Analytical Representation and Monetization of Distribution System Flexibility for TSO-DSO Coordination via Advanced Sampling Techniques

Burak Dindar, *Graduate Student Member, IEEE,* Can Berk Saner, *Member, IEEE,* Hüseyin K. Çakmak, and Veit Hagenmeyer, *Member, IEEE*

***Abstract*—As distribution system (DS) flexibility becomes crucial for transmission system operator (TSO) network management, data privacy concerns hinder effective interoperability. The notion of feasible operating region (FOR) has emerged as a promising privacy-preserving concept. However, effectively leveraging FOR in TSO operations remains challenging due to three key factors: its accurate determination in large-scale, meshed DS networks; its tractable analytical representation; and its economic valuation. The present paper proposes an AC optimal power flow (OPF)-based method to construct a three-dimensional PQV-FOR that explicitly accounts for voltage variability and diverse flexibility-providing unit (FPU) characteristics. The construction process employs a two-stage sampling strategy that combines bounding box projection and Fibonacci direction sampling to capture the FOR. Then, an implicit polynomial fitting approach is introduced to analytically represent the FOR. Furthermore, a quadratic cost function is derived over the PQV domain to monetize the FOR. The resulting framework enables single-round TSO-DSO communication without iterative exchanges. Case studies on meshed DS with up to 533 buses, integrated into TSs, demonstrate the method's efficiency compared to standard AC-OPF. On average, the proposed approach yields negligible cost deviations of at most 0.058% across test cases, while reducing online computation times by up to 58.11%.**

***Index Terms*—aggregated flexibility, analytical representation, feasible operating region, monetization, TSO-DSO coordination.**

I. Introduction

The rapid integration of distributed generators (DGs) has significantly increased the flexibility in distribution systems (DSs). This flexibility, delivered by flexibility-providing units (FPUs), offers substantial potential for overall network management. Therefore, effective coordination between Transmission System Operators (TSOs) and Distribution System Operators (DSOs) has become essential for reliable grid operation [1]. However, such coordination is often constrained by data privacy concerns among stakeholders [2], [3], motivating the development of privacy-preserving coordination frameworks.

In this context, the Feasible Operating Region (FOR), based on the PQ (active and reactive power) chart, has been proposed to represent aggregated DS flexibility. The FOR captures all feasible combinations of active and reactive power at the point of common coupling (PCC) while respecting DS technical constraints. By sharing this compact representation instead of sensitive DSO data (e.g., network topology or customer information), DSOs enable TSOs to directly incorporate DS flexibility into their operational planning [4].

A. Related Work

Literature classifies FOR computation into geometric, random sampling (RS), and optimization-based approaches, particularly those relying on optimal power flow (OPF). Geometric methods aggregate FPU flexibilities to form the FOR [5], [6]. However, they may not explicitly account for network constraints, as noted in [7]. RS approaches generate a large number of operating points and verify feasibility through power flow analysis [8]. While grid-aware, these methods are computationally intensive [9]. As a result, OPF-based methods have gained prominence as a more efficient alternative [10].

Among OPF-based approaches, linearized models are commonly used [11], [12], offering computational efficiency but sacrificing accuracy due to linearization of power flow equations [10]. Another category involves sampling strategies [13], typically classified as angle-based and set-point-based approaches, in which OPF is solved for varying objective function coefficients in the PQ domain [14], [15]. These methods are effective in identifying the FOR boundary [7], but most existing studies are limited to radial networks, small-scale systems, and fixed PCC voltage assumptions [4], [16]. In practice, DSs are increasingly operated as meshed networks and this should be considered. Moreover, to fully exploit DS flexibility, the PCC voltage should not be assumed constant; instead, analyses should be conducted in the three-dimensional PQV (active power, reactive power, and voltage) domain [17]. These factors highlight the need for more advanced approaches to accurately determine the FOR under realistic conditions.

Once the FOR is determined, a critical question arises: *how can the FORs computed by DSOs be effectively utilized by TSOs in their operational planning?* Iterative coordination schemes have been proposed, in which the DSO computes the FOR and the TSO proposes an operating point that is subsequently validated by the DSO [18]–[20]. However, such approaches often require multiple iterations, leading to high communication overhead. To avoid this, it is desirable to

This work was conducted within the framework of the Helmholtz Program Energy System Design (ESD).

Burak Dindar, Hüseyin K. Çakmak and Veit Hagenmeyer are with the Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany, (e-mail:burak.dindar@kit.edu; hueseyin.cakmak@kit.edu; veit.hagenmeyer@kit.edu).

Can Berk Saner is with the Department of Mathematics, National University of Singapore, Singapore 119076, (e-mail: sanerc@u.nus.edu).

TABLE I
OVERVIEW OF THE RELATED WORK.

Work	Uses Full AC-Power Flow Model	Meshed Distribution System	Considered Variables at PCC	Analytical FOR	Analytical Cost	TSO-DSO Coordination Scheme
[18]	✗	✗	P, Q	✓(Polygon)	✗	Iterative
[19]	✓	✓	P, Q	✗	✗	Iterative
[20]	✗	✗	P, Q	✓(Polygon)	✗	Iterative
[21]	✗	✗	P, Q	✓(Polygon)	✗	Iterative
[22]	✓	Not Specified	P, Q	✓(PQ Box)	✗	Single Round
[23]	✗	Not Specified	P, Q	✓(PQ Box)	✗	Iterative
[24]	✗	✗	P, Q, V	✓(Polyhedron)	✓	Not Specified
[25]	✓	✗	P, Q	✗	✗	Not Specified
[26]	✓	✓	P, Q	✗	✗	Not Specified
[27]	✓	✗	P or Q	✗	✓	Not Specified
[28]	✓	Not Specified	Q	✓(1D Q-Range)	✓	Single Round
This Work	✓	✓	P, Q, V	✓(Implicit Polynomial)	✓	Single Round

represent the FOR analytically so that it can be directly embedded into the TSO's OPF formulation. Due to the potentially non-convex nature of the FOR, this remains challenging, and simplifying assumptions are often employed in the literature.

For instance, [21] approximates the FOR as a polygon and integrates it into a linear OPF through linear inequalities. Similarly, [22] and [23] model the FOR using a rectangular PQ box. While computationally convenient, these simplifications neglect the complex geometry of the true FOR. The work presented in [24] employs the LinDistFlow model to compute a conservative approximation of the FOR, ensuring that the resulting bounds remain within the actual FOR, thus avoiding the risk of selecting infeasible setpoints, which is a common issue in convex hull-based approaches. As a result, existing methods often lead to over-conservative or inaccurate representations, highlighting the need for analytical FOR models that balance computational tractability and accuracy.

While the FOR characterizes feasibility, its integration into the TSO's OPF also requires an analytical representation of flexibility cost in the PQV domain. One of the earliest efforts is presented in [25], where monetization is approached through aggregation and disaggregation, and the cost is computed individually for each FOR operating point. Similarly, [26] follows a comparable approach. However, these methods lack an analytical cost representation, require repeated evaluations for new points, and therefore remain computationally intensive.

An analytical cost representation is proposed in [27], where piecewise linear functions are derived for active and reactive power, but the approach is validated only on a 34-bus radial system and not integrated into OPF. In [28], a quadratic cost is used for reactive power under fixed active power and voltage assumptions. Meanwhile, [24] emphasizes that the variables at the PCC, including P, Q, and V, must be considered holistically when defining the cost function; however, the cost is fitted using a quadratic function based on the simplified LinDistFlow model and is validated only on a small 15-bus radial test system. These studies highlight that, to fully leverage DS flexibility within TSO operations, the cost of the FOR must be represented in a form that is both analytically tractable and compatible with OPF formulations. Furthermore, these approaches must be extended to more realistic conditions, including meshed network topologies, large-scale systems, and AC-OPF. An overview of the related work is given in Table I.

B. Contributions

In the present paper, we address a key limitation of most existing flexibility models, which are defined in the PQ domain and assume a fixed PCC voltage. Since voltage is a controllable variable in practice, this assumption underestimates the available flexibility. Therefore, we construct the FOR in the three-dimensional PQV space, thereby capturing the broader DS flexibility. To efficiently capture the FOR boundary, we use bounding box projection and Fibonacci direction sampling techniques, enabling accurate estimation with a relatively small number of samples. Additionally, we incorporate diverse PQ characteristics, moving beyond the conventional assumption of rectangular PQ capability profiles.

Furthermore, while several studies provide geometric analytical FOR approximations, these are typically limited to simple shapes. In contrast, we represent the FOR using an implicit polynomial, which allows complex non-convex geometries to be expressed as a single constraint in the AC-OPF. Importantly, the polynomial is constructed conservatively to avoid including infeasible regions while still preserving as much usable flexibility as possible, thereby promoting feasible OPF solutions without significant loss of available flexibility.

Finally, a quadratic cost function is fitted over the PQV domain to complement the FOR representation and enable economic optimization of flexibility within the same OPF framework, unlike existing approaches. Comprehensive case studies are then conducted to demonstrate the effectiveness of the proposed method.

The key contributions of the present paper are as follows:

- A realistic characterization of DS flexibility is proposed by moving beyond conventional PQ-only models and explicitly defining flexibility in the three-dimensional PQV domain, capturing the impact of voltage variability.
- An analytical representation of DS flexibility is formulated to enable its direct integration into the TSO's AC-OPF, comprising two components: (i) an implicit polynomial to represent the non-convex FOR boundary, and (ii) an analytical quadratic cost function to monetize operating points, completing the analytical interface.
- A data-driven methodology is developed to construct analytical models through effective sampling strategies. A complementary approach, combining bounding box

projection and Fibonacci directions, efficiently captures the FOR boundary, while Latin hypercube sampling generates a homogeneous dataset for the cost function.
- The proposed analytical representations facilitate the realization of a single-round TSO-DSO communication scheme, without iterative coordination or computationally intensive disaggregation. The resulting interface is inherently privacy-preserving, as only analytical functions defined over non-sensitive coupling variables are exchanged, without requiring disclosure of sensitive data such as network topology and load profiles.

The rest of the paper is organized as follows: Section II outlines the proposed methodology and introduces the FOR-based AC-OPF. Building upon this formulation, Section III describes the sampling strategies used to extract representative boundary and interior points of the FOR. These samples form the basis of Section IV, where analytical models of the FOR and its associated cost function are constructed using implicit polynomials. Subsequently, Section V validates the proposed framework through case studies, including performance evaluation, sensitivity analyses, and comparisons with benchmark methods. Finally, Section VI concludes the paper.

II. Overview of the Proposed Methodology

Throughout this paper, we adopt the notation and symbol conventions summarized in Table II.

A. Standard AC-OPF for Integrated Transmission-Distribution System

For an integrated transmission-distribution system with n_{b} total buses, including n_{g} conventional generator and $n_{\mathrm{b,ts}}$ buses in TS, and comprising n_{ds} DSs, where the j-th DS contains $n_{\mathrm{dg},j}$ DGs[1], the standard AC-OPF problem, without any modifications to the constraints or objective function, can be formulated as follows:

$$\min_{\substack{\widehat{\boldsymbol{v}},\widehat{\boldsymbol{\theta}},\\ \breve{\boldsymbol{p}}_{\mathrm{g}},\breve{\boldsymbol{q}}_{\mathrm{g}},\\ \boldsymbol{p}_{\mathrm{dg},j},\\ \boldsymbol{q}_{\mathrm{dg},j}}} \quad \sum_{i=1}^{n_{\mathrm{g}}} C_i(\breve{\boldsymbol{p}}_{\mathrm{g}}^{(i)}) + \sum_{j=1}^{n_{\mathrm{ds}}} \sum_{k=1}^{n_{\mathrm{dg},j}} C_{jk}(\boldsymbol{p}_{\mathrm{dg},j}^{(k)}) \tag{1a}$$

$$\text{s.t.} \quad G_{\mathrm{P}}(\widehat{\boldsymbol{v}},\widehat{\boldsymbol{\theta}};\widehat{Y}) + \widehat{p}_{\mathrm{d}} - K\breve{\boldsymbol{p}}_{\mathrm{g}} - \sum_{j=1}^{n_{\mathrm{ds}}} H_j \boldsymbol{p}_{\mathrm{dg},j} = 0, \tag{1b}$$

$$G_{\mathrm{Q}}(\widehat{\boldsymbol{v}},\widehat{\boldsymbol{\theta}};\widehat{Y}) + \widehat{q}_{\mathrm{d}} - K\breve{\boldsymbol{q}}_{\mathrm{g}} - \sum_{j=1}^{n_{\mathrm{ds}}} H_j \boldsymbol{q}_{\mathrm{dg},j} = 0, \tag{1c}$$

$$G_{\mathrm{line}}(\widehat{\boldsymbol{v}},\widehat{\boldsymbol{\theta}};\widehat{Y}) \preceq \widehat{l}_{\mathrm{line,max}}, \tag{1d}$$

$$G_{\mathrm{dg},j}(\boldsymbol{p}_{\mathrm{dg},j},\boldsymbol{q}_{\mathrm{dg},j}) \preceq 0, \ \forall j \in \{1,..,n_{\mathrm{ds}}\}, \tag{1e}$$

$$\widehat{v}_{\min} \preceq \widehat{\boldsymbol{v}} \preceq \widehat{v}_{\max}, \ \widehat{\theta}_{\min} \preceq \widehat{\boldsymbol{\theta}} \preceq \widehat{\theta}_{\max}, \tag{1f}$$

$$\breve{p}_{\mathrm{g,min}} \preceq \breve{\boldsymbol{p}}_{\mathrm{g}} \preceq \breve{p}_{\mathrm{g,max}}, \ \breve{q}_{\mathrm{g,min}} \preceq \breve{\boldsymbol{q}}_{\mathrm{g}} \preceq \breve{q}_{\mathrm{g,max}}, \tag{1g}$$

$$p_{\mathrm{dg},j,\min} \preceq \boldsymbol{p}_{\mathrm{dg},j} \preceq p_{\mathrm{dg},j,\max}, \ \forall j \in \{1,.,n_{\mathrm{ds}}\}, \tag{1h}$$

$$q_{\mathrm{dg},j,\min} \preceq \boldsymbol{q}_{\mathrm{dg},j} \preceq q_{\mathrm{dg},j,\max}, \ \forall j \in \{1,..,n_{\mathrm{ds}}\}, \tag{1i}$$

