Temporal Surrogate Lagrangian Decomposition for Operational Hosting Capacity Assessment In Unbalanced Power Distribution Systems

Jingtao Qin, *Member, IEEE*, Hongbo Sun, *Senior Member, IEEE*, Nanpeng Yu, *Senior Member, IEEE*, Jianlin Guo, *Senior Member, IEEE*, Ye Wang, and Arvind Raghunathan, *Member, IEEE*

***Abstract*— The increasing penetration of distributed energy resources (DERs) necessitates advanced hosting capacity (HC) assessments to ensure reliable grid operation. Traditional static HC methods and Dynamic Operating Envelopes (DOEs) often rely on instantaneous voltage constraints, leading to overly conservative estimates. In this paper, we first propose a deterministic hosting capacity optimization model with a novel voltage violation duration constraint. Then, we build a graph convolutional recurrent network (GCRN)-based stochastic hosting capacity optimization (S-HC) framework. Finally, we propose a Surrogate Lagrangian Relaxation (SLR)-based temporal decomposition scheme to improve solving efficiency. Numerical results demonstrate the framework's effectiveness in enhancing operational HC assessment and maximizing DER integration.**

***Index Terms*— Hosting capacity, dynamic operating envelop, stochastic optimization, surrogate Lagrangian relaxation.**

I. Introduction

The integration of distributed energy resources (DERs), such as solar photovoltaic (PV) systems, has introduced new challenges and opportunities in the operation and planning of distribution systems. Central to understanding the impact of DERs is the concept of hosting capacity (HC), which is multifaceted and can be interpreted differently depending on the context. From a planning perspective, hosting capacity refers to the maximum installed capacity of DERs that the grid can physically accommodate, considering infrastructure constraints such as thermal ratings and equipment limits. From an operational perspective, hosting capacity focuses on the maximum accommodated capacity, i.e., the amount of DER generation the grid can handle under specific operating conditions without causing adverse impacts like voltage violations, thermal overloads, or instability.

In planning applications, hosting capacity assesses the capacity of DERs that can be installed under various scenarios, primarily through offline calculations [1]. While this approach is valuable for grid reinforcement and expansion planning, it may fail to capture the temporal and stochastic variability of DER generation and load profiles in near real-time. To address these limitations, Dynamic Operating Envelopes (DOEs) have emerged in recent years as a solution, providing a real-time framework for managing DER integration while also facilitating demand-side participation in future electricity markets [2]. DOEs can be viewed as a real-time adaptation of hosting capacity, designed to capture fluctuations in network conditions over short timescales [3]. They define operational boundaries within which DERs can function safely during a given period, mitigating adverse grid impacts [4]. In this context, dynamic operational HC can be viewed as the upper bound of feasible DER integration under real-time grid conditions.

Several methods for calculating DOEs and assessing the HC have been developed in recent years. Liu et al. [3] introduced a deterministic approach to compute robust DOEs, specifically addressing uncertainties in unbalanced distribution networks to support reliable integration of DERs. Additionally, unbalanced three-phase optimal power flow (UTOPF)-based techniques, as explored by Petrou et al. [5] and Liu et al. [6], have been developed to enhance DOE calculations and improve network capacity utilization. Mahmoodi et al. [7] further developed a hierarchical framework aimed at computing dynamic operating envelopes for unbalanced distribution networks. This method ensures compliance with network constraints while enabling consumers to autonomously manage their DERs, balancing grid security with flexibility in DER operation. Expanding the scope to HC forecasting, Islam et al. [8] introduced a sensitivity-enhanced recurrent neural network that dynamically incorporates exogenous and temporal factors, achieving superior predictive accuracy. In parallel, Santos et al. [9] combined optimized Volt–Var control with network reconfiguration to cost-effectively maximize DER hosting capacity while minimizing system losses. Addressing real-time operational challenges, Trinh and Chung [10] proposed coordinated active and reactive power control strategies that mitigate overvoltage and thermal violations, reducing DER curtailment and further expanding hosting limits.

Traditional static HC assessments [11] and DOEs studies usually use instantaneous voltage limits based on ANSI C84.1-2016 [12]. However, this standard defines voltage limits based on a 10-minute moving average, making instantaneous

Jingtao Qin and Nanpeng Yu are with the Electrical and Computer Engineering Department at the University of California, Riverside. (email: jingtao.qin@email.ucr.edu, nyu@ece.ucr.edu). Hongbo Sun, Jianlin Guo, Ye Wang, and Arvind Raghunathan are with Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA (e-mail: hongbosun@merl.com; guo@merl.com; yewang@merl.com; raghunathan@merl.com). Corresponding author: Nanpeng Yu.

evaluations overly conservative. To address this limitation, a dynamic HC analysis method was introduced in [13], which explicitly considers the duration of voltage violations, ensuring a more realistic and compliance-driven assessment. While this approach is based on yearly quasi-static time-series (QSTS) simulations and is primarily designed for planning-level HC assessments, it may not be directly applicable to dynamic operational HC assessments, which require real-time adaptability to changing grid conditions.

The first factor influencing operational HC is the reactive power support provided by the smart inverters of DERs. According to IEEE 1547-2018 standard [14], DERs equipped with smart inverters are required to inject or absorb reactive power to assist with voltage regulation, which can lead to a reduction in real power output. Other active network management (ANM) techniques, such as dynamic network reconfiguration (DNR) [15], [16] and tap-changing voltage regulators (VRs) [17], [18], can provide voltage support, thereby reducing DER curtailment. Recent studies have explored the joint optimization of DNR, VR tap settings, and DER reactive power control, demonstrating their potential to improve HC and grid stability [19], [20]. Various Alternating Direction Method of Multipliers (ADMM) decomposition methods have been investigated to solve inverter-based Volt-Var Control (VVC) problems within distribution system optimization. Daniel et al. [21] proposed a multi-period DER coordination approach utilizing a three-block ADMM-based distributed AC optimal power flow strategy, effectively addressing inverter-based VVC challenges with strong convergence and scalability. Yuntao et al. [22] introduced a bi-level consensus ADMM-based method for fully distributed inverter-based VVC in active distribution networks, highlighting improved convergence rates and enhanced operational reliability through hierarchical coordination. Additionally, an accelerated ADMM-based distributed inverter VVC strategy leveraging quasi-Newton techniques was presented by Xu et al. [23], demonstrating superior computational speed over conventional methods. However, the area-based decomposition schemes in these studies potentially face challenges related to communication delays, data privacy concerns, and robustness under real-time operational uncertainties. As an alternative, temporal decomposition methods may help preserve data privacy and enhance robustness in practical implementations.

Another factor contributing to the increased curtailment of DERs is the inherent conservativeness of robust optimization approaches. While these methods are designed to ensure system reliability under uncertainties, they often lead to overly conservative estimates of HC, resulting in unnecessary DER curtailment. To mitigate this issue, alternative approaches including probabilistic [24] and stochastic HC analysis [25], have been introduced to balance reliability with optimal DER integration. However, these methods typically rely on predefined statistical distributions of DERs' outputs, which may not accurately reflect real-world uncertainties. Additionally, they often fail to capture the complex spatiotemporal correlations between DERs' outputs across different buses. Graph Neural Networks (GNNs) have recently emerged as a powerful tool for time-series prediction of electrical loads and distributed energy resources (DERs) [26], demonstrating strong potential for mitigating conservativeness while ensuring system reliability.

HC assessment problems are predominantly formulated as mixed-integer linear programming (MILP) models. Despite their effectiveness, MILPs are notoriously challenging due to combinatorial complexity. Lagrangian Relaxation (LR) has long served as a promising strategy to mitigate these difficulties through problem decomposition [27]. A notable breakthrough in this direction is the Surrogate Lagrangian Relaxation (SLR) method proposed by Bragin and Tucker [28], which overcomes the persistent challenge of achieving linear-rate convergence without relying on prior knowledge or heuristic estimation of the optimal dual value. Compared with ADMM, the SLR method does not require solving augmented-Lagrangian subproblems or tuning penalty parameters, which substantially reduces computational overhead—especially in large-scale mixed-integer settings. Although ADMM provides solid $\mathcal{O}(1/k)$ convergence guarantees for convex continuous problems, its behavior in mixed-integer formulations is far less predictable. This makes SLR a more reliable and efficient choice for the stochastic HC problem considered in this work. More recently, Ben Hammouda et al. [29] extended the applicability of LR to continuous-time optimal control of coupled nonlinear systems within large-scale predictive control frameworks. Their contribution introduces a decomposable dual problem structure that facilitates efficient solution of large-scale multi-agent systems via Lagrangian duality, demonstrating strong potential for real-time distributed model predictive control. Nonetheless, advances in LR-based temporal decomposition remain underexplored and represent an important direction for future research.

To address the aforementioned limits, this paper makes the following key contributions:

1) To explicitly limit the duration of voltage violations, we introduce a novel voltage violation duration constraint, which, to our knowledge, is the first instance of such a formulation being integrated into a Mixed Integer Linear Programming (MILP) framework.
2) We develop an efficient three-phase linear power flow model that incorporates DNR and VRs to maximize the operational HC and propose a stochastic hosting capacity optimization (S-HC) model that leverages a Graph Convolutional Recurrent Network (GCRN) to effectively capture the spatiotemporal correlations between load demands and DER outputs across different buses.
3) We present a Surrogate Lagrangian Relaxation (SLR)-based temporal decomposition scheme and develop a Relaxed Stochastic Hosting Capacity (RS-HC) algorithm to improve computational efficiency. Numerical study results confirm that the proposed SLR approach substantially accelerates computation while preserving the same level of operational HC.

