IEEE PES General Meeting*, Denver, USA, Jul. 2022, pp. 1-5.
 - [13] F. Bu, Z. Ma, Y. Yuan *et al.*, "WECC composite load model parameter identification using evolutionary deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5407-5417, Nov. 2020.
 - [14] X. Wang, Y. Wang, D. Shi *et al.*, "Two-stage WECC composite load modeling: a double deep Q-learning networks approach," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4331-4344, Sept. 2020.
 - [15] P. Regulski, D. S. Vilchis-Rodriguez, S. Djurović *et al.*, "Estimation of composite load model parameters using an improved particle swarm optimization method," *IEEE Transactions on Power Delivery*, vol. 30, no. 2, pp. 553-560, Feb. 2014.
 - [16] J. Ma, R. He, Z. Y. Dong *et al.*, "Measurement-based load modeling using genetic algorithms," in *Proceedings of 2007 IEEE Congress on Evolutionary Computation*, Singapore, Sept. 2007, pp. 2909-2916.
 - [17] F. A. Hashim and A. G. Hussien, "Snake optimizer: a novel meta-heuristic optimization algorithm," *Knowledge-based Systems*, vol. 242, p. 108320, Apr. 2022.
 - [18] L. Yao, P. Yuan, C.-Y. Tsai *et al.*, "ESO: an enhanced snake optimizer for real-world engineering problems," *Expert Systems with Applications*, vol. 2023, p. 120594, Nov. 2023.
 - [19] N. Hansen, S. D. Müller, and P. Koumoutsakos, "Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES)," *Evolutionary Computation*, vol. 11, no. 1, pp. 1-18, Mar. 2003.
 - [20] K. Zhang, H. Zhu, and S. Guo, "Dependency analysis and improved parameter estimation for dynamic composite load modeling," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3287-3297, Nov. 2016.
 - [21] NERC, "Parameterization of the DER_A model for aggregate DER," North American Electric Reliability Corporation (NERC), Tech. Rep., 2023.
 - [22] P. Pourbeik, J. J. Sanchez-Gasca, J. Senthil *et al.*, "Generic dynamic models for modeling wind power plants and other renewable technologies in large-scale power system studies," *IEEE Transactions on Energy Conversion*, vol. 32, no. 3, pp. 1108-1116, Dec. 2016.
 - [23] S. Guo and T. J. Overbye, "Parameter estimation of a complex load model using phasor measurements," in *Proceedings of 2012 IEEE Power and Energy Conference at Illinois*, Champaign, USA, Feb. 2012, pp. 1-6.
 - [24] S. A. Foroutan and A. Srivastava, "Generator model validation and calibration using synchrophasor data," in *Proceedings of 2019 IEEE Industry Applications Society Annual Meeting*, Baltimore, USA, Sept. 2019, pp. 1-6.
 - [25] A. Chen, Z. Ren, W. Guo *et al.*, "An efficient adaptive differential grouping algorithm for large-scale black-box optimization," *IEEE Transactions on Evolutionary Computation*, vol. 27, no. 3, pp. 475-489, Jun. 2022.
 - [26] Y. Sun, M. Kirley, and S. K. Halgamuge, "A recursive decomposition method for large scale continuous optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 5, pp. 647-661, Oct. 2017.
 - [27] A. Srivastava, J. Zhao, H. Zhu *et al.*, "Distribution system behind-the-meter DERs: estimation, uncertainty quantification, and control," *IEEE Transactions on Power Systems*, vol. 40, no. 1, pp. 1060-1077, Jan. 2024.
 - [28] NERC, "Dynamic load modeling," North American Electric Reliability Corporation (NERC), Tech. Rep., 2016.
 - [29] NERC, "Parameterization of the DER A model," North American Electric Reliability Corporation (NERC), Tech. Rep., 2019.
 - [30] S. Li, S. Lewandowsky, and V. E. DeBrunner, "Using parameter sensitivity and interdependence to predict model scope and falsifiability," *Journal of Experimental Psychology: General*, vol. 125, no. 4, p. 360, Dec. 1996.
- Min-Seung Ko** received the B.S. and Ph.D. degrees in electrical engineering from Yonsei University, Seoul, Republic of Korea, in 2018 and 2024, respectively. He is currently a Postdoctoral Fellow of electrical and computer engineering at The University of Texas at Austin, Austin, USA. In 2024, he was a Senior Researcher with the Net-zero Laboratory, Gwangmyung, Republic of Korea. His current research interests include data-driven planning and operation of inverter-based resource (IBR)-dominant power system, stochastic stability analysis, and seamless integration of artificial intelligence (AI) datacenter.
- Seonghan Kim** received the B.S. degree in electrical engineering from Konkuk University, Seoul, Republic of Korea, in 2019. He is currently working toward the Ph.D. degree in electrical engineering from Yonsei University, Seoul, Republic of Korea. His research interests include power system dynamics, and power system modeling and stability analysis.
- Jae-Kyeong Kim** received the B.S. and Ph.D. degrees in electrical engineering from Yonsei University, Seoul, Republic of Korea, in 2012 and 2018, respectively. In 2018, he was a Postdoctoral Researcher at Yonsei University. Since December 2018, he has been with Korea Electrotechnology Research Institute (KERI), Uiwang, Republic of Korea. His research interests include power system dynamics, power system modeling and stability analysis, power system model validation, data-driven analysis, and application of deep learning.
- Taesik Nam** received the B.S. and Ph.D. degrees in electrical engineering from Yonsei University, Seoul, South Korea, in 2008 and 2015, respectively. He was an R&D engineer with LS Electric, Anyang, Republic of Korea, between 2015 and 2019. Since September 2019, he has been with the Korea Institute of Industrial Technology (KITECH), Gwangju, Republic of Korea. His research interests include grid-forming inverter, integration of variable renewable energy source, distribution grid optimization, power system operation, international standards for power system, and application of machine learning and deep learning in power system.
- Hao Zhu** received the B.S. degree from Tsinghua University, Beijing, China, in 2006, and the M.Sc. and Ph.D. degrees from the University of Minnesota, Minneapolis, USA, in 2009 and 2012, respectively. She is an Associate Professor of electrical and computer engineering at The University of Texas at Austin, Austin, USA. Her research interests include machine learning for power system operation and resilience enhancement, and nexus between power grid and new cyber/computing requirements such as generative artificial intelligence (GenAI) workload.
- Kyeon Hur** received the B.S. and M.S. degrees in electrical engineering from Yonsei University, Seoul, Republic of Korea, in 1996 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, Austin, USA, in 2007. He has rejoined Yonsei University since 2010 and leads a smart-grid research group. His research interests include dynamic performance modeling and analysis of power grid with high level of stochastic and power electronic-interfaced resources, and high-voltage direct current (HVDC) and flexible AC transmission system (FACTS) control.