[1]The presented framework is applicable to any type of FPUs. For ease of exposition, we refer to these units simply as DGs throughout the remainder of the paper.

TABLE II
Notation and Symbol Conventions

Symbol	Description
a, A	Parameters
$\boldsymbol{a}, \boldsymbol{A}$	Variables
a	Scalar and (column) vector
A	Matrix
$\mathcal{A}$	Set
$A(\cdot)$	Function
$a^{(n)}$	n-th element of vector a
$A^{(n,:)}$	n-th row of matrix A
$A^{(i,j)}$	Element in row i, column j of A
$\preceq, \succeq$	Element-wise comparison
$\leqslant, \geqslant$	Scalar comparison
$\widehat{\boldsymbol{a}}$	Integrated TS-DS variable
$\breve{\boldsymbol{a}}$	TS-only variable
$\boldsymbol{a}$	DS-only variable

where $\widehat{\boldsymbol{v}}, \widehat{\boldsymbol{\theta}}, \widehat{p}_{\mathrm{d}}, \widehat{q}_{\mathrm{d}} \in \mathbb{R}^{n_{\mathrm{b}}}$ denote the bus voltage magnitudes, voltage angles, and active and reactive power demand vectors, respectively. The bus admittance matrix is represented by $\widehat{Y} \in \mathbb{R}^{n_{\mathrm{b}} \times n_{\mathrm{b}}}$. Active and reactive power generation vectors for the TS are $\breve{\boldsymbol{p}}_{\mathrm{g}}, \breve{\boldsymbol{q}}_{\mathrm{g}} \in \mathbb{R}^{n_{\mathrm{b,ts}}}$, with the corresponding connection matrix $K \in \mathbb{R}^{n_{\mathrm{b}} \times n_{\mathrm{b,ts}}}$, where $K^{(t,v)} = 1$ if this element is in the TS, and zero otherwise. For the j-th DS, $\boldsymbol{p}_{\mathrm{dg},j}, \boldsymbol{q}_{\mathrm{dg},j} \in \mathbb{R}^{n_{\mathrm{dg},j}}$ denote the active and reactive power generation vectors of DGs. Their connection to the network is captured by matrix $H_j \in \mathbb{R}^{n_{\mathrm{b}} \times n_{\mathrm{dg},j}}$ with $H_j^{(m,n)} = 1$ if the n-th DG of DS j connects at bus m, and zero otherwise.

The OPF objective in (1a) minimizes the total generation cost, including that of DGs. Let $C_i(\cdot)$ and $C_{jk}(\cdot)$ be the generation cost functions for TS generators at bus i, and the k-th DG in DS j, respectively. Both functions are modeled as standard quadratic cost functions as $C_l(\boldsymbol{p}) = a_l \boldsymbol{p}^2 + b_l \boldsymbol{p} + c_l$, without loss of generality. For notational simplicity, the first n_{g} buses are correspond to conventional generators in the integrated system. Equations (1b)-(1c) enforce active and reactive power balance through functions $G_P(\cdot)$ and $G_Q(\cdot)$. The line flow limits are enforced in (1d) via $G_{line}(\cdot)$, bounded by line flow limit vector $\widehat{l}_{\mathrm{line,max}}$. Additionally, the function $G_{\mathrm{dg},j}(\cdot)$ in (1e) characterizes the operating limits of DGs that do not exhibit conventional rectangular PQ capability curves. Constraints on voltages and generations are imposed in (1f)-(1i).

B. FOR-Based AC-OPF for Privacy-Preserving DS Flexibility Utilization

A key observation from (1) is that utilizing DS flexibility within an OPF framework typically necessitates access to sensitive system information. For instance, the admittance matrix $\widehat{Y}$ encodes detailed grid topology, while the vectors $\widehat{p}_{\mathrm{d}}$ and $\widehat{q}_{\mathrm{d}}$ contain detailed load profiles. As the OPF problem is typically coordinated by the TSO, DSOs are often reluctant to disclose such data due to privacy concerns. To overcome this limitation, we propose a FOR-based AC-OPF formulation that enables the integration of DS flexibility without requiring the exchange of sensitive information between TSOs and DSOs.

In the proposed framework, each DS is represented via two analytical functions: one representing its FOR and the

other modeling the associated cost of FOR. These functions are defined over non-sensitive coupling variables that are already exchanged between TSOs and DSOs. Specifically, we consider n_{ds} DSs, where the j-th DS comprises $n_{\mathrm{dg},j}$ DGs, and is connected to a designated transmission buses tb_j, which serve as points of common coupling (PCCs). These PCCs are assumed to be *empty buses*, meaning they do not host any directly connected generation or load units. Each DS j is therefore characterized by functions defined over the coupling variables in the PQV domain, $\boldsymbol{p}_j, \boldsymbol{q}_j, \boldsymbol{v}_j \in \mathbb{R}$, namely active and reactive power flow at the PCC-directed from DS towards TS-and voltage magnitude at the corresponding PCC. For notational convenience, we define a concatenated vector of coupling variables as $\boldsymbol{x}_j = \begin{bmatrix}\boldsymbol{p}_j & \boldsymbol{q}_j & \boldsymbol{v}_j\end{bmatrix}^\top \in \mathbb{R}^3$. With this, the FOR-based AC-OPF problem can be expressed as follows:

$$\min_{\substack{\breve{\boldsymbol{v}},\breve{\boldsymbol{\theta}},\\ \breve{\boldsymbol{p}}_{\mathrm{g}},\breve{\boldsymbol{q}}_{\mathrm{g}},\\ \boldsymbol{x}_j}} \quad \sum_{i=1}^{n_{\mathrm{g}}} C_i(\breve{\boldsymbol{p}}_{\mathrm{g}}^{(i)}) + \sum_{j=1}^{n_{\mathrm{ds}}} C_j(\boldsymbol{x}_j) \tag{2a}$$

$$\text{s.t.} \quad G_{\mathrm{P}}(\breve{\boldsymbol{v}}, \breve{\boldsymbol{\theta}}; \breve{Y}) + \breve{p}_{\mathrm{d}} - \breve{\boldsymbol{p}}_{\mathrm{g}} = 0, \tag{2b}$$

$$G_{\mathrm{Q}}(\breve{\boldsymbol{v}}, \breve{\boldsymbol{\theta}}; \breve{Y}) + \breve{q}_{\mathrm{d}} - \breve{\boldsymbol{q}}_{\mathrm{g}} = 0, \tag{2c}$$

$$G_{\mathrm{line}}(\breve{\boldsymbol{v}}, \breve{\boldsymbol{\theta}}; \breve{Y}) \preceq \breve{l}_{\mathrm{line,max}}, \tag{2d}$$

$$\breve{v}_{\min} \preceq \breve{\boldsymbol{v}} \preceq \breve{v}_{\max},\ \breve{\theta}_{\min} \preceq \breve{\boldsymbol{\theta}} \preceq \breve{\theta}_{\max}, \tag{2e}$$

$$\breve{p}_{\mathrm{g,min}} \preceq \breve{\boldsymbol{p}}_{\mathrm{g}} \preceq \breve{p}_{\mathrm{g,max}},\ \breve{q}_{\mathrm{g,min}} \preceq \breve{\boldsymbol{q}}_{\mathrm{g}} \preceq \breve{q}_{\mathrm{g,max}}, \tag{2f}$$

$$FOR_j(\boldsymbol{x}_j) \leqslant 0,\ \forall j \in \{1,..,n_{\mathrm{ds}}\}, \tag{2g}$$

$$\breve{\boldsymbol{p}}_{\mathrm{g}}^{(tb_j)} = \boldsymbol{p}_j,\ \breve{\boldsymbol{q}}_{\mathrm{g}}^{(tb_j)} = \boldsymbol{q}_j,\ \breve{\boldsymbol{v}}^{(tb_j)} = \boldsymbol{v}_j, \tag{2h}$$

$$x_{j,\min} \preceq \boldsymbol{x}_j \preceq x_{j,\max},\ \forall j \in \{1,..,n_{\mathrm{ds}}\}, \tag{2i}$$

$$\boldsymbol{x}_j = \begin{bmatrix}\boldsymbol{p}_j\ \boldsymbol{q}_j\ \boldsymbol{v}_j\end{bmatrix}^\top, \forall j \in \{1,..,n_{\mathrm{ds}}\}. \tag{2j}$$

Examining (2b) - (2f), it is evident that only TS-related variables are explicitly included, while DS-related variables are encapsulated within the functions $FOR_j(\boldsymbol{x}_j)$, as defined in (2g). $FOR_j(\boldsymbol{x}_j)$ characterize the FOR of the DSs, ensuring compliance with internal technical constraints such as voltage and line flow limits. Specifically, $FOR_j(\boldsymbol{x}_j) \leqslant 0$ holds if and only if $\boldsymbol{x}_j$ lies within the FOR. Otherwise, it indicates a violation of DS constraints. Additionally, the function $C_j(\boldsymbol{x}_j)$ represent the cost associated with the FOR. Furthermore, (2h) enforces the physical coupling between TS and DS by ensuring that the coupling variables match at the PCCs. Lastly, (2i) defines the bounds on the coupling variables.

Both $FOR_j(\cdot)$ and $C_j(\cdot)$ must be represented in analytical form to ensure their tractability within the AC-OPF framework. Crucially, these functions are constructed solely using non-sensitive coupling variables, thereby preserving data privacy between TSOs and DSOs. A schematic overview of the proposed framework is illustrated in Fig. 1. In this architecture, each DSO computes its respective $FOR_j(\cdot)$ and $C_j(\cdot)$ and communicates them to the TSO. The TSO subsequently incorporates these functions into the FOR-based AC-OPF in (2). Once an optimal solution within the feasible region is obtained, each DSO independently solves its local AC-OPF to determine the internal dispatch of its DGs. This process eliminates the need for computationally intensive disaggregation or iterative coordination. As a result, the proposed method enables cost-effective integration of DS flexibility into system-level decision-making, while maintaining data confidentiality and complying with the technical constraints of both TSs and DSs. However, the practical implementation of this framework relies on the availability of high-quality data that characterize the FOR. Consequently, before these analytical functions can be constructed, a systematic data generation process is required, as detailed in the following section.