The remainder of this paper is structured as follows: Section II presents the dynamic operational HC model for unbalanced three-phase distribution systems. Section III introduces the SLR-based temporal decomposition scheme for the stochastic HC model outlined in Section II. Section IV provides case

studies to validate the proposed approach, and Section V concludes the paper with key findings.

II. Dynamic Operational Hosting Capacity Model for Unbalanced Three-Phase Distribution System

A. Three-phase Power Flow Model with Voltage Regulators

In this subsection, we give the three-phase power flow model with voltage regulators. For each branch (i,k), the voltage drop is:

$$\boldsymbol{u}_k = \boldsymbol{u}_i - \boldsymbol{z}_{ik}[(\boldsymbol{P}_{ik} - j\boldsymbol{Q}_{ik}) \oslash \boldsymbol{u}_i^*], \tag{1}$$

where $\boldsymbol{u}_i = [V_i^A, V_i^B, V_i^C]^{\mathcal{T}}, \boldsymbol{u}_k = [V_k^A, V_k^B, V_k^C]^{\mathcal{T}} \in \mathbb{C}^{3\times 1}$ are the three-phase voltage vector of bus i and k, $\boldsymbol{P}_{ik}, \boldsymbol{Q}_{ik} \in \mathbb{R}^{3\times 1}$ are the three-phase active and reactive power flow vector of branch (i,k), $\boldsymbol{z}_{ik} = \boldsymbol{r}_{ik} + j\boldsymbol{x}_{ik} \in \mathbb{C}^{3\times 3}$ are the three-phase impedance matrix of branch (i,k), $\oslash$ is the element-wise division.

Assuming that the voltage magnitudes between the phases are nearly balanced [30], we multiply both sides of (1) by its conjugate to obtain:

$$|\boldsymbol{u}_k|^2 = |\boldsymbol{u}_i|^2 - 2(\tilde{\boldsymbol{r}}_{ik}\boldsymbol{P}_{ik} + \tilde{\boldsymbol{x}}_{ik}\boldsymbol{Q}_{ik}) + \boldsymbol{c}_{ik}^u(\boldsymbol{P}_{ik}, \boldsymbol{Q}_{ik}), \tag{2}$$

where $\tilde{\boldsymbol{r}}_{ik} = \mathrm{Re}\{\boldsymbol{T}\boldsymbol{T}^H\} \otimes \boldsymbol{r}_{ik} + \mathrm{Im}\{\boldsymbol{T}\boldsymbol{T}^H\} \otimes \boldsymbol{x}_{ik}$ is the equivalent resistance matrix, $\tilde{\boldsymbol{x}}_{ik} = \mathrm{Re}\{\boldsymbol{T}\boldsymbol{T}^H\} \otimes \boldsymbol{x}_{ik} - \mathrm{Im}\{\boldsymbol{T}\boldsymbol{T}^H\} \otimes \boldsymbol{r}_{ik}$ is the equivalent reactance matrix, $\boldsymbol{T} = [1, e^{-j2\pi/3}, e^{j2\pi/3}]$, and $\boldsymbol{c}_{ik}^u(\boldsymbol{P}_{ik}, \boldsymbol{Q}_{ik})$ is the high-order linear approximation, $\otimes$ is the element-wise product.

Consider a three-phase, wye-wye solidly grounded voltage regulator situated in branch (i,k), which can be equivalently represented by an ideal transformer in series with an impedance [31], as illustrated in Fig. 1. In this model, bus s is a virtual node between bus i and bus k, connected to the secondary side of the ideal transformer, and $\boldsymbol{z}$ represents the combined impedance of the voltage regulator and the branch.

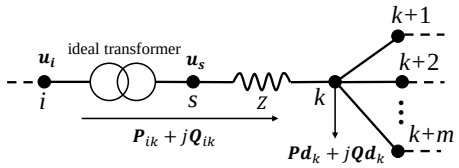


Fig. 1. Branch with an equivalent voltage regulator model.

To obtain a linear formulation of voltage regulators, we first apply Special Ordered Set type 1 (SOS1) to model the voltage relationship between bus i and s as:

$$U_s^\varphi = U_i^\varphi \cdot \sum_{m=1}^{M_{ik}} x_{ik}^m w_{ik}^m, \quad \sum_{m=1}^{M_{ik}} x_{ik}^m \cdot w_{ik}^m = (\eta_{ik}^m)^2, \tag{3}$$

$$\sum_{m=1}^{M_{ik}} x_{ik}^m \cdot m = \tau_{ik}, \quad \sum_{m=1}^{M_{ik}} x_{ik}^m = 1, \tag{4}$$

where $U_i^\varphi = |V_i^\varphi|^2$ is the squared voltage magnitude, η_{ik} denotes the turns ratio of the voltage regulator, τ_{ik} represents the tap position, M_{ik} is the total number of taps, and x_{ik}^m, w_{ik}^m are a binary variable and its associated non-negative weight, respectively.

Then, we define a new variable $y_{ik}^m = U_i^\varphi \cdot x_{ik}^m$ and apply big-M method to relax constraint (3) as:

$$0 \le y_{ij}^m \le M \cdot x_{ij}^m, \quad 0 \le U_i^\varphi - y_{ij}^m \le M \cdot (1 - x_{ij}^m) \tag{5}$$

A similar procedure can be applied to the power flow balance equation of branch (i,k), which will give us:

$$\begin{cases} \boldsymbol{P}_{ik} = \sum\limits_{j\in\mathbb{D}(k)} \boldsymbol{P}_{kj} + \boldsymbol{P}\boldsymbol{d}_k - \boldsymbol{P}\boldsymbol{g}_k + c_{ik}^p(\boldsymbol{P}, \boldsymbol{Q}) \\ \boldsymbol{Q}_{ik} = \sum\limits_{j\in\mathbb{D}(k)} \boldsymbol{Q}_{kj} + \boldsymbol{Q}\boldsymbol{d}_k - \boldsymbol{Q}\boldsymbol{g}_k + c_{ik}^q(\boldsymbol{P}, \boldsymbol{Q}), \end{cases} \tag{6}$$

where $\boldsymbol{P}\boldsymbol{d}_k, \boldsymbol{Q}\boldsymbol{d}_k$ are the active and reactive load at bus k, $\boldsymbol{P}\boldsymbol{g}_k, \boldsymbol{Q}\boldsymbol{g}_k$ are the active and reactive DER generation at bus k, $c_{ik}^p(\boldsymbol{P}, \boldsymbol{Q}), c_{ik}^q(\boldsymbol{P}, \boldsymbol{Q})$ are the higher-order linear approximation of active and reactive power loss, $\mathbb{D}(k)$ is the set of downstream buses connected to bus k. The details of approximation (2) and (6) can be found in Appendix A and Appendix B in reference [30].

B. Deterministic Hosting Capacity Optimization Model

In this subsection, we develop a deterministic dynamic operational HC optimization model for DERs. The objective is to maximize the total active operational HC of all DERs that the system can accommodate, or equivalently, to minimize the total active power curtailment of all DERs. To reduce the frequency of switch and voltage regulator operations, we include their operational costs in the objective function.

1) Objective function:

$$\begin{aligned} \text{Obj.} \min_{Pg,\gamma,\rho} \sum_{t\in\mathcal{T}} \Bigg(& \sum_{i\in\mathcal{G}} \sum_{\varphi\in\psi_i} (\widehat{Pg}_{i,t}^\varphi - Pg_{i,t}^\varphi) \\ & + \sum_{(i,k)\in\mathcal{E}^s} \omega^s \gamma_{ik,t} + \sum_{(i,k)\in\mathcal{E}^r} \omega^r \rho_{ik,t} \Bigg), \end{aligned} \tag{7}$$

where $\mathcal{T}$ is the set of time periods, $\widehat{Pg}_{i,t}^\varphi, Pg_{i,t}^\varphi$ are the maximum available and actual active power of DER on bus i at time t, $\mathcal{G}$ is the set of buses which have DERs, $\mathcal{E}^s, \mathcal{E}^r$ are the set of branches that have switches and VRs, ω^s, ω^r are the operation cost coefficient of switches and voltage regulators, $\gamma_{ik,t}, \rho_{ik,t}$ are the operation times of switch and VR on branch (i,k) at time t. The following constraints are included in the optimization problem:

2) DER output constraints:

$$\begin{cases} Pg_{i,t}^\varphi = \zeta_{i,t}^\varphi \cdot \widehat{Pg}_{i,t}^\varphi, \ 0 \le \zeta_{i,t}^\varphi \le 1 \\ -Sg_i^\varphi \le Pg_{i,t}^\varphi \le Sg_i^\varphi, -Sg_i^\varphi \le Qg_{i,t}^\varphi \le Sg_i^\varphi \\ -\sqrt{2}Sg_i^\varphi \le Pg_{i,t}^\varphi + Qg_{i,t}^\varphi \le \sqrt{2}Sg_i^\varphi \\ -\sqrt{2}Sg_i^\varphi \le Pg_{i,t}^\varphi - Qg_{i,t}^\varphi \le \sqrt{2}Sg_i^\varphi \\ \forall i \in \mathcal{G}, \forall t \in \mathcal{T}, \forall \varphi \in \psi_i, \end{cases} \tag{8}$$

where $\zeta_{i,t}^\varphi$ is the capacity efficiency factor of DER on bus i at time t, $Qg_{i,t}^\varphi$ is the reactive power of DER on bus i at time t, $Sg_{i,t}^\varphi$ is the installed capacity of DER on bus

i. To simplify the original quadratic capacity constraint, we employ the circle constraint linearization method proposed in [32], which provides an efficient approximation. The same approach is also applied to the thermal capacity constraints of the branches.