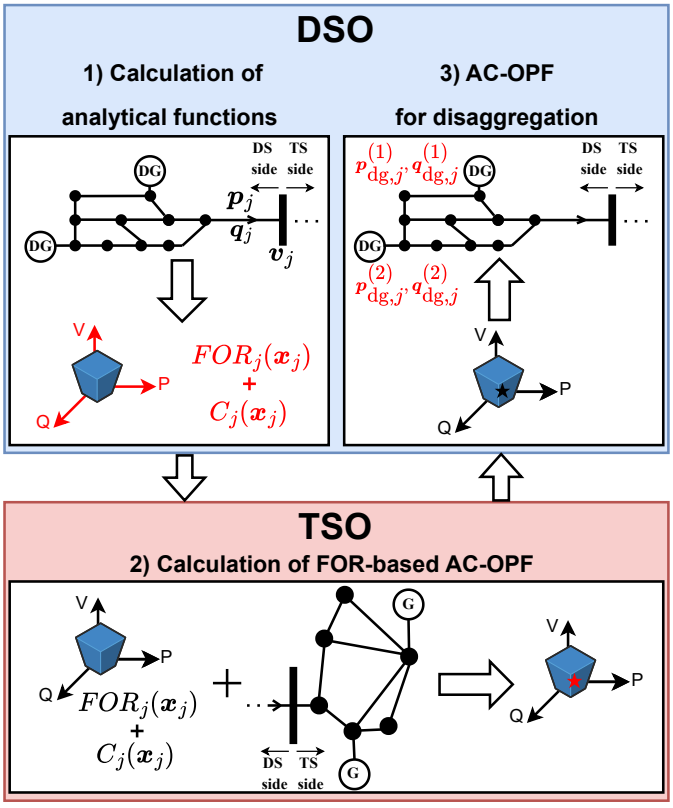

Fig. 1. Schematic representation of the proposed method. For clarity, only a single DS is illustrated; however, the framework supports multiple DSs. The results obtained at each step are highlighted in red.

III. Sampling Strategies for Flexibility Modeling

Based on the architectural requirements established in Section II, constructing the analytical functions $FOR_j(\cdot)$ and $C_j(\cdot)$ requires two distinct datasets because they serve fundamentally different modeling purposes. The FOR representation aims to accurately capture the boundary of the feasible region, which requires dense sampling on the surface of the region. In contrast, the cost functions $C_j(\cdot)$ aims to approximate the cost distribution within the interior of the feasible region, which requires a homogeneous sampling of interior points. Therefore, we employ efficient and computationally tractable strategies to generate both datasets with minimal overhead, as detailed in the following subsections.

It is important to note that the sampling is not performed by independently prescribing active power, reactive power, and voltage at the PCC. Instead, all PQV samples are obtained as feasible solutions of AC-OPF problems, in which the physical

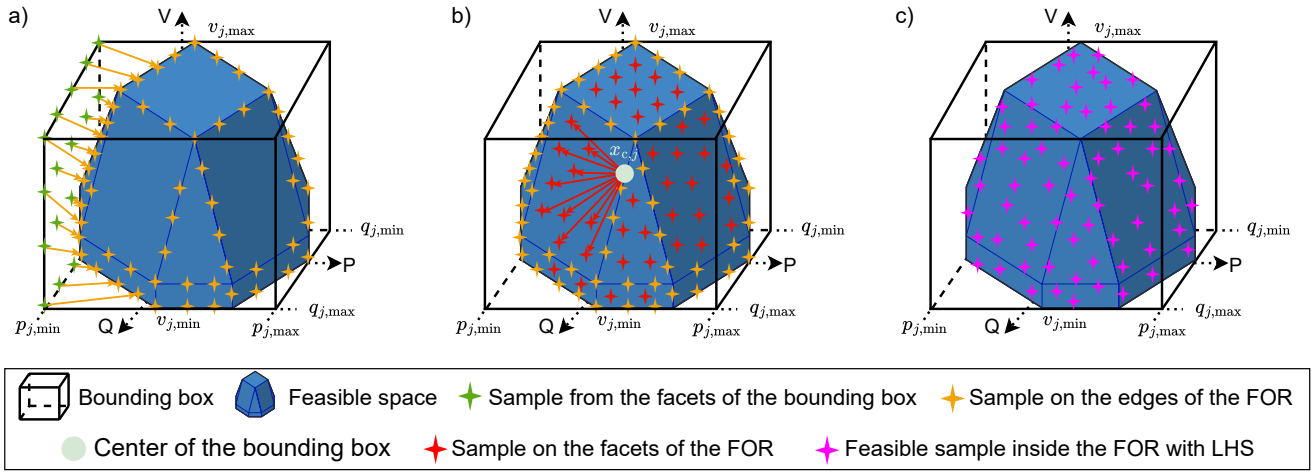

Fig. 2. Illustration of the sampling procedure. a) BBPS: Firstly, a tight bounding box is established to enclose the FOR. Initial samples (green) are then taken from its facets and projected onto the FOR boundary (amber) by solving (3). This strategy effectively concentrates samples along the edges of the FOR (see Algorithm 1). b) FDS: An origin point ($x_{c,j}$) is defined at the center of the bounding box, from which virtual arrows are cast toward the boundary along systematically chosen directions using a Fibonacci lattice. By solving (4), boundary samples (red) are obtained across the facets of the FOR, ensuring comprehensive and uniform surface coverage (see Algorithm 2). c) LHS: Candidate points are generated inside the bounding box using LHS. By solving (5), only the feasible samples (magenta) are retained within the FOR. This ensures a homogeneous dataset that accurately represents the interior region, providing a reliable basis for approximating the cost distribution (see Algorithm 3). For clarity in the illustration, arrows are shown for only one side.

interdependence among these variables is fully enforced by the AC power flow equations and operational constraints. Accordingly, the constructed PQV datasets represent physically achievable PCC operating points rather than arbitrary combinations of P, Q, and V.

A. Boundary Sampling for the PQV-FOR

As highlighted in the literature (see Section I-A), AC-OPF-based methods are particularly effective for identifying the FOR boundary. However, capturing the FOR's complete geometry, which includes both the edges and facets, requires a comprehensive strategy. To this end, we develop an approach for generating FOR boundary data that combines two complementary methods: *Bounding Box Projection Sampling (BBPS)* to effectively capture the edges, and *Fibonacci Direction Sampling (FDS)* to ensure full coverage of the facets. The combination of these methods, which are detailed in the following subsections, yields a well-distributed and representative dataset that accurately characterizes the FOR's complex shape.

1) Bounding Box Projection Sampling (BBPS): The BBPS method, detailed in Algorithm 1, is the first of two complementary strategies. The process is executed in two phases.

In the first phase, we establish a tight bounding box that encloses the FOR. This is achieved by solving six AC-OPF problems to find the minimum and maximum feasible values for each of the three coupling variables ($\boldsymbol{p}_j, \boldsymbol{q}_j, \boldsymbol{v}_j$) at the PCC, resulting in the limit vectors $x_{j,\min}, x_{j,\max} \in \mathbb{R}^3$. This targeted approach avoids inefficient sampling of the inherently large infeasible space and focuses the data generation effort on the most relevant region.

In the second phase, we generate the boundary dataset itself. A set of $n_{\text{bbps},j}$ points, denoted individually as $x_{\text{lhsf},j} \in \mathbb{R}^3$, is sampled on the facets of the previously determined bounding box using Latin Hypercube Sampling (LHS) [29]. Each external point $x_{\text{lhsf},j}$ is then projected onto the FOR boundary by solving the optimization problem defined in (3). This problem finds the closest feasible point $x_j^* \in \mathbb{R}^3$ on the FOR by minimizing the Euclidean (L_2) distance, an approach similar to the one introduced in our previous work [30].

The resulting optimal points x_j^* are collected to form the dataset $D_{\text{bbps},j} \in \mathbb{R}^{n_{\text{bbps},j} \times 3}$. The points generated via BBPS tend to concentrate along the sharp edges of the FOR (see Fig. 2a), effectively capturing these critical features.

Algorithm 1 The BBPS-based Algorithm for Generating FOR Boundary Data

Input: Power system data for DS j
Output: $x_{j,\min}, x_{j,\max}$, $D_{\text{bbps},j}$

Phase 1: Determine Bounding Box
for $idx \leftarrow 1$ to 3 **do**
  $x_{j,\min}^{(idx)} \leftarrow$ Solve $\min \boldsymbol{x}_j^{(idx)}$ s.t. AC-OPF constraints
  $x_{j,\max}^{(idx)} \leftarrow$ Solve $\max \boldsymbol{x}_j^{(idx)}$ s.t. AC-OPF constraints
end for
Phase 2: Project Samples onto FOR Boundary
$D_{\text{bbps},j} \leftarrow [\,]$;
for $k \leftarrow 1$ to $n_{\text{bbps},j}$ **do**
  $x_{\text{lhsf},j} \leftarrow$ sample a vector from the facets of the bounding box defined by $x_{j,\min}$ and $x_{j,\max}$ using LHS
  Define optimization problem:

$$\min_{\boldsymbol{x}_j} \ \|x_{\text{lhsf},j} - \boldsymbol{x}_j\|_2^2 \tag{3a}$$

  s.t. Standard AC-OPF constraints for DS j

  Solve (3) and obtain optimal point x_j^*
  $D_{\text{bbps},j} \leftarrow [D_{\text{bbps},j}; x_j^{*\top}]$;
end for

2) Fibonacci Direction Sampling (FDS): While the BBPS method effectively captures the edges of the FOR, a complementary approach is needed to sample its facets. To this end, we introduce the FDS method, detailed in Algorithm 2, which ensures comprehensive coverage of the FOR's entire surface (see Fig. 2b).

Algorithm 2 The FDS-based Algorithm for Generating FOR Boundary Data

Input: Power system data for DS j, $x_{j,\min}, x_{j,\max}$
Output: $D_{\text{fds},j}$

Define an optimization problem for a given direction d_k:

$$\max_{\boldsymbol{t}_k} \ \boldsymbol{t}_k \tag{4a}$$

$$\text{s.t. } \boldsymbol{x}_j = x_{\text{c},j} + d_k \boldsymbol{t}_k \tag{4b}$$

Standard AC-OPF constraints for DS j

$D_{\text{fds},j} \leftarrow [\,]$;
$x_{\text{c},j} \leftarrow 0.5 \times (x_{j,\max} + x_{j,\min})$
$\phi \leftarrow \pi(3 - \sqrt{5})$ ▷ Golden angle for Fibonacci lattice
for $k \leftarrow 0$ to $n_{\text{fds},j} - 1$ **do**
  $d_k \leftarrow \left[r\cos(\theta)\ r\sin(\theta)\ z\right]^\top$, where $z = 1 - \frac{2k+1}{n_{\text{fds},j}}$, $\theta = \phi \cdot k$, $r = \sqrt{1 - z^2}$
  Solve (4) to obtain the optimal step size t_k^*
  Calculate the boundary point: $x_j^* \leftarrow x_{\text{c},j} + d_k t_k^*$
  $D_{\text{fds},j} \leftarrow [D_{\text{fds},j}; x_j^{*\top}]$;
end for

The core idea is to cast *virtual arrows* from the FOR's interior to its boundary along systematically chosen directions. The process begins by defining an origin point, $x_{\text{c},j} \in \mathbb{R}^3$, at the center of the bounding box. From this origin, a set of $n_{\text{fds},j}$ direction vectors, $d_k \in \mathbb{R}^3$, is generated using the Fibonacci lattice method, which ensures uniform angular coverage in three-dimensional space. For each direction d_k, we solve the optimization problem in (4) to find the maximum feasible step size, $\boldsymbol{t}_k \in \mathbb{R}$, along that direction. The resulting intersection point with the FOR boundary, $x_j^* \in \mathbb{R}^3$, is then recorded. This procedure is repeated for all directions, yielding the dataset $D_{\text{fds},j} \in \mathbb{R}^{n_{\text{fds},j} \times 3}$.

Finally, to create a complete representation of the FOR's geometry, the datasets from both sampling methods are combined into a single dataset: $D_{\text{FOR},j} = [D_{\text{bbps},j}; D_{\text{fds},j}]$, where, $D_{\text{FOR},j} \in \mathbb{R}^{n_{\text{FOR},j} \times 3}$. For this, $n_{\text{FOR},j}$ is defined as $n_{\text{FOR},j} = n_{\text{bbps},j} + n_{\text{fds},j}$. This final dataset captures both the sharp edges and the smooth facets of the FOR with high fidelity, serving as the foundation for the data-driven construction of the analytical model $FOR(\boldsymbol{x})$.

B. Interior Sampling for the Cost of the FOR

To create the cost function, a dataset of feasible operating points from the interior of the FOR is required. Unlike the boundary-focused methods, this step necessitates a homogeneous sampling of the entire FOR. We achieve this using the LHS method, as illustrated in Fig. 2c and detailed in Alg. 3.

The process begins by generating a candidate operating point, $x_{\text{lhsc},j} \in \mathbb{R}^3$, from within the bounding box using LHS.

Algorithm 3 The LHS-based Algorithm for Generating Cost Data

Input: Power system data for DS j, $x_{j,\min}, x_{j,\max}$
Output: $D_{\text{cost},j}$, $y_{\text{cost},j}$

Define an optimization problem for a candidate point $x_{\text{lhsc},j}$:

$$\min\ C_j = \sum_{k=1}^{n_{\text{dg},j}} C_{jk}(\boldsymbol{p}_{\text{dg},j}^{(k)}) \tag{5a}$$

$$\text{s.t. } \boldsymbol{x}_j = x_{\text{lhsc},j} \tag{5b}$$

Standard AC-OPF constraints for DS j

$D_{\text{cost},j} \leftarrow [\,]$;
$y_{\text{cost},j} \leftarrow [\,]$;
$idx \leftarrow 1$;
while $idx \leqslant n_{\text{cost},j}$ **do**
  $x_{\text{lhsc},j} \leftarrow$ sample a vector from the bounding box using LHS
  if (5) with $x_{\text{lhsc},j}$ is feasible **then**
    Obtain the optimal total cost C_j^*
    $D_{\text{cost},j} \leftarrow [D_{\text{cost},j}; x_{\text{lhsc},j}{}^\top]$;
    $y_{\text{cost},j} \leftarrow [y_{\text{cost},j}; C_j^*]$;
    $idx \leftarrow idx + 1$;
  end if
end while

This point is then tested for feasibility by solving the AC-OPF problem defined in (5), where the PCC variables are fixed to the candidate's values. If feasible, the candidate point is valid, and the point itself is stored in the feature dataset $D_{\text{cost},j} \in \mathbb{R}^{n_{\text{cost},j} \times 3}$, and the corresponding optimal total cost, C_j^*, is stored in the target vector $y_{\text{cost},j} \in \mathbb{R}$. This procedure is repeated until the desired number of samples, $n_{\text{cost},j}$, is collected, resulting in a dataset that homogeneously covers the cost characteristics of the FOR interior. These generated datasets provide the essential numerical foundation for the analytical modeling process. By capturing both the boundary and the interior cost distribution, these discrete samples enable the transition from raw data to the continuous functional representation described in the following section.

IV. Construction of the Analytical Functions

Based on the boundary and interior datasets $D_{\text{FOR},j}$ and $D_{\text{cost},j}$ generated in Section III, we now detail the procedure for constructing the analytical functions $FOR_j(\cdot)$ and $C_j(\cdot)$.

We utilize *polynomial representations* for several practical and theoretical reasons. First, fitting polynomial functions requires short training times and minimal hyperparameter tuning, making it computationally attractive for large-scale power system applications. Second, once fitted, the function is fully characterized by a set of numerical coefficients. This enables a privacy-preserving exchange of flexibility information while ensuring communication is trivial in size and highly data-efficient. Third, polynomials are smooth and differentiable, which makes them well-suited for gradient-based optimization solvers used in AC-OPF. This ensures numerical stability and

seamless integration into the OPF formulation while maintaining computational efficiency. Finally, as detailed in the theoretical insights provided in Section IV-C, the use of polynomials is mathematically substantiated by their strong approximation capabilities over bounded domains. This provides a theoretical foundation for the proposed analytical representations, which are capable of capturing complex, non-convex flexibility characteristics across different network configurations.

A. Implicit Polynomial Representation of the PQV-FOR

For notational brevity, the DS subscript j is omitted in the remainder of this section.

The objective of this subsection is to approximate the boundary of the three-dimensional FOR using a single implicit polynomial function. For this, an implicit polynomial of degree d_{FOR} is used. A point $x = (p, q, v)$ is considered to be on the boundary of the FOR if it satisfies $FOR(x) = 0$. The polynomial is of the form:

$$FOR(\boldsymbol{x}) = \sum_{\alpha+\beta+\theta \leqslant d_{\mathrm{FOR}}} a_{\alpha\beta\theta} \boldsymbol{p}^{\alpha} \boldsymbol{q}^{\beta} \boldsymbol{v}^{\theta}. \tag{6}$$

The coefficients $\boldsymbol{a}_{\alpha\beta\theta}$ are the unknown parameters we seek. To facilitate a linear algebraic solution, we establish an explicit ordering for the monomial terms. We define a bijective mapping $\sigma : \{(\alpha, \beta, \theta) \in \mathbb{N}_0^3 \mid \alpha + \beta + \theta \leqslant d_{\mathrm{FOR}}\} \rightarrow \{1, \ldots, K\}$ that uniquely maps each exponent triplet (α, β, θ) to an index s, where, $K = \binom{d_{\mathrm{FOR}}+3}{3}$ is the total number of monomial terms. This mapping allows the coefficients to be arranged into a single column vector $a \in \mathbb{R}^K$, where the s-th element, $\boldsymbol{a}^{(s)}$, corresponds to the coefficient $\boldsymbol{a}_{\alpha\beta\theta}$ such that $\sigma(\alpha, \beta, \theta) = s$.

1) Volumetric Constraint Generation: To ensure that the polynomial defines a true volume, yielding negative values inside the FOR, positive values outside, and values (near) zero on the boundary, we augment the original boundary data with additional interior and exterior points. This volumetric constraint strategy, inspired by the 3L algorithm for fitting implicit surfaces [31], enforces a clear separation between the inside and outside of the region, which cannot be achieved using boundary points alone.

Let the original data matrix be denoted $D_{\mathrm{bnd}} \equiv D_{\mathrm{FOR}}$. We generate additional constraint sets by scaling the boundary points relative to their centroid, c, defined as:

$$c = \frac{1}{n_{\mathrm{FOR}}} \sum_{l=1}^{n_{\mathrm{FOR}}} x_l \tag{7}$$

where x_l is the l-th row (sample point) of D_{bnd}. From each point x_{bnd} in the rows of D_{bnd}, we generate new points by applying a scaling factor γ. An inner point x_{in} is generated using a shrink factor $\gamma_{\mathrm{in}} < 1$, and one or more sets of outer points $\{x_{\mathrm{out},m}\}$ are generated using growth factors $\gamma_{\mathrm{out},m} > 1$. The transformation is given by:

$$x_{\mathrm{new}} = c + \gamma(x_{\mathrm{bnd}} - c). \tag{8}$$

This procedure yields an inner data matrix D_{in} and a set of S outer data matrices $\{D_{\mathrm{out},m}\}_{m=1}^{S}$, each of size $n \times 3$.

2) Linear System Formulation and Solution: The problem of finding the coefficient vector $\boldsymbol{a}$ is framed as solving an overdetermined system of linear equations. For each constraint data matrix, we construct a corresponding monomial matrix. Let D' be a generic data matrix. The monomial matrix $M(D') \in \mathbb{R}^{|D'| \times K}$ is constructed such that each row corresponds to a point x_l (the l-th row of D') and each column s corresponds to a unique monomial basis term, ordered according to the mapping σ. The entry for the l-th row and the s-th column is given by the evaluation of the s-th monomial at point x_l as $M(D')^{(l,s)} = p_l^{\alpha} q_l^{\beta} v_l^{\theta}$, where $s = \sigma(\alpha, \beta, \theta)$.

Using this definition, we construct a monomial matrix for each of our constraint sets: $M_{\mathrm{in}} = M(D_{\mathrm{in}})$, $M_{\mathrm{bnd}} = M(D_{\mathrm{bnd}})$, and $M_{\mathrm{out},m} = M(D_{\mathrm{out},m})$ for $m = 1, \ldots, S$. These matrices are vertically concatenated to form a single composite system matrix M_{FOR}. Correspondingly, we define a constraint vector b_{FOR} by concatenating vectors of target values for each region. These target vectors consist of small constants chosen to enforce the desired sign of the polynomial: $c_{\mathrm{in}} < 0$, $c_{\mathrm{bnd}} \approx 0$, and $c_{\mathrm{out},m} > 0$. The complete linear system is thus formulated as $M_{\mathrm{FOR}}\boldsymbol{a} = b_{\mathrm{FOR}}$, where

$$M_{\mathrm{FOR}} = \begin{bmatrix} M_{\mathrm{in}} \\ M_{\mathrm{bnd}} \\ M_{\mathrm{out},1} \\ \vdots \\ M_{\mathrm{out},S} \end{bmatrix}, \; b_{\mathrm{FOR}} = \begin{bmatrix} c_{\mathrm{in}} \cdot 1 \\ c_{\mathrm{bnd}} \cdot 1 \\ c_{\mathrm{out},1} \cdot 1 \\ \vdots \\ c_{\mathrm{out},S} \cdot 1 \end{bmatrix}. \tag{9}$$

Thereby, 1 is a column vector of ones of appropriate dimension. Since this system is overdetermined, an exact solution generally does not exist. We therefore seek the coefficient vector a^* that minimizes the squared Euclidean norm of the residual, $\|M\boldsymbol{a} - b\|^2$. This least-squares solution is found using the Moore-Penrose pseudoinverse, denoted by $M_{\mathrm{FOR}}^{\dagger}$:

$$a^* = M_{\mathrm{FOR}}^{\dagger} b_{\mathrm{FOR}}. \tag{10}$$

The resulting vector a^* contains the coefficients of the implicit polynomial $FOR(\boldsymbol{x})$ in (6), providing an analytical representation of the FOR.

B. Quadratic Representation of the Cost of the FOR

To monetize the FOR, we model the cost function $C(\boldsymbol{x})$ as a trivariate quadratic polynomial. This is a special case of the polynomial fitting described in the preceding section, with the degree fixed to $d_{\mathrm{cost}} = 2$. The cost function takes the general form:

$$C(\boldsymbol{x}) = \sum_{\alpha+\beta+\theta \leqslant 2} w_{\alpha\beta\theta} \boldsymbol{p}^{\alpha} \boldsymbol{q}^{\beta} \boldsymbol{v}^{\theta}. \tag{11}$$

The fitting procedure seeks the coefficients $\boldsymbol{w}_{\alpha\beta\theta}$ that best match the data generated by Algorithm 3. It follows the same linear least-squares formulation previously described. The monomial matrix $M_{\mathrm{cost}} \in \mathbb{R}^{n_{\mathrm{cost}} \times K}$ is constructed from the feature dataset D_{cost}, where $K = 10$ for a second-degree trivariate polynomial. The vector y_{cost} serves directly as the target for the regression. The coefficient vector $w^* \in \mathbb{R}^{10}$, which contains the ordered coefficients $\boldsymbol{w}_{\alpha\beta\theta}$, is then found

by solving the overdetermined system $M_{\text{cost}}\boldsymbol{w} = y_{\text{cost}}$ using the Moore-Penrose pseudoinverse:

$$w^* = M_{\text{cost}}^{\dagger} y_{\text{cost}}. \tag{12}$$

The resulting vector w^* provides the coefficients for the analytical cost function $C(\boldsymbol{x})$ in (11).

C. Theoretical Justification for the Analytical Representations

Recall that the FOR is defined over the coupling variables $\boldsymbol{x}$, which are bounded as in (2i), and therefore lie in a compact domain. Let $FOR^{\star}(\boldsymbol{x})$ denote an exact, but generally unknown, scalar function such that

$$FOR^{\star}(\boldsymbol{x}) \leqslant 0 \iff \boldsymbol{x} \text{ is feasible.} \tag{13}$$

Assume that $FOR^{\star}(\cdot)$ is continuous over this domain. Then, by the Stone-Weierstrass theorem [32], for any $\varepsilon > 0$, there exists a multivariate polynomial $FOR(\boldsymbol{x})$ such that

$$\sup_{\boldsymbol{x}} \left| FOR^{\star}(\boldsymbol{x}) - FOR(\boldsymbol{x}) \right| < \varepsilon. \tag{14}$$

This implies that the feasible region can be approximated arbitrarily well, up to an arbitrarily small boundary error, by polynomial sublevel sets of the form $FOR(\boldsymbol{x}) \leqslant 0$. Therefore, implicit polynomial representations, as adopted in (6), provide a theoretically justified and expressive framework for approximating the FOR.

A similar argument holds for the analytical representation of the flexibility cost. Given that the underlying generation costs are continuous and the operating domain is compact, the Stone-Weierstrass theorem ensures that the aggregated cost function can be approximated arbitrarily well by a polynomial proxy $C(\boldsymbol{x})$. This provides a theoretical justification for using a polynomial cost model, while the choice of a trivariate quadratic form reflects a trade-off between approximation accuracy and computational tractability.

D. Remarks and Discussion

In the following, we provide several remarks to clarify key modeling assumptions and discuss practical considerations related to the construction of the analytical functions.