3) Branch power flow balance constraints:

$$\begin{cases} P^{\varphi}_{ik,t} = \sum\limits_{j\in\mathbb{D}(k)} P^{\varphi}_{kj,t} + Pd^{\varphi}_{k,t} - Pg^{\varphi}_{k,t} \\ Q^{\varphi}_{ik,t} = \sum\limits_{j\in\mathbb{D}(k)} Q^{\varphi}_{kj,t} + Qd^{\varphi}_{k,t} - Qg^{\varphi}_{k,t} \\ \forall k \in \mathcal{N}, \forall t \in \mathcal{T}, \forall \varphi \in \psi_i, \end{cases} \tag{9}$$

where $\mathcal{N}$ is the set of all buses. Here, we neglect the higher-ordered term in (6).

For branches without a voltage regulator:

$$\begin{cases} U^{\varphi}_{k,t} \le U^{\varphi}_{i,t} - 2(\tilde{\boldsymbol{r}}^{\varphi}_{ik}\boldsymbol{P}_{ik,t} + \tilde{\boldsymbol{x}}^{\varphi}_{ik}\boldsymbol{Q}_{ik,t}) + M(1-\alpha_{ik,t}) \\ U^{\varphi}_{k,t} \ge U^{\varphi}_{i,t} - 2(\tilde{\boldsymbol{r}}^{\varphi}_{ik}\boldsymbol{P}_{ik,t} + \tilde{\boldsymbol{x}}^{\varphi}_{ik}\boldsymbol{Q}_{ik,t}) - M(1-\alpha_{ik,t}) \\ \forall (i,k) \in \mathcal{E}/\mathcal{E}^r, \forall t \in \mathcal{T}, \forall \varphi \in \psi_i, \end{cases} \tag{10}$$

where $\alpha_{ik,t}$ is a binary indicator variable which equals 1 if branch (i,k) is connected at time t, $\mathcal{E}$ is the set of all branches. Again, we neglect the higher-order term in (2).

For branches with a voltage regulator:

$$\begin{cases} (3)-(5) \\ U^{\varphi}_{k,t} \le U^{\varphi}_{s,t} - 2(\tilde{\boldsymbol{r}}^{\varphi}_{ik}\boldsymbol{P}_{ik,t} + \tilde{\boldsymbol{x}}^{\varphi}_{ik}\boldsymbol{Q}_{ik,t}) + M(1-\alpha_{ik,t}) \\ U^{\varphi}_{k,t} \ge U^{\varphi}_{s,t} - 2(\tilde{\boldsymbol{r}}^{\varphi}_{ik}\boldsymbol{P}_{ik,t} + \tilde{\boldsymbol{x}}^{\varphi}_{ik}\boldsymbol{Q}_{ik,t}) - M(1-\alpha_{ik,t}) \\ \forall (i,k) \in \mathcal{E}^r, \forall t \in \mathcal{T}, \forall \varphi \in \psi_i, \end{cases} \tag{11}$$

4) Voltage magnitude and imbalance constraints:

$$\begin{cases} U^{\varphi}_{i,t} = U_{\text{ref}}, \forall i \in \mathcal{S} \\ U^{\min}_i \le U^{\varphi}_{i,t} \le U^{\max}_i, \forall i \in \mathcal{N}/\mathcal{S} \\ -\epsilon \le \dfrac{U^{\varphi}_{i,t} - U^{avg}_{i,t}}{U^{avg}_{i,t}} \le \epsilon,\ U^{avg}_{i,t} = \dfrac{1}{3}\sum\limits_{\varphi\in\psi_i} U^{\varphi}_{i,t}, \forall i \in \mathcal{N}_3 \\ \forall t \in \mathcal{T}, \forall \varphi \in \psi_i, \end{cases} \tag{12}$$

where $U^{\min}_i, U^{\max}_i$ are range B service voltage in the ANSI standard [12], $\mathcal{S}$ and $\mathcal{N}_3$ are the set of substations and the set of three-phase buses, $U^{avg}_{i,t}$ is the average voltage magnitude on bus i at time t. The third equation is used to regulate the imbalance level for all three-phase buses. ϵ is the voltage imbalance limit factor.

5) Voltage violation duration limits: In addition to the instantaneous voltage magnitude limits, we also incorporate novel voltage violation duration limits to capture dynamic information. Specifically, we limit both the overall duration of voltage violations and the duration of consecutive voltage violations, formulated as follows:

$$\begin{cases} \overline{U}_i \cdot a^{\varphi}_{i,t} + \varepsilon^{\varphi}_{i,t} \le U^{\varphi}_{i,t} \le \overline{U}_i + M \cdot a^{\varphi}_{i,t} \\ \underline{U}_i \cdot (1-b^{\varphi}_{i,t}) \le U^{\varphi}_{i,t} \le \underline{U}_i + M \cdot (1-b^{\varphi}_{i,t}) - \varrho^{\varphi}_{i,t} \\ a^{\varphi}_{i,t} + b^{\varphi}_{i,t} \le 1,\ \forall i \in \mathcal{N}_d, \forall t \in \mathcal{T}, \forall \varphi \in \psi_i \end{cases} \tag{13}$$

$$\begin{cases} \sum\limits_{t\in\mathcal{T}} \left(a^{\varphi}_{i,t} + b^{\varphi}_{i,t}\right) \le d_1,\ \forall i \in \mathcal{N}_d, \forall \varphi \in \psi_i \\ \sum\limits_{k=t}^{t+d_2} \left(a^{\varphi}_{i,k} + b^{\varphi}_{i,k}\right) \le d_2, \forall i \in \mathcal{N}_d, \forall t \in \mathcal{T}_d, \forall \varphi \in \psi_i \end{cases} \tag{14}$$

where $\overline{U}_i, \underline{U}_i$ are range A service voltage in the ANSI standard [12], $a^{\varphi}_{i,t}, b^{\varphi}_{i,t}$ are the binary indicators that voltage on bus i at time t for phase φ violate range A service voltage, $\varepsilon^{\varphi}_{i,t}, \varrho^{\varphi}_{i,t}$ are non-negative auxiliary variables, d_1, d_2 are the duration limits, $\mathcal{N}_d$ is the monitor bus set, $\mathcal{T}_d$ is the monitor time set, $|\mathcal{T}_d| = |\mathcal{T}| - d_2$.

6) Branch thermal capacity constraints:

$$\begin{cases} -\alpha_{ik,t}S^{\max}_{ik} \le P^{\varphi}_{ik,t} \le \alpha_{ik,t}S^{\max}_{ik} \\ -\alpha_{ik,t}S^{\max}_{ik} \le Q^{\varphi}_{ik,t} \le \alpha_{ik,t}S^{\max}_{ik} \\ -\sqrt{2}\alpha_{ik,t}S^{\max}_{ik} \le P^{\varphi}_{ik,t} + Q^{\varphi}_{ik,t} \le \sqrt{2}\alpha_{ik,t}S^{\max}_{ik} \\ -\sqrt{2}\alpha_{ik,t}S^{\max}_{ik} \le P^{\varphi}_{ik,t} - Q^{\varphi}_{ik,t} \le \sqrt{2}\alpha_{ik,t}S^{\max}_{ik}, \end{cases} \tag{15}$$

where $S^{\max}_{ik}$ is the thermal capacity of branch (i,k).

7) Radiality constraints:

$$\begin{cases} \sum\limits_{(i,k)\in\mathcal{E}} \alpha_{ik,t} = |\mathcal{N}| - |\mathcal{S}| \\ \beta_{ik,t} + \beta_{ki,t} = \alpha_{ik,t} \end{cases} \tag{16}$$

$$\begin{cases} \sum\limits_{k\in\mathbb{N}(i)} \beta_{ki,t} = 1, \forall i \in \mathcal{N}/\mathcal{S} \\ \beta_{ki,t} = 0, \forall i \in \mathcal{S}, \end{cases} \tag{17}$$

where $|\mathcal{N}|$ and $|\mathcal{S}|$ are the total number of buses and substation buses, $\beta_{ki,t}$ is a binary variable that equals 1 if k is the upstream bus of i.