1) Dependence of PQV Variables: The FOR is represented in the PQV domain using the coupling variables $\boldsymbol{x} = (\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{v})$. These variables are not treated as independent. As described in Section III, all PQV samples are obtained as feasible solutions of AC-OPF problems, ensuring that the physical interdependence among active power, reactive power, and voltage is fully preserved. Consequently, the constructed function $FOR(\boldsymbol{x})$ represents a projection of the underlying AC-OPF feasible manifold onto the PQV space, rather than an approximation over arbitrary combinations of $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{v})$.

2) Interpretation of Voltage in the PQV Domain: Within the proposed formulation, voltage is treated as a controllable variable subject to operational limits, as defined in (2i). This is consistent with its role in AC-OPF, where voltage is regulated within a safe operating range (e.g., 0.95-1.05 p.u.). The focus of this work is on feasibility and flexibility. Therefore, potential long-term effects associated with operating at extreme voltage levels are not explicitly modeled. If desired, such preferences or penalties can be incorporated into the analytical cost function $C(\boldsymbol{x})$. While the case studies in this work primarily evaluate the accuracy of the FOR and its cost representation, the proposed analytical structure allows such quality-of-service considerations to be incorporated in a straightforward manner.

3) Choice of Polynomial Order: As discussed in Section IV-C, polynomial functions possess strong approximation capabilities over bounded domains. The accuracy of the implicit polynomial representation in (6) depends on the selected degree d_{FOR}. In this sense, the polynomial degree can be interpreted as a hyperparameter. Standard hyperparameter tuning procedures, such as cross-validation or validation-based selection, can be employed to identify an appropriate degree for a given system. The impact of the polynomial degree is further examined empirically in Section V-D.

4) Re-computation of the Analytical functions under System Variations: The FOR depends on system operating conditions, including load levels, renewable generation, network topology, and operational constraints. Changes in these factors may alter the shape of the FOR. Typical examples include significant load variations (e.g., peak vs off-peak demand), fluctuations in renewable generation (e.g., PV intermittency or wind variability), network reconfiguration due to switching actions, and topology changes/equipment outages caused by faults or maintenance activities. Such events directly affect the feasible operating region defined by the underlying AC-OPF constraints. These changes necessitate recomputation of the FOR, including both the boundary dataset D_{FOR} and the corresponding model fitting $FOR(\boldsymbol{x})$.

In addition, changes in the feasible region also impact the associated flexibility cost. Since the cost dataset is generated from feasible operating points within the FOR, as described in Section III-B, any modification to the FOR requires recomputation of the interior dataset D_{cost}, followed by re-fitting of $C(\boldsymbol{x})$. Variations in DG cost coefficients may further require generation of a new dataset and recomputation of the cost function.

Despite this, the proposed sampling and fitting procedures are computationally efficient. As demonstrated in Section V-C, the total data generation time, including both boundary and interior sampling, remains within typical operational time frames and can be further accelerated via parallelization. This allows the analytical representations to be computed offline (e.g., day-ahead or intra-day) and updated periodically as system conditions evolve.

Consequently, both the FOR and the associated cost function can be interpreted as parametric objects that adapt to changing system conditions. Future work may explore adaptive or parametric representations to further improve robustness under uncertainty and reduce the need for full re-computation.

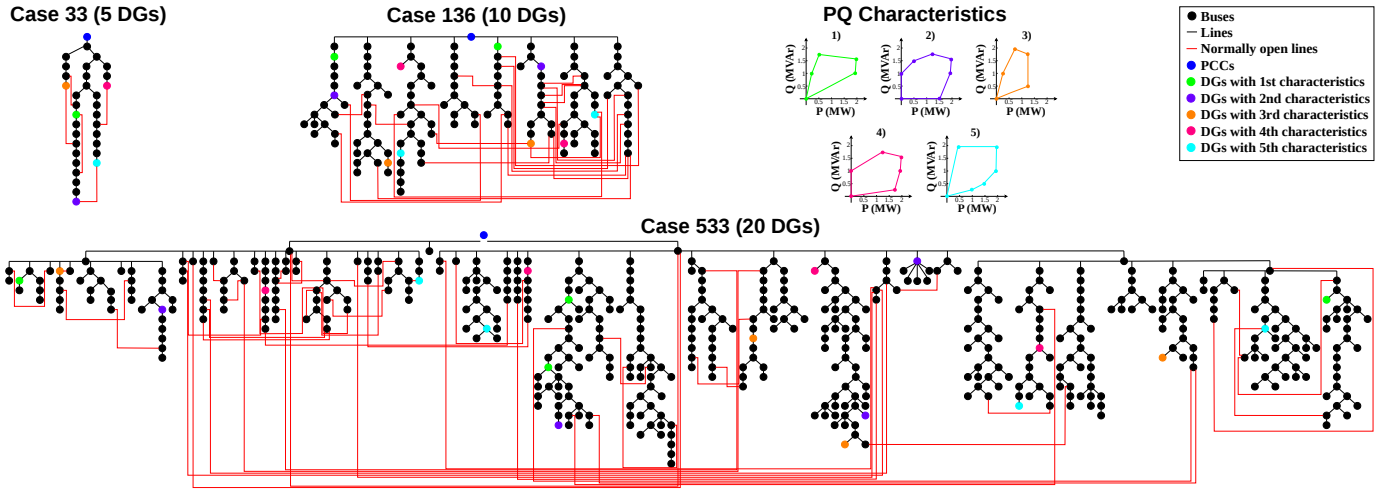

Fig. 3. Single line diagrams of the DSs and characteristics of the DGs.

TABLE III
PARAMETERS FOR THE DESIGN OF THE FUNCTIONS

DS	n_{bbps}	n_{fds}	n_{cost}	d_{FOR}	d_{cost}	γ_{in}	$\gamma_{\text{out},1}$	$\gamma_{\text{out},2}$	c_{in}	c_{bnd}	$c_{\text{out},1}$	$c_{\text{out},2}$
Case 33bw	10^3	10^4	10^3	8	2	0.999	1.005	1.07	-0.15	0	0.1	0. 2
Case 136	10^3	10^4	10^3	8	2	0.999	1.02	1.03	-0.05	0	0.1	0. 2
Case 533	10^3	10^4	10^3	8	2	0.999	1.005	1.07	-0.08	0	0.15	0. 6

The analytical formulations presented in this section complete the theoretical modeling of the FOR and its cost function. In the following section, the accuracy of these models are validated through comprehensive case studies.

V. CASE STUDIES AND DISCUSSION

To provide a rigorous validation of the analytical methodology developed in the preceding sections, we assess the performance of the proposed framework through a comprehensive set of case studies. First, DSs with diverse PQ characteristics are considered, and their FORs and associated costs are approximated using analytical functions. The accuracy, computation time, and sensitivity of these approximations are then evaluated, including the impact of sampling strategies, polynomial degree, and PCC voltage variability. Finally, the analytically represented DSs are integrated into specified TSs, where the proposed FOR-based AC-OPF framework is applied and benchmarked against both a centralized AC-OPF and an iterative privacy-preserving coordination scheme. The overall performance is evaluated in terms of feasibility, optimality, computational efficiency, and communication requirements.

All simulations are conducted in the MATLAB environment, utilizing the MATPOWER toolbox [33] with the KNITRO solver [34] for the AC-OPF problems. All case studies, except for the analysis presented in Section V-C, are conducted on a PC equipped with an Intel Core i7-10700K CPU @ 3.80 GHz and 32 GB RAM.

A. Distribution Systems Specifications

To rigorously evaluate the performance and scalability of the proposed method, we conduct case studies on 33-, 136-, and 533-bus DSs (corresponding to $j = 1, 2, 3$, respectively, where $n_{\text{ds}} = 3$). These DSs are augmented with 5, 10, and 20 DGs, respectively. Notably, the voltage at all PCCs is allowed to vary within a range of 0.95 to 1.05 p.u., rather than being fixed to a nominal value as commonly assumed in the literature. The configurations and key specifications of these systems are illustrated in Fig. 3.

As depicted in Fig. 3, all normally open lines in the selected DSs are closed, enabling meshed network topologies. This allows the proposed method to be tested under more complex meshed configurations, in contrast to many existing studies that focus solely on radial networks. Moreover, instead of assuming idealized rectangular PQ characteristics, each DS is equipped with diverse non-ideal convex PQ characteristics, in line with the modeling approach presented in [35]. Specifically, five distinct types of DG characteristics are incorporated to reflect the heterogeneity of PQ characteristic. This results in complex, three-dimensional FORs varying with voltage, allowing a comprehensive assessment of the proposed methodology.

B. Derivation of Analytical FOR and Cost Functions for DSs

In this subsection, we construct analytical approximations of the FOR and the corresponding cost functions for each DS. These approximations are derived from the datasets $D_{\text{bbps},j}$, $D_{\text{fds},j}$, i.e., $D_{\text{FOR},j}$ and $D_{\text{cost},j}$, which are obtained through the sampling strategies described earlier in Algorithms 1, 2, 3. The number of generated data points for each DS is summarized in Table III. Note that, due to the inclusion of voltage as an additional variable, constructing the three-dimensional FOR requires relatively more data points com-

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3709037

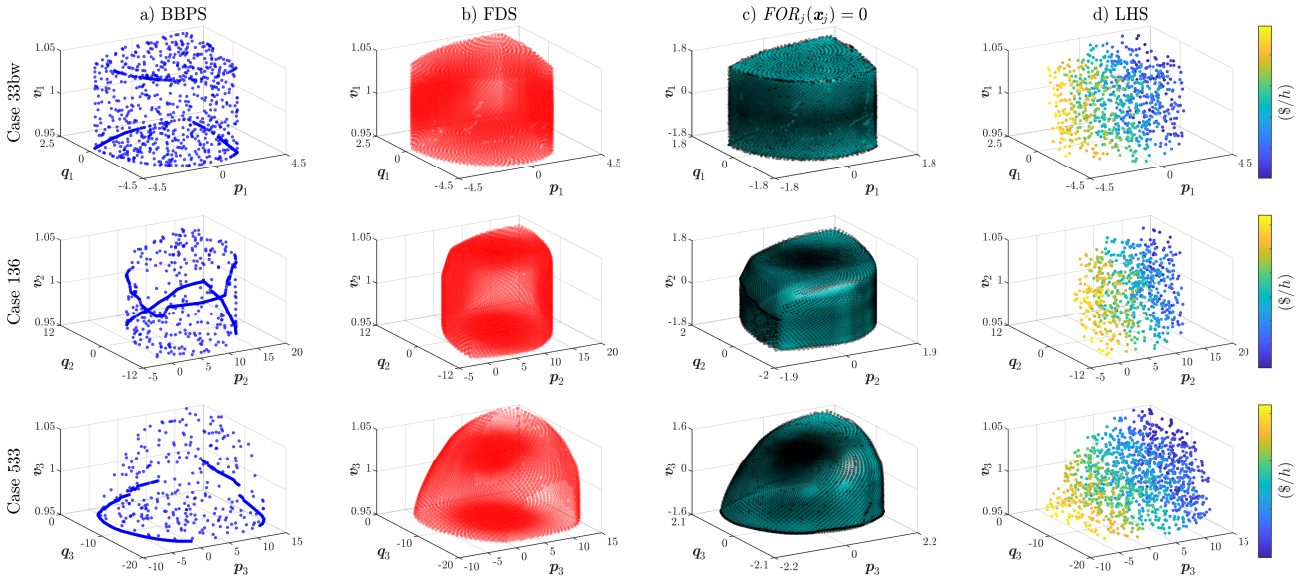

Fig. 4. a) Dataset generated using BBPS b) Dataset generated using FDS c) Constructed FOR functions (with z-score normalization) d) Cost distribution obtained via LHS. Active and reactive power are expressed in MW and MVAR, respectively, while voltage is expressed in p.u.

pared to its two-dimensional counterpart. Furthermore, the datasets are visualized in the three-dimensional PQV domain in Fig. 4a and 4b. As observed, the FOR and its associated cost characteristics can be effectively captured. As expected, the BBPS method tends to populate points along the edges of the FOR, whereas the FDS method generates samples that lie predominantly on its facets. Merging these datasets yields an effective representation of the FOR. The visualizations also reveal the emergence of complex FOR geometries, due to meshed topologies and the presence of a large number of DGs.

For the cost modeling, samples are drawn homogeneously from the interior of the FOR using the LHS method (see Fig. 4d). The cost function reflects only the generation costs of the DGs; the cost of power imported through the PCC is excluded, as it is already incorporated into the generation costs when the DS model is integrated to the TS. Here, operating points corresponding to higher power export from the DS to the TS, represented by more negative active power values under the adopted sign convention, require increased DG production and therefore incur higher costs.