8) Switch and voltage regulator operation constraints:

$$\begin{cases} \gamma_{ik,t} \ge \alpha_{ik,t} - \alpha_{ik,t-1} \\ \gamma_{ik,t} \ge \alpha_{ik,t-1} - \alpha_{ik,t} \\ \rho_{ik,t} \ge \tau_{ik,t} - \tau_{ik,t-1} \\ \rho_{ik,t} \ge \tau_{ik,t-1} - \tau_{ik,t} \end{cases} \tag{18}$$

$$\begin{cases} \sum\limits_{t\in\mathcal{T}} \gamma_{ik,t} \le \gamma^{\max}_{ik} \\ \sum\limits_{t\in\mathcal{T}} \rho_{ik,t} \le \rho^{\max}_{ik}, \end{cases} \tag{19}$$

where $\gamma^{\max}_{ik}, \rho^{\max}_{ik}$ are the daily operation limit for switch and voltage regulator on branch (i,k).

C. Stochastic Hosting Capacity Optimization Model

Unlike traditional methods that often rely on simplified statistical assumptions and fail to capture spatiotemporal dependencies, we employ a Graph Convolutional Recurrent Network (GCRN) to address the uncertainties in load demands and DER outputs. Subsequently, we develop a stochastic HC optimization model powered by the predictions of the GCRN.

Let $\vartheta_{t-K+1}, \vartheta_{t-K+2}, \ldots, \vartheta_t \in \mathbb{R}^{n\times 3}$ represent a sequence of length-K historical load demand or DER output shape matrices in the unbalanced distribution network. The predictions of loads or DER outputs at time $t+1$ can be obtained as:

$$\hat{\vartheta}_{t+1} = F_\Theta(\vartheta_{t-K+1}, \ldots, \vartheta_t), \tag{20}$$

where Θ is the set of trainable weight matrices. Specifically, the GCRN integrates spatial and temporal modeling, followed by prediction generation, as detailed below:

1) Spatial convolution: At each time step t, the input matrix ϑ_t is processed by a graph convolution network (GCN) to capture spatial dependencies:

$$H_t = \sigma\left(D^{-1/2}(A+I)D^{-1/2}\vartheta_t W\right), \tag{21}$$

where A is the adjacency matrix of the graph built from the power distribution network, I is the identity matrix used to add self-loops to the graph, D is the diagonal degree matrix with $D_{ii} = \sum_j (A_{ij} + I_{ij})$, W is the trainable weight matrix, and $\sigma(\cdot)$ is the activation function.

2) Temporal convolution: The spatially encoded output H_t from the GCN is then fed into a Recurrent Neural Network (RNN) to model temporal dependencies across time steps. The hidden state h_t is updated as:

$$h_t = \sigma\left(W_x^T H_t^T + W_h^T h_{t-1} + \boldsymbol{b}\right), \tag{22}$$

where h_t is the hidden state at time t, which captures temporal information accumulated up to time t, W_x, W_h are the trainable weight matrices, and $\boldsymbol{b}$ is a bias vector.

3) Prediction generation: The hidden state h_t, which encapsulates both spatial and temporal dependencies, is passed through a fully connected layer to predict the load or DER output shape for each bus at time $t+1$:

$$\hat{\vartheta}_{t+1} = \phi\left(W_o^T h_t\right), \tag{23}$$

where W_o is a trainable weight matrix, $\phi(\cdot) : \mathbb{R}^{3|\mathcal{C}|\times n} \rightarrow \mathbb{R}^{n\times 3\times|\mathcal{C}|}$ is a reshape function, $\mathcal{C}$ is the set of scenarios. This step provides multiple predictions of load demand/DER output shape values.

To introduce stochasticity for HC optimization, the probability mass function (PMF) of load demand or DER output shape at time $t+1$ on bus i is given by:

$$\text{PMF}(\vartheta^{\varphi}_{i,t+1} = \hat{\vartheta}^{\varphi,c}_{i,t+1}) = \frac{1}{|\mathcal{C}|}, \tag{24}$$

where $c \in \mathcal{C}$ represents the scenario index. Here, we assume an equal probability distribution across all scenarios. Additionally, we assume the active and reactive load demand shapes are the same for each bus at time $t+1$, which means:

$$Pd^{\varphi}_{i,t+1} = \overline{Pd}^{\varphi}_{i,t+1} \cdot \vartheta^{\varphi}_{i,t+1}, Qd^{\varphi}_{i,t+1} = \overline{Qd}^{\varphi}_{i,t+1} \cdot \vartheta^{\varphi}_{i,t+1} \tag{25}$$

where $\overline{Pd}^{\varphi}_{i,t+1}, \overline{Qd}^{\varphi}_{i,t+1}$ are the nominal active and reactive load on bus i at time $t+1$.

Once the PMF for all load demands and DER outputs is obtained, multiple scenarios are generated for each time period. To improve clarity, all decision variables and parameters are categorized into two distinct groups, represented as vectors. The variables in group $\boldsymbol{x}$ vary across different scenarios, capturing the system's stochastic nature, while the variables in group $\boldsymbol{y}$ remain consistent across all scenarios, ensuring stability in the optimization process.

$$\begin{aligned}
&\boldsymbol{x}: \Big\{ Pg^{\varphi}_{i,t}, Qg^{\varphi}_{i,t}, \widehat{Pg}^{\varphi}_{i,t}, Pd^{\varphi}_{i,t}, Qd^{\varphi}_{i,t}, P^{\varphi}_{ik,t}, Q^{\varphi}_{ik,t}, \\
&\quad U^{\varphi}_{i,t}, \bar{U}^{\varphi}_{i,t}, a^{\varphi}_{i,t}, b^{\varphi}_{i,t}, \bar{\varepsilon}^{\varphi}_{i,t}, \underline{\varepsilon}^{\varphi}_{i,t} \Big\} \\
&\boldsymbol{y}: \left\{ \zeta^{\varphi}_{i,t}, x^{m}_{ik,t}, y^{m}_{ik,t}, \alpha_{ik,t}, \beta_{ik,t}, \gamma_{ik,t}, \tau_{ik,t}, \rho_{ik,t} \right\}
\end{aligned} \tag{26}$$

Accordingly, the stochastic hosting capacity optimization model (S-HC) can be formulated as follows:

$$\text{S-HC:} \begin{cases} \text{Obj. } \min\limits_{Pg,\gamma,\rho} \sum\limits_{t\in\mathcal{T}} \left(\dfrac{1}{|\mathcal{C}|} \sum\limits_{c\in\mathcal{C}} \sum\limits_{i\in\mathcal{G}} \sum\limits_{\varphi\in\psi_i} (\widehat{Pg}^{\varphi,c}_{i,t} - Pg^{\varphi,c}_{i,t}) \right. \\ \qquad \left. + \sum\limits_{(i,k)\in\mathcal{E}^s} \omega^s \gamma_{ik,t} + \sum\limits_{(i,k)\in\mathcal{E}^r} \omega^r \rho_{ik,t} \right) \\ \text{s.t. } (8)-(19),\ \boldsymbol{x} \rightarrow \boldsymbol{x}^c,\ \forall c \in \mathcal{C}, \end{cases} \tag{27}$$

All decision variables in group $\boldsymbol{x}$ will be extended by adding a scenario-specific dimension, and the expanded variables will be represented as $\boldsymbol{x}^c$.

III. Surrogate Lagrangian Relaxation-based Temporal Decomposition for Stochastic Hosting Capacity Assessment

In this section, we introduce a temporal decomposition approach utilizing the Surrogate Lagrangian Relaxation (SLR) algorithm [33] to enhance the solving efficiency of S-HC. The temporal decomposition consists of the following steps.

A. Dualization

First, we dualize the temporal daily operation constraint (19) and the voltage duration limit in (14), and augment the objective function of S-HC by incorporating the product of the Lagrange multipliers and these temporal constraints. We derive a relaxed stochastic hosting capacity optimization model (RS-HC) after rearrangement as:

$$\text{RS-HC:} \begin{cases} \text{Obj. } \min\limits_{Pg,\gamma,\rho} \sum\limits_{j\in\mathcal{J}} \left(\dfrac{1}{|\mathcal{C}|} \sum\limits_{c\in\mathcal{C}} \sum\limits_{i\in\mathcal{G}} \sum\limits_{t\in\mathcal{T}_j} \sum\limits_{\varphi\in\psi_i} (\widehat{Pg}^{\varphi,c}_{i,t} - Pg^{\varphi,c}_{i,t}) \right. \\ \quad + \sum\limits_{c\in\mathcal{C}} \sum\limits_{i\in\mathcal{N}_d} \sum\limits_{\varphi\in\psi_i} \left(\nu^{\varphi,c}_i \sum\limits_{t\in\mathcal{T}_j} (a^{\varphi,c}_{i,t} + b^{\varphi,c}_{i,t}) - \nu^{\varphi,c}_i d_1 \right. \\ \quad \left. + \sum\limits_{t\in\mathcal{T}_j} \sum\limits_{k=K_1}^{K_2} \upsilon^{\varphi,c}_{i,k} (a^{\varphi,c}_{i,t} + b^{\varphi,c}_{i,t}) - \sum\limits_{t\in\mathcal{T}_j/\mathcal{T}_d} \upsilon^{\varphi,c}_{i,t} d_2 \right) \\ \quad + \sum\limits_{(i,k)\in\mathcal{E}^s} \left((\lambda_{ik} + \omega^s) \sum\limits_{t\in\mathcal{T}_j} \gamma_{ik,t} - \lambda_{ik}\gamma^{\max}_{ik} \right) \\ \quad \left. + \sum\limits_{(i,k)\in\mathcal{E}^r} \left((\lambda_{ik} + \omega^r) \sum\limits_{t\in\mathcal{T}_j} \rho_{ik,t} - \lambda_{ik}\rho^{\max}_{ik} \right) \right) \\ \text{s.t. } (8)-(13), (15)-(18), \boldsymbol{x} \rightarrow \boldsymbol{x}^c, \forall c \in \mathcal{C}, \forall j \in \mathcal{J}, \end{cases} \tag{28}$$

where $\mathcal{T}_j$ is a subset of time horizons, $\mathcal{J}$ is the set of subproblems, $K_1 = \max(1, t-d), K_2 = \min(t, |\mathcal{T}|-d)$, $\nu^{\varphi}_{i,k}, \upsilon^{\varphi}_{i,k}$ are the dual variables associated with voltage duration limit (14), λ_{ik} is the dual variable associated with daily operation limit (19) for switches or VRs.