After generating the datasets, we proceed to construct the analytical functions by representing them in polynomial form. The hyperparameters used for this fitting process are summarized in Table III. Specifically, one inner data matrix D_{in} and two outer data matrices D_{out} (i.e., $S = 2$) are generated. These hyperparameters are deliberately chosen to ensure that the resulting polynomial defines a conservative feasible region. Such conservatism is essential in power system operations: while classifying a feasible operating point as infeasible results in only a small cost penalty, misclassifying an infeasible operating point as feasible can lead to severe operational and reliability issues. The fitted polynomials, shown in Fig. 4c together with the boundary data points, demonstrate that the proposed approach achieves a highly accurate approximation

TABLE IV
PERFORMANCE METRICS OF THE FOR FUNCTIONS

Function	Fitting time (s)	Accuracy	Recall	Specificity
$FOR_1(\boldsymbol{x}_1)$	0.36	99.67%	99.16%	100%
$FOR_2(\boldsymbol{x}_2)$	0.37	99.68%	98.97%	100%
$FOR_3(\boldsymbol{x}_3)$	0.39	99.50%	98.14%	100%

of the FOR with minimal conservativeness, ensuring that the approximation remains entirely within the true FOR. Additionally, to further improve interpretability, two-dimensional slice plots for the Case 533 corresponding to Fig. 4c are provided in Fig. 5, which illustrate both the sampled boundary points and the fitted implicit surface more clearly. Due to space limitations, these plots are shown only for Case 533.

To evaluate the performance of the functions $FOR_j(\boldsymbol{x}_j)$, we generate random 10^5 samples within the bounding box. These samples naturally contain specific amount of feasible and infeasible samples. Using these samples, we estimate the volume of the FOR that is captured by the polynomial approximation. For this purpose, we adopt the classical confusion matrix framework. The quality of the approximation is then quantified using standard performance metrics derived from the confusion matrix, and the corresponding results are reported in Table IV.

The specificity metric evaluates the ability of the function to correctly identify infeasible points. Since all functions achieve 100% specificity, no infeasible points are misclassified as feasible, indicating that the approximation does not allow any infeasible operating points. In contrast, the recall metric measures performance on feasible samples. For example, for $FOR_1(\boldsymbol{x}_1)$, the recall is 99.16%, meaning that 99.16% of feasible points are correctly identified, while 0.84% of the FOR volume is not captured. This illustrates the trade-off be-

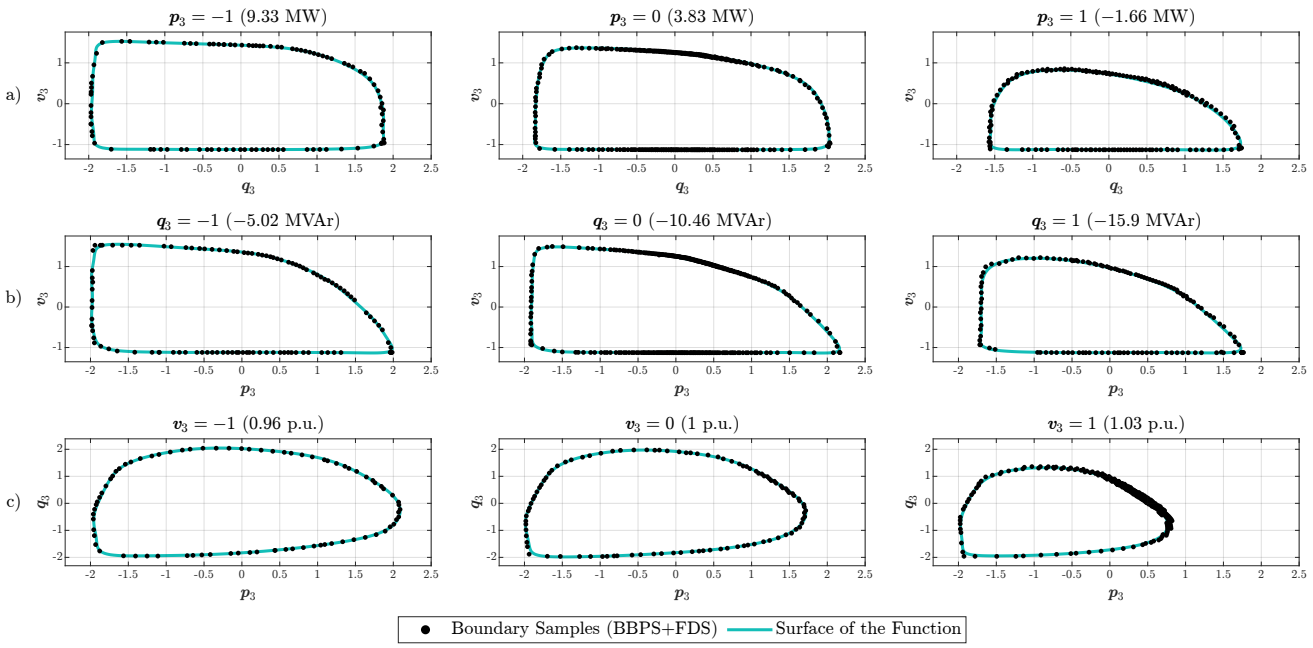

Fig. 5. Two-dimensional slice plots of the FOR for Case 533, illustrating the sampled boundary points and the surface of the fitted function. Slices are taken at fixed (a) active power, (b) reactive power, and (c) voltage, respectively. All axis values correspond to z-score normalized quantities. Figure titles correspond to the slice value in normalized and actual units.

TABLE V
PERFORMANCE METRICS OF THE COST FUNCTIONS

Function	Fitting time (s)	RMSE	MAE
$C_1(\boldsymbol{x}_1)$	0.041	0.0021	0.0016
$C_2(\boldsymbol{x}_2)$	0.045	0.0055	0.0042
$C_3(\boldsymbol{x}_3)$	0.057	0.0341	0.0266

tween ensuring complete feasibility and fully representing the volume of the FOR. Overall, when the results are examined, the proposed approximation performs as intended: a small portion of the FOR is sacrificed to reach 100% feasibility. Note that, the impact of this slight volume loss on the overall operational cost is analyzed in the following section.

Furthermore, the cost functions are constructed using the hyperparameters specified in Table III, with their performance metrics summarized in Table V. As shown, the cost functions exhibit high accuracy, indicating that they can effectively approximate the cost of the FOR. The corresponding cost distribution is illustrated in Fig. 4d. Also, the fitting times of both $FOR_j(\boldsymbol{x}_j)$ and $C_j(\boldsymbol{x}_j)$ are quite low, demonstrating the computational efficiency of the proposed approach. Overall, given that the DSs considered involve a large number of buses and DGs with complex meshed structures, the analytical approximations are obtained with consistently high performance.

C. Analysis of the Sampling Efficiency of the Proposed Method

As shown in Table IV and V, once the sampling datasets are generated, the fitting of the analytical FOR and cost functions is completed in less than one second. Consequently, the computation time of the proposed method is dominated by the sampling stage, while the construction of the analytical functions and their subsequent integration into the TSO's AC-OPF introduce negligible overhead (The computational impact of integrating the analytical functions into the TSO's OPF is separately analyzed in Section V-E).

As discussed earlier, the proposed sampling strategy is specifically designed to reduce computational complexity by focusing on boundary-relevant points and by limiting the number of required AC-OPF solves. Moreover, the three datasets $D_{\text{bbps},j}$, $D_{\text{fds},j}$, and $D_{\text{cost},j}$ are independent and can be generated in parallel. In addition, each dataset generation process is fully parallelizable internally. To evaluate the computation time of the sampling process, all simulations are performed on a server equipped with two Intel Xeon Platinum 8176M CPUs (2.10 GHz, 28 cores each).

Table VI shows the sample generation times under single-core and 56-core execution. Among the datasets, for the largest test system (Case 533), the total generation time is 200.2 s under 56-core execution. This generation time fits within typical operational time frames; for instance, market-based operations are commonly conducted in 15-minute intervals. Furthermore, since the FOR construction can be performed offline (e.g., day-ahead or ahead of market clearing), the computational burden does not pose a practical limitation for operational deployment in such cases.

The results also demonstrate near-linear scalability with respect to parallelization. Specifically, speedups of 48.46, 48.63, and 45.06 are achieved for Cases 33, 136, and 533, respectively, when comparing single-core and 56-core execution. Despite an approximately 16-fold increase in system size from Case 33 to Case 533, the observed speedup scales proportionally, indicating that the proposed method remains computationally tractable for large-scale DSs.

Finally, it should be noted that recalculation of the analyt-

TABLE VI
SAMPLE GENERATION TIME (S) FOR DIFFERENT DISTRIBUTION SYSTEMS

DS	Single-Core				56-Core			
	$D_{\text{bbps},j}$	$D_{\text{fds},j}$	$D_{\text{cost},j}$	Total	$D_{\text{bbps},j}$	$D_{\text{fds},j}$	$D_{\text{cost},j}$	Total
Case 33bw	95.91	1044.31	93.60	1233.82	2.53	20.46	2.47	25.46
Case 136	203.55	1997.67	196.39	2397.61	4.28	40.41	4.61	49.30
Case 533	789.01	7577.64	653.87	9020.52	18.81	155.54	25.85	200.20

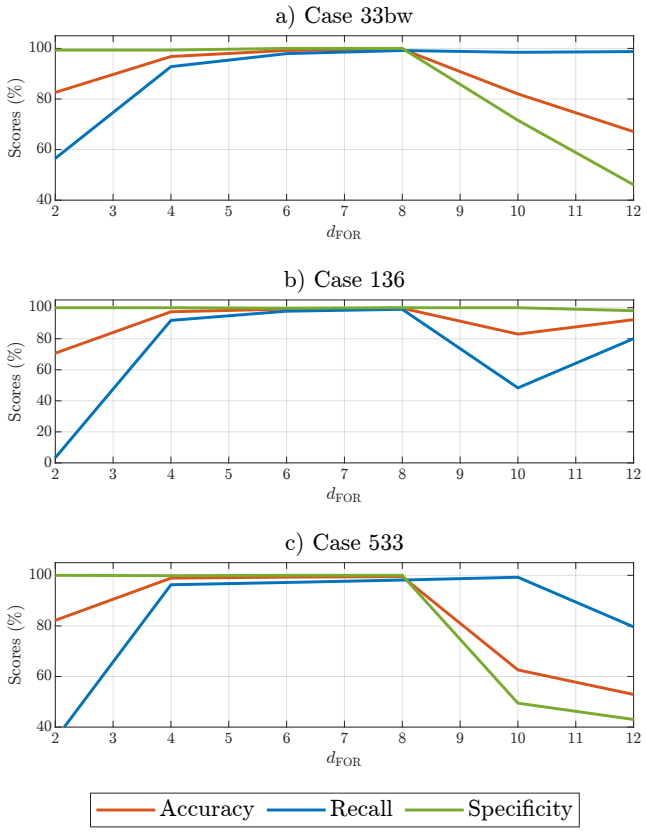

Fig. 6. Change of performance metrics according to polynomial degree.

ical functions is required when significant changes occur in network topology, operational limits, or forecasted operating conditions. As demonstrated by the computational results, such recalculations can be completed within a relatively short time and well within typical operational time windows. Therefore, the proposed framework is suitable for practical deployment in current power system operations.

D. Analysis of the Impact of Polynomial Degree

The most influential parameter affecting the performance of the analytical FOR representation is the polynomial degree, denoted by d_{FOR}. This parameter directly affects the expressive capability of the implicit polynomial and, consequently, the accuracy and reliability of the resulting FOR. To systematically assess its impact, a sensitivity analysis is conducted by varying d_{FOR} while keeping all other parameters fixed.

Fig. 6 illustrates the evolution of key performance metrics for different values of d_{FOR} across all DSs. As discussed earlier, the specificity metric is of critical importance, since values below 100% indicate that infeasible operating points may be incorrectly classified as feasible. Such misclassification would compromise system security when the FOR is embedded into the TSO-level AC-OPF. Therefore, only polynomial degrees achieving 100% specificity are considered acceptable.

From Fig. 6, it can be observed that lower polynomial degrees are insufficient to fully capture the non-convex geometry of the PQV-FOR, resulting in reduced recall and accuracy. Furthermore, excessively high polynomial degrees lead to a degradation in all metrics. Across all DSs, a polynomial degree of $d_{\text{FOR}} = 8$ consistently achieves perfect specificity while simultaneously maximizing accuracy. Consequently, $d_{\text{FOR}} = 8$ is selected as the nominal value for constructing the analytical FOR and is used throughout the subsequent integration of the FOR into the TSO-level AC-OPF formulations.

E. Incorporation of Analytical DS Representations into TS and Benchmark Against AC-OPF

In this section, the approximated DSs are integrated into the TS as formulated in Equation (2). To this end, Case 33bw, Case 136, and Case 533 DSs are integrated with the corresponding TS benchmark models from the PGLib-OPF library, namely Case 30, Case 57, and Case 162, respectively [36]. In the Case 30, buses 11, 16, and 20 are designated as the PCCs; in the Case 57, buses 7, 34, and 48 serve as the PCCs; and in the Case 162, buses 75, 109, and 129 are selected as the PCCs. In this way, three distinct TS-DS test systems are constructed, each consisting of one TS interconnected with three DSs.

After that, we conduct simulations using 1,000 randomly generated sets of cost coefficients. The results are compared against the standard AC-OPF in terms of both total cost and computational time. Fig. 7 presents histograms of the cost and time differences, using the standard AC-OPF as the reference. The feasibility ratio is consistently 100%, indicating that the proposed method never produces infeasible solutions. The cost difference remains negligible across all cases, with average deviations of 0.058%, 0.031%, and 0.002%, respectively, where these values indicate that the proposed method results in higher cost compared to the reference AC-OPF. Although the FOR of the DSs is constructed in a relatively conservative manner, its impact on the total cost is observed to be minimal. As the size of the TS increases, the influence of the DSs diminishes, leading to smaller cost deviations. In terms of computational time, the proposed method outperforms the standard AC-OPF in the majority of tests, demonstrating high computational efficiency. The average time differences are -44.22%, -58.11%, -47.62% (-0.23, -0.24, and -0.33 seconds), respectively. Note that, negative values indicate that the proposed method is faster

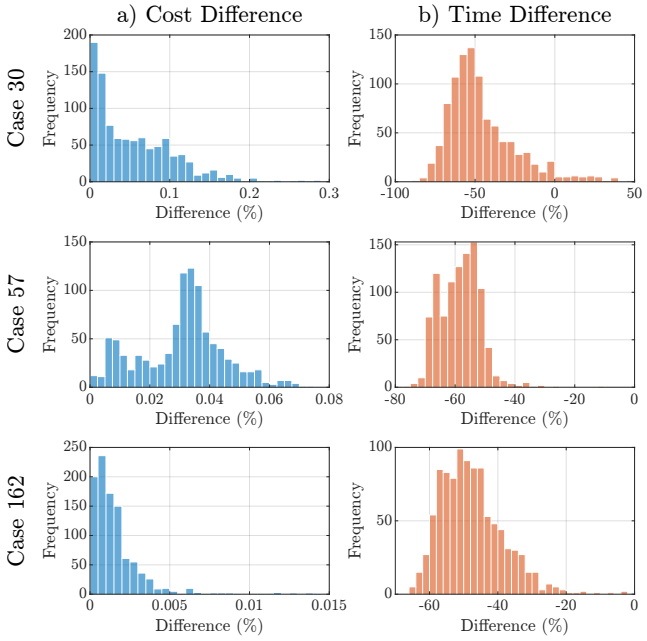

Fig. 7. Histogram of total cost and computational time differences, with AC-OPF as the reference, for integrated TS-DS. Each TS is combined with three analytically represented DSs (Case 33bw, Case 136, and Case 533). Positive cost differences indicate that the proposed method yields higher cost values than the reference, while negative time differences indicate instances where the proposed method is faster.

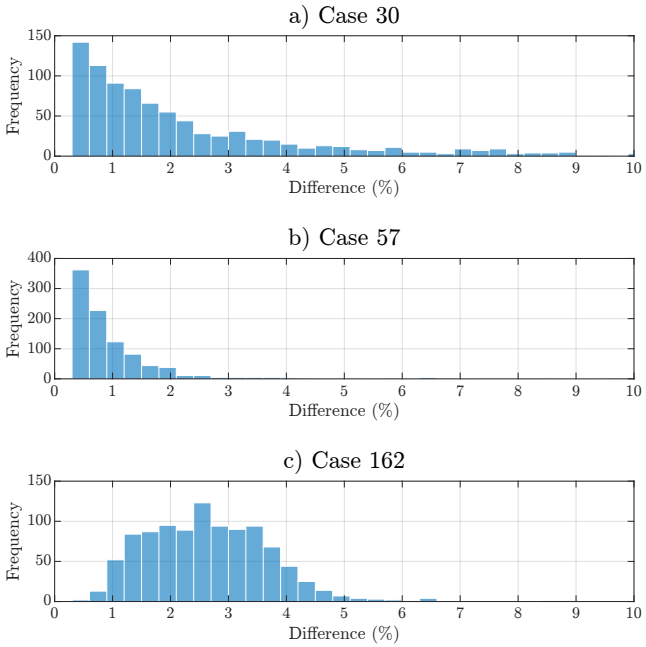

Fig. 8. Distribution of cost differences between the proposed PQV-FOR and the PQ-FOR approach with fixed PCC voltage, where PQV-FOR is used as the reference. Positive values indicate higher costs associated with the PQ-FOR approach. Outliers are omitted from the plots.

than the standard AC-OPF. These efficiency gains are achieved by representing the complex DSs with compact polynomial approximations.

Note that, within the proposed method, the TSO determines an operating point located on the FOR by considering the overall system cost. The corresponding cost value at this operating point is already available to the DSO through the previously constructed cost function. Consequently, disaggregation at the DS level can be seamlessly carried out by solving a standard AC-OPF that incorporates this operating point as an input.

The results demonstrate that the proposed analytical representations enable the tractable integration of DS flexibility into the TSO-level OPF without significant loss of optimality. Although the FOR is constructed conservatively, its impact on the total cost remains negligible, even as system size increases. By replacing complex, high-dimensional DS models with compact polynomials, the method achieves a consistent computational speedup compared to a centralized AC-OPF, while maintaining 100% feasibility and preserving DS-level data privacy. While this section focuses on the efficiency of the analytical modeling approach; the impact on coordination and communication requirements is examined in Section V-G.

F. Analysis of the Impact of the Varying Voltage at the PCCs

Most existing FOR formulations characterize DS flexibility solely in terms of active and reactive power, i.e., in the PQ space, assuming a fixed voltage at the PCC. Under this assumption, voltage-dependent network phenomena are neglected. In practice, however, voltage deviations at the PCC directly influence network losses, reactive power flows, and operational constraints, thereby affecting both the shape and size of the FOR (see Fig. 4c and Fig. 5c).

Allowing voltage to vary expands the FOR in the PQV space, resulting in a larger flexibility volume. This expanded region enables the TS to exploit DS flexibility more effectively, which in turn can improve overall system operation and reduce operational costs. Consequently, incorporating variable voltage is essential for accurately capturing DS flexibility.

To quantify the impact of PCC voltage variability, a comparative analysis is conducted between the proposed PQV-based FOR and a conventional PQ-based FOR. The analysis is performed on the same TS-DS test systems used in the previous section. For the PQ-based FOR, the PCC voltage is fixed at 1.0 p.u., while all other parameters and modeling assumptions are kept identical to isolate the effect of voltage variability. For each test system, 1,000 randomly generated sets of cost coefficients are used, following the same procedure as in the previous analysis. The cost differences are evaluated by taking the PQV-based FOR as the reference and are summarized in Fig. 8.

The results indicate that the mean cost differences are 2.01%, 1.29%, and 2.65% for Cases 30, 57, and 162, respectively, while the maximum observed differences reach 21.53%, 30.16%, and 11.11%. Here, positive values indicate that the PQ-FOR with fixed PCC voltage results in higher costs compared to the PQV-FOR reference. In all cases, feasibility is 100%, confirming that both FOR representations remain operationally valid. However, fixing the PCC voltage consistently leads to higher system costs, demonstrating that restricting voltage results in suboptimal utilization of DS flexibility.

These results show that incorporating voltage variability at

TABLE VII
COMPARISON OF STANDARD (CENTRALIZED), ADMM-BASED, AND PROPOSED TSO-DSO COORDINATION METHODS FOR THE CASE 162 TEST SYSTEM WITH 3 DSS.

	Offline phase	Online phase				Time	Cost
Method	Total offline time (s)	Online iterations	Avg. local comput. time (s/iter)	Avg. coord. time (s/iter)	Total online time (s)	Total exec. time (s)	Diff. vs standard AC-OPF
Standard AC-OPF	–	1	–	–	1.06	1.06	–
ADMM	–	130	1.48	0.88	319.95	319.95	0.36%
Proposed	200.59	1	–	–	0.73	201.32	0.07%

the PCCs is crucial not only for accurately representing the FOR geometry but also for improving the economic efficiency of TSO-DSO coordination. The proposed PQV-FOR therefore provides a more realistic and effective flexibility representation than conventional PQ-based approaches.

G. Benchmark Against an Iterative Communication Scheme

While Section V-E evaluates the performance of the proposed analytical representation against a centralized AC-OPF, the following analysis focuses on the coordination perspective by benchmarking the proposed framework against an iterative communication scheme.

The proposed analytical representation framework facilitates a single-round TSO-DSO communication scheme by shifting the bulk of the execution time to offline and highly parallelizable operations, such as sample generation and model fitting by DSOs. The only online execution involves a single operation where the TSO solves the FOR-based AC-OPF. This approach offers structural advantages over iterative communication schemes that require online local computations and continuous data exchange at every iteration, which imposes strict requirements on communication reliability and latency. To quantify these benefits, the proposed framework is benchmarked against an iterative communication scheme based on the alternating direction method of multipliers (ADMM). The ADMM-based scheme is developed using DPLib, an open-source MATLAB-based library for distributed power system optimization [37].

In the ADMM-based scheme, the TSO and DSOs are modeled as multiple regions coupled through tie lines. For each interface, every incident region maintains local copies of the boundary bus voltage magnitude and phase angle, together with auxiliary variables representing neighboring regions. At each iteration, all regions solve local AC-OPF problems in parallel using the most recent interface variables and dual multipliers. Afterward, interface voltage magnitudes and phase angles are exchanged and aggregated, and the dual variables are updated to enforce interregional consistency. This iterative exchange continues until interface consistency convergence criteria are satisfied. To reflect a *best case* situation for ADMM, an *oracle* approach is adopted, where dispatch results from 1,000 iterations are evaluated using a centralized AC-OPF to identify the iteration that yields the lowest cost while remaining feasible. The penalty parameter (ρ) is also optimized by a trial-and-error approach. This oracle-based evaluation avoids penalizing ADMM for suboptimal stopping criteria and separates coordination and communication overhead from convergence tuning effects. The ADMM benchmark employs a standard implementation, while more advanced variants with improved convergence properties exist in the literature. The purpose of this comparison is not to optimize ADMM performance, but to highlight differences in online coordination and communication requirements.

The results on the TS-DS test system with Case 162 as the TS are summarized in Table VII and include offline computation time, online computation time, iteration counts, and total execution time under an assumed one-way communication delay of 50 ms. Compared to the ADMM-based scheme, which requires 130 online iterations and incurs a total execution time of 319.95 s, the proposed framework completes online coordination in a single iteration and reduces the online execution time from 319.95 s to 0.73 s, which corresponds to a reduction of more than 99.7%. Despite an offline model construction cost of 200.59 s, the proposed method reduces the total execution time by 118.63 s, which corresponds to a 37.1% reduction, while achieving a smaller cost deviation of 0.07% compared to 0.36% for ADMM.

We remark that the communication payload per exchange is negligible for both approaches. In ADMM, each online iteration exchanges two interface variables, namely the PCC voltage magnitude and angle, per DS in each direction, which results in $4\,n_{\mathrm{DS}}$ double-precision scalars per iteration. In contrast, the proposed framework exchanges $\left(\binom{d_{\mathrm{FOR}}+3}{3}+10\right)n_{\mathrm{DS}}$ scalars once in the offline phase to communicate the analytical FOR and cost models, with no iterative online data exchange. For an 8th-degree FOR polynomial, this equals $175\,n_{\mathrm{DS}}$ scalars. These payloads are on the order of 10 to 1,000 double-precision scalars and do not pose a bandwidth limitation. The key difference lies in the frequency of data exchange required during online coordination.

These results confirm that the proposed framework substantially lowers the online computation and communication burden while preserving solution quality.