B. Decomposition

$$\text{Sub-RS-HC:}\begin{cases}\text{Obj. } \min\limits_{Pg,\gamma,\rho} \sum\limits_{c\in\mathcal{C}}\sum\limits_{i\in\mathcal{G}}\sum\limits_{t\in\mathcal{T}_j}\sum\limits_{\varphi\in\psi_i}\left(\frac{1}{|\mathcal{C}|}(\widehat{Pg}_{i,t}^{\varphi,c}-Pg_{i,t}^{\varphi,c,n})\right.\\ \quad+\sum\limits_{c\in\mathcal{C}}\sum\limits_{i\in\mathcal{N}_d}\sum\limits_{\varphi\in\psi_i}\left(\nu_i^{\varphi,c}\sum\limits_{t\in\mathcal{T}_j}(a_{i,t}^{\varphi,c}+b_{i,t}^{\varphi,c})\right.\\ \quad\left.\left.+\sum\limits_{t\in\mathcal{T}_j}\sum\limits_{k=K_1}^{K_2}\upsilon_{i,k}^{\varphi,c}(a_{i,t}^{\varphi,c}+b_{i,t}^{\varphi,c})\right)\right.\\ \quad+\sum\limits_{(i,k)\in\mathcal{E}^s}(\lambda_{ik}^n+\omega_{ik}^s)\sum\limits_{t\in\mathcal{T}_j}\gamma_{ik,t}^n\\ \quad\left.+\sum\limits_{(i,k)\in\mathcal{E}^r}(\lambda_{ik}^n+\omega_{ik}^r)\sum\limits_{t\in\mathcal{T}_j}\rho_{ik,t}^n\right.\\ \text{s.t. } (8)-(13),(15)-(18),\boldsymbol{x}\to\boldsymbol{x}^{c,n},\boldsymbol{y}\to\boldsymbol{y}^n,\forall c\in\mathcal{C}\end{cases}\tag{29}$$

Then, we decompose augmented RS-HC into multiple individual period-level subproblems. A subproblem for sub-period j at iteration n is defined as (29), which is formulated by fixing the variables in the other subproblems to the values obtained from the previous iteration.

C. Coordination

Unlike traditional Lagrangian decomposition algorithms, the multipliers in (29) can be updated after solving each subproblem. To make sure that the multipliers are updated in acute angles, the following optimality condition needs to be satisfied:

$$L(\boldsymbol{x}^{c,n},\boldsymbol{y}^n,\boldsymbol{\lambda}^n,\boldsymbol{\nu}^{c,n},\boldsymbol{\upsilon}^{c,n})<L(\boldsymbol{x}^{c,n-1},\boldsymbol{y}^{n-1},\boldsymbol{\lambda}^n,\boldsymbol{\nu}^{c,n},\boldsymbol{\upsilon}^{c,n}),\tag{30}$$

where $L(\boldsymbol{x}^{c,n},\boldsymbol{y}^n,\boldsymbol{\lambda}^n,\boldsymbol{\lambda}^n)$ is the objective value of RS-HC at iteration n. Here we adopt the step-sizing formula developed in [33] and update the multipliers as follows:

$$\begin{aligned}\lambda_{ik}^{n+1}&=[\lambda_{ik}^n+\varsigma^n\tilde{g}_{ik}^n]^+,\ \forall(i,k)\in\mathcal{E}^s\cup\mathcal{E}^r\\ \nu_i^{\varphi,c,n+1}&=\left[\nu_i^{\varphi,c,n}+\varsigma^n\cdot\tilde{g}_{\nu,i}^{\varphi,c,n}\right]^+,\forall i\in\mathcal{N}_d\\ \upsilon_{i,k}^{\varphi,c,n+1}&=\left[\upsilon_{i,k}^{\varphi,c,n}+\varsigma^n\cdot\tilde{g}_{\upsilon,i,k}^{\varphi,c,n}\right]^+,\forall i\in\mathcal{N}_d,\forall k\in\mathcal{T}_d\\ \varsigma^n&=\xi\cdot\frac{\varsigma^{n-1}||\tilde{\boldsymbol{g}}^{n-1}||}{||\tilde{\boldsymbol{g}}^n||},\end{aligned}\tag{31}$$

where ς^n is the stepsize at iteration n, $\xi<1$ is a scaling constant, $[\cdot]^+$ represents projection onto the positive orthant. $\tilde{\boldsymbol{g}}^{\boldsymbol{n}}$ is the surrogate subgradient vector at iteration n, which is given as:

$$\begin{aligned}\tilde{g}_{ik}^n&=\sum_{t\in\mathcal{T}}\gamma_{ik,t}^n-\gamma_{ik}^{\max},\ \forall(i,k)\in\mathcal{E}^s\\ \tilde{g}_{ik}^n&=\sum_{t\in\mathcal{T}}\rho_{ik,t}^n-\rho_{ik}^{\max},\ \forall(i,k)\in\mathcal{E}^r\\ \tilde{g}_{\nu,i}^{\varphi,c,n}&=\sum_{t\in\mathcal{T}}(a_{i,t}^{\varphi,c,n}+b_{i,t}^{\varphi,c,n})-d_1,\forall i\in\mathcal{N}_d\\ \tilde{g}_{\upsilon,i,k}^{\varphi,c,n}&=\sum_{t=k}^{k+d}(a_{i,t}^{\varphi,c,n}+b_{i,t}^{\varphi,c,n})-d_2,\forall i\in\mathcal{N}_d,\forall k\in\mathcal{T}_d.\end{aligned}\tag{32}$$

The key steps of SLR are outlined in Algorithm 1. Since we initialize $\boldsymbol{\gamma}^0,\boldsymbol{\rho}^0,\boldsymbol{a}^0,\boldsymbol{b}^0$ as $\mathbf{1}$, the stopping criterion in line 8 guarantees that the solution obtained by SLR is feasible when the algorithm terminates early.

Algorithm 1 SLR for RS-HC model

Initialize $\boldsymbol{\lambda}^0,\boldsymbol{\nu}^0,\boldsymbol{\upsilon}^0,\boldsymbol{g}^0,\boldsymbol{x}^{c,0},\boldsymbol{y}^0$, multipliers are initialized as small constants, surrogate subgradients are initialized as 0, $\{\boldsymbol{\gamma}^0,\boldsymbol{\rho}^0,\boldsymbol{a}^0,\boldsymbol{b}^0\}=\mathbf{1}$.

1: **for** $n=1,\ldots,N$ **do**
2: For sub-period $j=(n \bmod |\mathcal{J}|)+1$, solve Sub-RS-HC model j and update $\boldsymbol{\lambda}^n,\boldsymbol{\nu}^n,\boldsymbol{\upsilon}^n,\boldsymbol{g}^n,\boldsymbol{x}^{c,n},\boldsymbol{y}^n$.
3: **if** Optimality condition (30) is satisfied **then**
4: Continue iteration.
5: **else**
6: Set $\boldsymbol{x}^{c,n}=\boldsymbol{x}^{c,n-1},\boldsymbol{y}^n=\boldsymbol{y}^{n-1}$
7: **end if**
8: **if** $\boldsymbol{g}^n\le\mathbf{0}$ and $||\Delta\boldsymbol{x}||+||\Delta\boldsymbol{y}||\le 1e^{-5}$ **then**
9: Stop iteration.
10: **end if**
11: **end for**

IV. Case Studies

In this section, we first outline the numerical setup, followed by an evaluation of the proposed GCRN model's performance in DER and load prediction. Next, we compare the solving efficiency of MILP and SLR for the S-HC and RS-HC models and present the operational HC and device operation results for both algorithms. Finally, we discuss the selection and impact of monitoring buses.

A. Numerical Setup

We utilize the modified IEEE 123-bus unbalanced distribution system, as shown in Fig. 2, for our first test system. This system includes 4 voltage regulators, 6 tie switches, and 14 solar PV systems. The parameters of the 123-bus system are detailed in Table I. The value of the big-M constant is empirically tuned and set to $M=5$ to balance computational efficiency and model accuracy. The historical time series data for load and solar PV are sourced from [34], an open-source dataset developed by the National Renewable Energy Laboratory (NREL), which includes data spanning one year.

TABLE I
PARAMETERS OF 123-BUS SYSTEM

Parameter	value	Parameter	value
Sg_i^{φ}	50 KVA	ω^s	$1e^{-4}$
S_{ik}^{φ}	2,000 KVA	ω^r	$1e^{-4}$
$U_i^{\max}$	1.10 p.u.	$M_{i,k}$	5
$U_i^{\min}$	0.90 p.u.	$\gamma_{i,k}^{\max}$	4
$\overline{U}_i$	1.05 p.u.	$\rho_{i,k}^{\max}$	8
$\underline{U}_i$	0.95 p.u.	d_1	8
ϵ	0.06	d_2	4

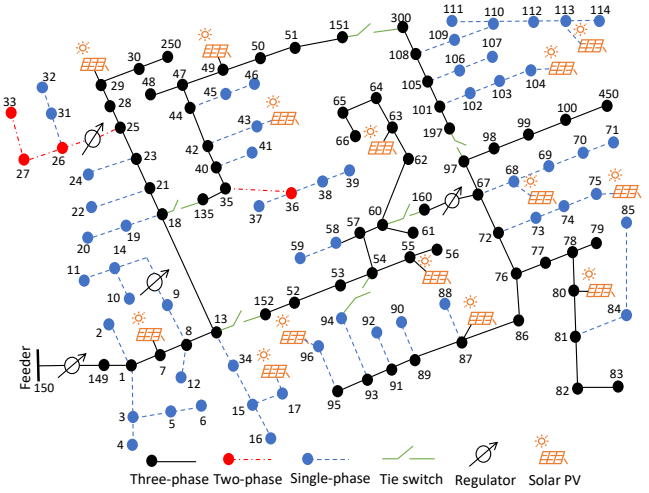


Fig. 2. Modified IEEE 123-bus system.

We further validate our approach using a real-world distribution network from the Midwestern United States, first introduced in [35]. This system consists of three feeders with 240 buses, serving 1,120 customers. Each customer is equipped with a smart meter recording hourly energy consumption (kWh). For this study, we scale the full year of 2017 smart meter data provided in [35]. Since the dataset does not include DERs, we integrate solar PV data from the IEEE 123-bus test system into the 240-bus network. Specifically, solar PV units are installed at buses {1013, 1017, 2025, 2040, 2048, 2058, 3052, 3061, 3067, 3091, 3130, 3135, 3154, 3161}, with system parameters aligned to those in the IEEE 123-bus system. The original network includes one voltage regulator and nine tie switches; to better demonstrate the effectiveness of our algorithm, we add four additional voltage regulators at branches (1008,1009), (2032,2033), (3022,3030), and (3107,3140). Optimization is carried out using Gurobi 12.0.1 [36], with a time limit of one hour for the 123-bus system and two hours for the 240-bus system, and a MIP gap tolerance is set to 0.01% for both.

The proposed GCRN architecture consists of one Graph Convolutional Long Short-Term Memory (GCLSTM) layer, as introduced in [37], followed by two Fully Connected Network (FNN) layers. The corresponding hyperparameters are listed in Table II. We use the mean squared error (MSE) as the loss function. Regularization and early stopping were not applied because preliminary experiments indicated stable convergence without significant overfitting. The model parameters corresponding to the minimum training loss are retained for final evaluation to ensure reliable and consistent performance.

TABLE II
HYPERPARAMETERS OF THE GCRN

Model	GCLSTM	FNN-1	FNN-2
Input dimension	24	64	64
Output dimension	64	64	1
Learning rate	0.001		
Sequence length K	24		
Training epochs	500		

B. Performance of GCRN on solar PV and load prediction

In this subsection, we evaluate the performance of the proposed GCRN in predicting solar PV generation and load demand. To benchmark the GCRN, we use the Gaussian Mixture Model (GMM) introduced in [38]. Two approaches are considered for employing the GMM to characterize uncertainties: one models the uncertainties spatially, and the other models them temporally. These models are referred to as GMM-S and GMM-T, respectively. The probabilistic density functions (PDF) of load demand or DER output for these models are expressed as:

$$\begin{cases} \text{GMM-S} : \text{PDF}_t(\vartheta_t) = \sum_{i=1}^{K} \phi_i \mathbf{N}(\vartheta_t | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \\ \text{GMM-T} : \text{PDF}_i(\vartheta_i) = \sum_{i=1}^{K} \phi_i \mathbf{N}(\vartheta_i | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \end{cases} \tag{33}$$

where K is the number of Gaussian components, ϕ_i represents the mixing coefficient for the i-th component, satisfying $\sum_{i=1}^{K} \phi_i = 1$, $\mathbf{N}(\cdot|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ denotes the multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$. In GMM-S, the PDF is modeled spatially by capturing the uncertainty across the network at a single time t. In contrast, GMM-T characterizes temporal uncertainties by capturing variations for a specific bus i over time.

1) IEEE 123-bus system: For the IEEE 123-bus system, we train the models using 90 days of historical data and evaluate their performance over a subsequent 7-day testing period. The average mean squared errors (MSEs) for load demand and solar PV generation predictions during the testing phase are reported in Table III and IV. Notably, the proposed GCRN consistently achieves the lowest MSE under all scenarios, demonstrating its superior predictive accuracy.

TABLE III
AVERAGE MSE OF LOAD DEMAND PREDICTION ON 123-BUS SYSTEM

Models	$\lvert\mathcal{C}\rvert = 1$	$\lvert\mathcal{C}\rvert = 3$	$\lvert\mathcal{C}\rvert = 5$	$\lvert\mathcal{C}\rvert = 10$	$\lvert\mathcal{C}\rvert = 20$
GMM-S	0.00181	0.00333	0.00314	0.00296	0.00311
GMM-T	0.00305	0.00306	0.00303	0.00280	0.00276
GCRN	**0.000316**	**0.000319**	**0.000320**	**0.000334**	**0.000327**

TABLE IV
AVERAGE MSE OF SOLAR PV PREDICTION ON 123-BUS SYSTEM

Models	$\lvert\mathcal{C}\rvert = 1$	$\lvert\mathcal{C}\rvert = 3$	$\lvert\mathcal{C}\rvert = 5$	$\lvert\mathcal{C}\rvert = 10$	$\lvert\mathcal{C}\rvert = 20$
GMM-S	0.0127	0.0161	0.0124	0.0108	0.0115
GMM-T	0.0233	0.0141	0.0141	0.0128	0.0125
GCRN	**0.000812**	**0.000797**	**0.000802**	**0.000800**	**0.000799**

The complexity of the S-HC problem scales linearly with the number of scenarios. To optimize computation time, it is essential to select a small set of "representative scenarios". As shown in Tables III and IV, the performance of the proposed GCRN remains stable across different scenario counts. Therefore, we select $|C| = 3$ to balance computational efficiency and representativeness.

Then, we present the daily solar PV generation and load demand shape profiles for bus 29, along with the GCRN-predicted scenarios, in Fig. 3. The figure illustrates that the scenarios generated by GCRN effectively capture the actual variations in solar PV generation and load demand.

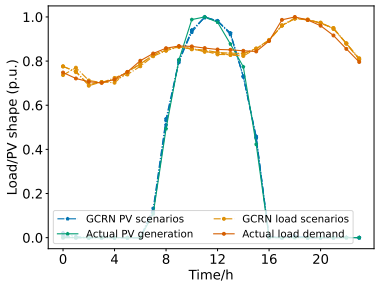


Fig. 3. The daily solar PV generation and load demand with prediction scenarios on the 123-bus system.

2) Real-world 240-bus system: For the real-world 240-bus system, the models are also trained on 90 days of historical data and evaluated over a subsequent 7-day testing period. The average MSEs for active and reactive load predictions during testing are reported in Tables V and VI, respectively. Unlike the 123-bus system, we assume that the active and reactive loads follow distinct load profiles. Once again, the proposed GCRN consistently achieves the lowest MSE across all scenarios, underscoring its superior predictive accuracy.

TABLE V
AVERAGE MSE OF ACTIVE LOAD PREDICTION ON 240-BUS SYSTEM

Models	$\|\mathcal{C}\| = 1$	$\|\mathcal{C}\| = 3$	$\|\mathcal{C}\| = 5$	$\|\mathcal{C}\| = 10$	$\|\mathcal{C}\| = 20$
GMM-S	0.0213	0.0212	0.0224	0.0226	0.0230
GMM-T	0.0200	0.0216	0.0231	0.0223	0.0228
GCRN	**0.00442**	**0.00440**	**0.00441**	**0.00443**	**0.00441**

TABLE VI
AVERAGE MSE OF REACTIVE LOAD PREDICTION ON 240-BUS SYSTEM

Models	$\|\mathcal{C}\| = 1$	$\|\mathcal{C}\| = 3$	$\|\mathcal{C}\| = 5$	$\|\mathcal{C}\| = 10$	$\|\mathcal{C}\| = 20$
GMM-S	0.0229	0.0221	0.0231	0.0227	0.0234
GMM-T	0.0218	0.0226	0.0229	0.0226	0.0235
GCRN	**0.00735**	**0.00735**	**0.00735**	**0.00736**	**0.00736**

C. Computation efficiency of MILP and SLR

1) IEEE 123-bus system: In this subsection, we compare the computation efficiency of the MILP and SLR methods in solving the S-HC and RS-HC models for the first testing day. Given the radial structure of the test feeder, we empirically select downstream buses that are far from the substation as monitor buses. Specifically, we choose buses 85 and 114. Fig. 4 illustrates the evolution of the objective values for both approaches throughout the solving process. For MILP, the upper and lower bounds, as reported by Gurobi, are depicted. For SLR, we evaluate its robustness by varying the sub-horizon length $|\mathcal{T}_j|$ across four values: 1, 2, 3, and 4. To ensure clarity, we denote SLR with $|\mathcal{T}_j| = 1$ as SLR-1 and collectively refer to SLR-1, SLR-2, SLR-3, and SLR-4 as SLRs. In the case of SLRs, each marker represents the objective value at each iteration, whereas for MILP, each marker corresponds to the objective value reported by Gurobi.