VI. CONCLUSION

In the present paper, we address key challenges in enabling effective coordination between transmission system operators (TSOs) and distribution system operators (DSOs) through a privacy-preserving representation of distribution system (DS) flexibility. We propose a comprehensive framework for constructing and utilizing a three-dimensional PQV feasible operating region (FOR), that explicitly accounts for voltage variability at the point of common coupling (PCC) and heterogeneous flexibility-providing unit (FPU) characteristics. Our method employs an AC optimal power flow (OPF)-based sampling strategy, utilizing bounding box projection

and Fibonacci direction sampling techniques, and introduces a tractable polynomial representation constructed through an implicit polynomial fitting approach. This enables a conservative yet sufficiently accurate analytical approximation of the FOR with a relatively small data set, ensuring system feasibility without excessive conservatism.

Additionally, we construct an analytical cost function associated with the FOR, enabling the economic valuation and seamless integration of DS flexibility into TSO-level decision-making processes. To operationalize this integration, we develop a FOR-based AC-OPF framework. Within this scheme, the TSO determines the optimal operating point at the PCC using the analytical models provided by DSs. Subsequently, each DSO performs local FPU dispatch by solving its own AC-OPF based on the TSO's decision. This single-round communication mechanism eliminates the need for computationally intensive disaggregation or iterative information exchange, thereby reducing online coordination and communication requirements.

We evaluate the proposed framework through extensive case studies involving large-scale, meshed DSs integrated into TSs. The analyses assess the performance of the analytical FOR and cost representations, the impact of PCC voltage variability on flexibility and economic performance, the computation time of the sampling and fitting procedures, the coordination efficiency relative to iterative benchmark and the overall efficiency of the proposed framework against the standard AC-OPF. Across all test cases, the proposed method achieves average cost deviations below 0.06% while delivering online computational speedups of up to 58% relative to standard AC-OPF. These results demonstrate that DS flexibility can be effectively integrated into power system operation under realistic modeling and privacy constraints.

Finally, it should be noted that while the results demonstrate consistent performance across a wide range of operating conditions, formal theoretical guarantees regarding the existence or smoothness of the FOR are not provided. The feasible set of the AC-OPF may exhibit non-smooth characteristics under certain conditions, and a rigorous theoretical analysis of these aspects, as well as of the sufficiency of the dataset construction strategy, constitutes an important direction for future research. The development of uncertainty-aware and topology-adaptive update mechanisms for the analytical representations is another promising direction for future work.

REFERENCES

[1] A. G. Givisiez, K. Petrou, and L. F. Ochoa, "A review on TSO-DSO coordination models and solution techniques," *Electr. Power Syst. Res.*, vol. 189, p. 106659, 2020.

[2] A. Mohammadi, M. Mehrtash, and A. Kargarian, "Diagonal quadratic approximation for decentralized collaborative TSO+ DSO optimal power flow," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2358–2370, 2018.

[3] I. Canyakmaz, C. B. Saner, and A. Varvitsiotis, "Physics-informed neural networks for privacy-preserving model sharing in power systems," in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*. IEEE, 2024, pp. 1–5.

[4] F. Capitanescu, "TSO–DSO interaction: Active distribution network power chart for TSO ancillary services provision," *Electr. Power Syst. Res.*, vol. 163, pp. 226–230, 2018.

[5] S. Kundu, K. Kalsi, and S. Backhaus, "Approximating flexibility in distributed energy resources: A geometric approach," in *Power Syst. Comput. Conf. (PSCC)*. IEEE, 2018, pp. 1–7.

[6] E. Öztürk, K. Rheinberger, T. Faulwasser, K. Worthmann, and M. Preißinger, "Aggregation of demand-side flexibilities: A comparative study of approximation algorithms," *Energies*, vol. 15, no. 7, p. 2501, 2022.

[7] D. A. Contreras and K. Rudion, "Computing the feasible operating region of active distribution networks: Comparison and validation of random sampling and optimal power flow based methods," *IET Gener. Transm. Distrib.*, vol. 15, no. 10, pp. 1600–1612, 2021.

[8] S. Alsharif, M. Sarstedt, and E. M. Veith, "Comparison of heuristic optimization-based methods for determining the flexibility potential at vertical system interconnections," in *IEEE PES Innov. Smart Grid Technol. Conf. Europe (ISGT EUROPE)*. IEEE, 2023, pp. 1–6.

[9] M. Heleno, R. Soares, J. Sumaili, R. J. Bessa, L. Seca, and M. A. Matos, "Estimation of the flexibility range in the transmission-distribution boundary," in *2015 IEEE Eindhoven PowerTech*. IEEE, 2015, pp. 1–6.

[10] M. Sarstedt, L. Kluß, J. Gerster, T. Meldau, and L. Hofmann, "Survey and comparison of optimization-based aggregation methods for the determination of the flexibility potentials at vertical system interconnections," *Energies*, vol. 14, no. 3, p. 687, 2021.

[11] M. Kalantar-Neyestanaki, F. Sossan, M. Bozorg, and R. Cherkaoui, "Characterizing the reserve provision capability area of active distribution networks: A linear robust optimization method," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2464–2475, 2019.

[12] P. Fortenbacher and T. Demiray, "Reduced and aggregated distribution grid representations approximated by polyhedral sets," *Int. J. Electr. Power Energy Syst.*, vol. 117, p. 105668, 2020.

[13] R. Vijay and P. Mathuria, "Feasibility and flexibility regions estimation at TSO–DSO interconnection node using grid structure optimization," *Sustain. Energy Grids Netw.*, vol. 32, p. 100952, 2022.

[14] J. Silva, J. Sumaili, R. J. Bessa, L. Seca, M. A. Matos, V. Miranda, M. Caujolle, B. Goncer, and M. Sebastian-Viana, "Estimating the active and reactive power flexibility area at the TSO-DSO interface," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4741–4750, 2018.

[15] L. Lopez, A. Gonzalez-Castellanos, D. Pozo, M. Roozbehani, and M. Dahleh, "Quickflex: a fast algorithm for flexible region construction for the TSO-DSO coordination," in *2021 International Conference on Smart Energy Systems and Technologies (SEST)*. IEEE, 2021, pp. 1–6.

[16] A. Rabiee, R. J. Bessa, J. Sumaili, A. Keane, and A. Soroudi, "Exploiting the determinant factors on the available flexibility area of ADNs at TSO-DSO interface," *IET Renew. Power Gener.*, vol. 18, no. 14, 2024.

[17] L. Stark and L. Hofmann, "Novel approach for flexibility aggregation at multiple vertical TSO-DSO interconnections," in *Transformation der Stromversorgung–Netzregelung und Systemführung; 15. ETG/GMA-Fachtagung, Netzregelung und Systemführung "*. VDE, 2024.

[18] K. Yoon, J. Leem, S. H. Lee, and J.-W. Park, "Cooperative control of TSO and DSO: Management of line congestion and frequency response," *IEEE Access*, 2024.

[19] L. Stark, M. Sarstedt, and L. Hofmann, "Determination of interdependent feasible operation regions at multiple TSO-DSO interconnections," in *2023 IEEE PES Innov. Smart Grid Technol. Conf. Europe (ISGT EUROPE)*. IEEE, 2023, pp. 1–6.

[20] J. Liu, Z. Tang, Y. Liu, Y. Zhou, P. P. Zeng, and Q. Wu, "Region-inspired distributed optimal dispatch of flexibility providers in coordinated transmission-distribution framework," *Energy*, p. 134985, 2025.

[21] D. A. Contreras, S. Müller, and K. Rudion, "Congestion management using aggregated flexibility at the TSO-DSO interface," in *2021 IEEE Madrid PowerTech*. IEEE, 2021, pp. 1–6.

[22] F. Capitanescu, "OPF integrating distribution systems flexibility for TSO real-time active power balance management," in *Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2018)*. IET, 2018, pp. 1–5.

[23] M. Kalantar-Neyestanaki and R. Cherkaoui, "Grid-cognizant TSO and DSO coordination framework for active and reactive power flexibility exchange: The Swiss case study," *Electr. Power Syst. Res.*, vol. 235, p. 110747, 2024.

[24] M. B. Bandeira, T. Faulwasser, and A. Engelmann, "An ADP framework for flexibility and cost aggregation: Guarantees and open problems," *Electr. Power Syst. Res.*, vol. 234, p. 110818, 2024.

[25] M. Sarstedt and L. Hofmann, "Monetarization of the feasible operation region of active distribution grids based on a cost-optimal flexibility disaggregation," *IEEE Access*, vol. 10, pp. 5402–5415, 2022.

[26] A. Churkin, M. Sanchez-Lopez, M. I. Alizadeh, F. Capitanescu, E. A. M. Ceseña, and P. Mancarella, "Impacts of distribution network reconfiguration on aggregated DER flexibility," in *2023 IEEE Belgrade PowerTech*. IEEE, 2023, pp. 1–7.

[27] F. Capitanescu, "Computing cost curves of active distribution grids aggregated flexibility for TSO-DSO coordination," *IEEE Trans. Power Syst.*, vol. 39, no. 1, pp. 2381–2384, 2023.
[28] J. Bozionek, T. Wolgast, and A. Nieße, "Design and evaluation of a multi-level reactive power market," *Energy Informatics*, vol. 5, no. 1, p. 6, 2022.
[29] D. Huntington and C. Lyrintzis, "Improvements to and limitations of Latin hypercube sampling," *Probabilistic engineering mechanics*, vol. 13, no. 4, pp. 245–253, 1998.
[30] B. Dindar, C. B. Saner, H. K. Çakmak, and V. Hagenmeyer, "Machine learning-driven multi-agent-based AC optimal power flow with effective dataset creation for data privacy and interoperability," *Sustain. Energy Grids Netw.*, vol. 42, p. 101672, 2025.
[31] M. M. Blane, Z. Lei, H. Civi, and D. B. Cooper, "The 3L algorithm for fitting implicit polynomial curves and surfaces to data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 3, pp. 298–313, 2002.
[32] M. H. Stone, "The generalized Weierstrass approximation theorem," *Mathematics Magazine*, vol. 21, no. 5, pp. 237–254, 1948.
[33] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, 2011.
[34] R. H. Byrd, J. Nocedal, and R. A. Waltz, "Knitro: An integrated package for nonlinear optimization," *Large-scale nonlinear optimization*, pp. 35–59, 2006.
[35] B. Dindar, C. B. Saner, H. K. Çakmak, and V. Hagenmeyer, "Privacy-preserving utilization of distribution system flexibility for enhanced TSO-DSO interoperability: A novel machine learning-based optimal power flow approach," *Applied Energy*, vol. 414, p. 127848, 2026.
[36] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang *et al.*, "The power grid library for benchmarking AC optimal power flow algorithms," *arXiv preprint arXiv:1908.02788*, 2019.
[37] M. Hasanzadeh and A. Kargarian, "DPLib: A standard benchmark library for distributed power system analysis and optimization," *arXiv preprint arXiv:2506.20819*, 2025.

Burak Dindar (Graduate Student Member, IEEE) received his B.S degree from Electrical Engineering Department of Kocaeli University in 2016, M.Sc. degree from Electrical Engineering Department of Istanbul Technical University in 2020. He works at Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology as a research assistant and he is a Ph.D. student in same department. His main research interests include power quality and machine learning applications in electric power systems.

Can Berk Saner (Member, IEEE) was born in Istanbul, Türkiye. He received his B.Sc. degree in Electrical Engineering from Istanbul Technical University, Türkiye, in 2015; his first M.Sc. degree in Engineering Business Management from the University of Sussex, U.K., in 2016; his second M.Sc. degree in Electrical Engineering from Istanbul Technical University in 2019; and his Ph.D. degree in Electrical and Computer Engineering from the National University of Singapore in 2024. His research interests include optimization and machine learning applications in electric power systems.

Hüseyin K. Çakmak received his Ph.D. degree from the University of Karlsruhe, Germany in 2000. He is head of the working group "Energy System Analysis" at the Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Karlsruhe, Germany. His research areas include sector-coupled energy system analysis (modeling, simulation, optimization, and visualization), 3D, virtual and augmented reality, data analysis, machine learning, real-time systems, and high-performance and distributed computing.

Veit Hagenmeyer (Member, IEEE) received the Ph.D. degree from Université Paris XI, Paris, France, in 2002. He is currently a Professor of energy informatics with the Faculty of Informatics, and the Director of the Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Karlsruhe, Germany. His research interests include modeling, optimization and control of sector integrated energy systems, machine-learning based forecasting of uncertain demand and production in energy systems mainly driven by renewables, and integrated cyber-security of such systems.