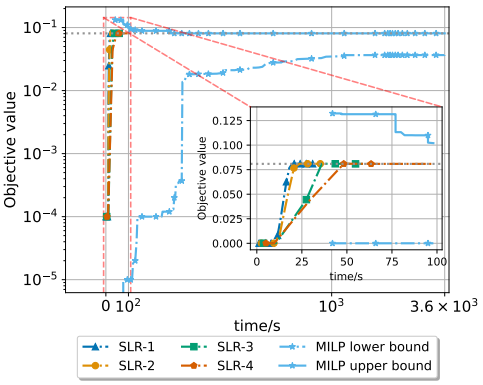


Fig. 4. Progress on the upper and lower bounds for solving the 24-hour S-HC and RS-HC.

From Fig. 4, we observe that SLRs quickly converge to the nearly optimal objective value in a significantly shorter time compared to MILP. Although MILP identifies the best upper bound in approximately 300 s, the presence of a large constant M in numerous constraints weakens the LP relaxation, making it considerably less effective. As a result, MILP requires a much longer time to reach the best lower bound. Additional experiments confirm that, given sufficient time, the MILP lower bound eventually reaches the current best upper bound.

2) Real-world 240-bus system: For the real-world 240-bus system, we empirically select downstream buses {1015, 2060, 3052, 3139} as monitoring points. Fig. 5 depicts the evolution of the objective values for both solution approaches over the course of the solving process. In this large-scale system, the MILP solver identifies its first feasible solution at approximately 5,700 seconds, yet the MIP gap remains substantial within the 1-hour time limit. In contrast, the SLR methods rapidly converge to sub-optimal solutions that are close to the MILP's performance, as illustrated in the figure.

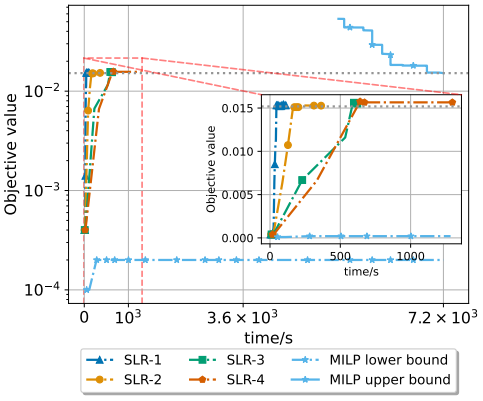


Fig. 5. Progress on the upper and lower bounds for solving the 24-hour S-HC and RS-HC.

D. Operational HC and device operation results

1) IEEE 123-bus system: The operational HC of solar PV at selected buses, along with the total value at $t = 14h$, as

obtained by MILP and SLRs, is presented in Table VII. The results indicate that SLRs achieve the same operational HC as MILP for all solar PV systems.

TABLE VII
PV OPERATIONAL HC OF DIFFERENT BUSES AT $t = 14h$ (KW)

Algorithm	bus 7	bus 29	bus 63	bus 80	bus 104	Total
MILP	43.37	46.24	36.88	36.48	46.26	565.89
SLRs	43.37	46.24	36.88	36.48	46.26	565.89

Next, we present the switch reconfiguration and tap setting results for MILP and SLR-4 in Fig. 6 and Fig. 7, respectively. Switch IDs 1 to 6 correspond to the switches located between the following bus pairs: (13, 152), (18, 135), (60, 160), (54, 94), (151, 300), and (97, 197). To maintain the radiality of the initial system topology, the six switches are initially set to [1,1,1,0,0,1]. The switch reconfiguration results indicate that both MILP and SLR-4 require two operations on switches 1 and 5, but at different time steps.

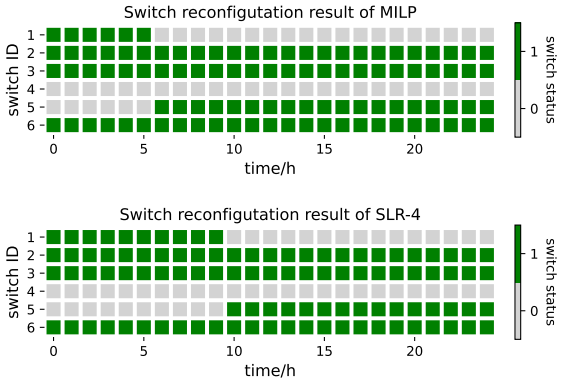


Fig. 6. Switch reconfiguration results of SLR-4 and MILP.

The initial tap positions are set to 0 for all voltage regulators. VR IDs 1 to 4 correspond to the voltage regulators located between the bus pairs (150, 149), (9, 14), (25, 26), and (160, 67). The results show that SLR-4 requires more tap-changing operations than MILP, with three operations for SLR-4 compared to two operations for MILP. Given that we use small operation cost coefficients ($1e^{-4}$), the objective value obtained by SLR-4 can be considered approximately optimal.

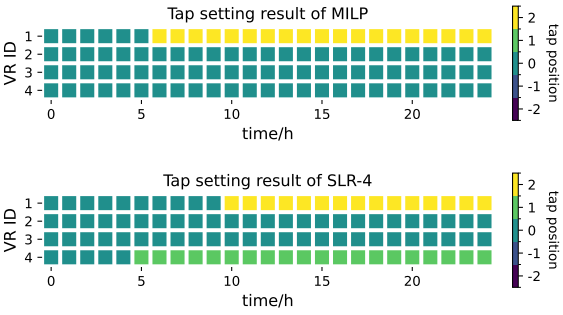


Fig. 7. Tap setting results of SLR-4 and MILP.

Fig. 8 and 9 depict the voltage profiles at monitor buses 85 and 114, respectively. These results reveal that, across all evaluated algorithms, instances of voltage violations are both infrequent and transient—persisting for no more than four consecutive time steps and cumulatively not exceeding eight intervals over the simulation horizon. This indicates a high level of compliance with operational constraints, underscoring the effectiveness of the proposed voltage violation duration constraint in maintaining voltage stability.

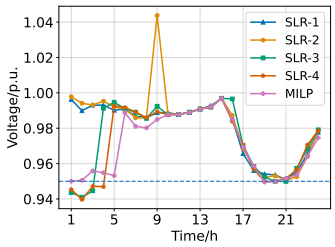


Fig. 8. Voltage profiles of phase C on bus 85 in scenario 1

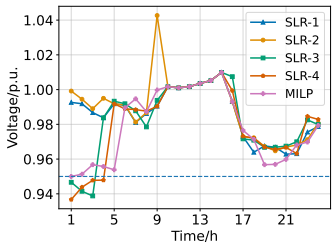


Fig. 9. Voltage profiles of phase A on bus 114 in scenario 1

2) Real-world 240-bus system: Table VIII presents the operational hosting capacity (HC) of solar PV at selected buses, along with the total capacity at $t = 14$ h, as obtained by the MILP and SLR approaches. For this larger real-world system, the results show that the SLR methods can achieve higher operational HC than the MILP at certain PV buses, as well as a greater total HC overall.

TABLE VIII
PV OPERATIONAL HC OF DIFFERENT BUSES AT $t = 14h$ (KW)

Algorithm	bus 2040	bus 2058	bus 3052	bus 3130	Total
MILP	29.85	20.65	25.20	20.14	416.78
SLRs	**34.73**	**23.64**	**31.55**	**32.83**	**443.69**

We then present the switch reconfiguration and tap setting results for both the MILP and SLR-2 approaches in Fig. 10 and Fig. 11, respectively. Switch IDs 1 through 9 correspond to the switches located between the following bus pairs: (1, 1001), (1010, 2057), (1, 2001), (2012, 2013), (2021, 2026), (2013, 3005), (1, 3001), (3075, 3076), and (3081, 2016). To preserve the radiality of the initial system topology, the nine switches are initialized with the configuration [1, 0, 1, 1, 1, 0, 1, 1, 0]. The reconfiguration results reveal that both MILP and SLR-2 require a total of 13 switching operations, albeit occurring at different time steps.

Consistent with the 123-bus system, the initial tap positions of all voltage regulators in the 240-bus network are set to zero. The five voltage regulators, denoted as VR IDs 1–5, are installed at bus pairs (0, 1), (1008, 1009), (2032, 2033), (3022, 3030), and (3107, 3140). Simulation results demonstrate that

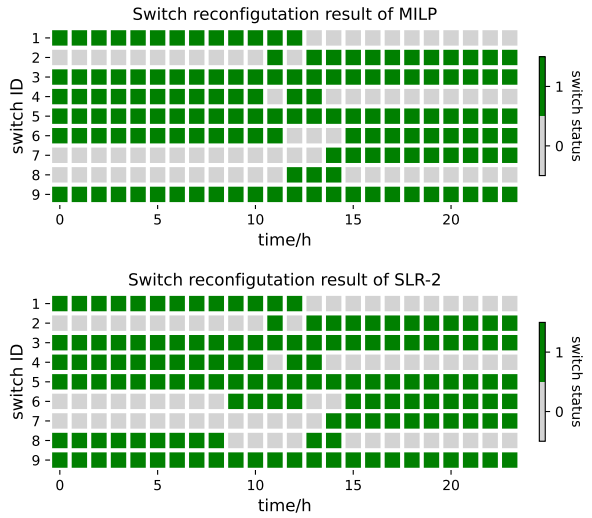


Fig. 10. Switch reconfiguration results of SLR-2 and MILP.

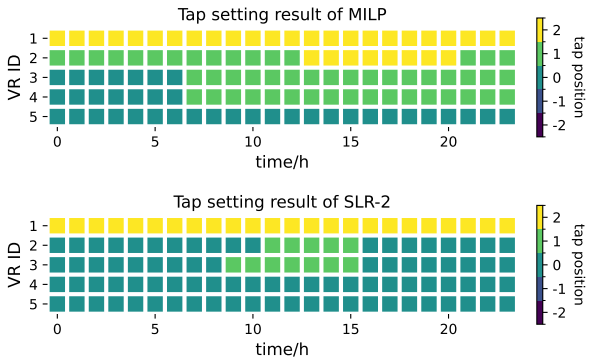


Fig. 11. Tap setting results of SLR-2 and MILP.

the MILP formulation requires a total of six tap-changing operations, whereas the SLR-2 approach achieves comparable performance with only four operations. This reduction highlights the improved operational efficiency of SLR-2 relative to MILP in regulating voltage within large-scale distribution systems.

Fig. 12 and 13 present the voltage profiles at monitoring buses 2060 and 3139, respectively. Consistent with the observations from the 123-bus system, voltage violations across all evaluated algorithms are infrequent and short-lived—persisting for no more than four consecutive time steps and occurring in fewer than eight intervals over the entire simulation horizon.

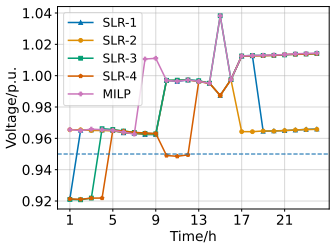


Fig. 12. Voltage profiles of phase B on bus 2060 in scenario 1

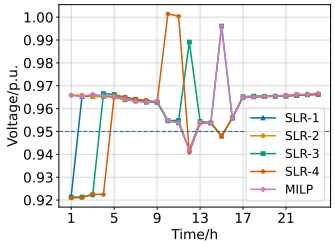


Fig. 13. Voltage profiles of phase A on bus 3139 in scenario 1

E. Selection and impact of monitor buses

In Section IV-C, we empirically select three downstream buses as monitor buses on the 123-bus system. However, this approach may lead to voltage violations at other buses, potentially exceeding the voltage violation duration limits. To ensure that the voltages at all buses comply with the voltage violation duration limit, we can manually identify buses that exceed this constraint. A subset of these violated buses can then be empirically selected and added to the monitor bus set $\mathcal{N}_d$.

Table IX shows that SLR-4 and MILP results in 65 and 54 buses violating the voltage duration limit when using the initial monitor set $\mathcal{N}_d = \{85, 114\}$, respectively. To address this, we empirically select seven additional buses from the violated set and add them to $\mathcal{N}_d$. With the updated monitor set, we resolve the S-HC and RS-HC models and find that all buses now remain within the specified limits. Note that while the total solar PV operational hosting capacity across all time periods remains unchanged under both the initial and updated monitor sets, the timing and frequency of switch and voltage regulator operations differ between the two cases. Additionally, for MILP with the updated monitor set, Gurobi terminates at $3,600$ seconds with a 75.9% MIP gap, once again due to the weak LP relaxation caused by the increased number of big-M constraints. Finally, when selecting all buses, except for the substation bus, as the monitor set, MILP fails to find the optimal solution within $3,600$ seconds. In contrast, SLR-4 successfully finds a nearly optimal solution in approximately 15 minutes. In terms of optimality gap, the MILP achieves a smaller gap than the SLR when using the initial monitor set. However, as the monitoring set expands to $\mathcal{N}_d = \mathcal{N}/\mathcal{S}$, the MILP is unable to obtain a high-quality solution within $3,600$ seconds, whereas the SLR method achieves a tighter optimality gap under the same computational limit.

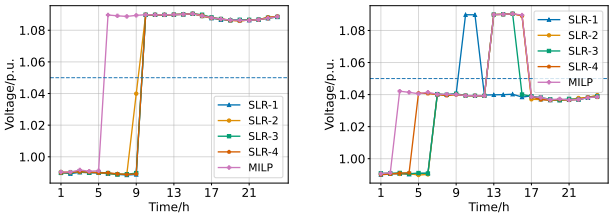
Fig. 14 and 15 depict the voltage profiles at buses 1 and 96 under both the initial and updated monitor sets. With the updated monitor set, the voltages are effectively reduced to comply with the voltage violation limits.

F. Sensitivity analysis of key parameters

In this subsection, we perform a sensitivity analysis of key parameters on the 123-bus system. We first examine the impact of the cost coefficients ω^s and ω^r, while fixing the voltage violation duration limits at $d_1 = 4$ and $d_2 = 8$. Specifically, we assess how varying these parameters influences the operational PV hosting capacity (HC) and device operation frequency.

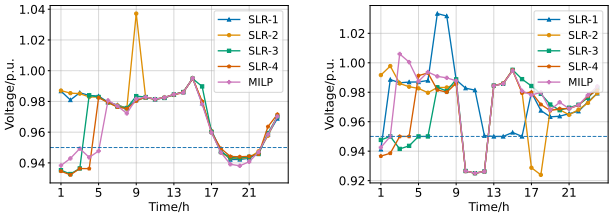
TABLE IX
PERFORMANCE COMPARISON UNDER DIFFERENT MONITOR SETS $\mathcal{N}_d$ ON 123-BUS SYSTEM

Algorithm	$\mathcal{N}_d$	Number of integer variables	Number of violate buses	Total operational PV HC (KW)	Total device operation times	Solve time (s)	MIP gap (%)	Optimality gap (%)
MILP	{85,114}	13,824	**54**	4152.22	**4**	3,600	54.60	0.00
	{39,66,75,83, 85,94,96,114, 149,160,300,450}	18,144	0	4152.22	7	3,600	75.90	0.37
	$\mathcal{N}/\mathcal{S}$	67,392	0	4118.05	13	3,600	100	9.57
SLR-4	{85,114}	13,824	65	4152.22	5	**97**	N/A	0.12
	{39,66,75,83, 85,94,96,114, 149,160,300,450}	18,144	0	4152.22	7	**383**	N/A	0.37
	$\mathcal{N}/\mathcal{S}$	67,392	0	**4152.22**	**9**	**896**	N/A	0.62



(a) Use original monitor set (b) Use updated monitor set

Fig. 14. Voltage profiles of phase A on bus 1 in scenario 1.



(a) Use original monitor set (b) Use updated monitor set

Fig. 15. Voltage profiles of phase B on bus 96 in scenario 1.

Table X summarizes the total PV HC, device operations, and solution times under different settings. The results indicate that larger values of ω^s and ω^r not only lead to fewer device operations but also reduce the achievable PV hosting capacity. Across all scenarios, the proposed SLR-4 consistently matches or outperforms the MILP benchmark. In terms of hosting capacity, SLR-4 yields equal or higher values than MILP under small to moderate cost coefficients. More importantly, SLR-4 demonstrates a clear advantage in computational efficiency: while MILP consistently hits the 3600-second time limit, SLR-4 converges within 1,000 seconds in every case, and achieves its fastest solution in just 70 seconds at $\omega^s, \omega^r = 0.1$. Overall, these findings underscore the scalability and efficiency of the proposed SLR-4 method, which achieves high-quality solutions while dramatically reducing computation time.

Next, we fix $\omega^s = \omega^r = 0.0001$ and examine the impact of varying d_1 and d_2. As shown in Table XI, smaller values of d_1 and d_2 force PV inverters to supply additional reactive power for voltage support, thereby reducing the total operational hosting capacity of solar PV. In contrast, relaxing the

TABLE X
TOTAL OPERATIONAL PV HC AND DEVICE OPERATION TIMES UNDER DIFFERENT ω^s, ω^r ON 123-BUS SYSTEM

ω^s, ω^r	Total operational PV HC (KW)		Total device operation times		Solve time (s)	
	MILP	SLR-4	MILP	SLR-4	MILP	SLR-4
0.0001	4118.05	**4152.22**	13	**9**	3600	**896**
0.001	4090.92	**4113.61**	**7**	11	3600	**712**
0.01	4015.29	**4016.13**	5	5	3600	**348**
0.1	3897.78	3897.78	1	1	3600	**70**

duration limits from $(d_1, d_2) = (2, 4)$ to $(12, 16)$ allows the system to accommodate higher PV hosting capacity. Across all scenarios, the proposed SLR-4 consistently achieves higher or comparable PV HC relative to MILP, while requiring fewer device operations. For example, at $(d_1, d_2) = (8, 12)$, SLR-4 achieves 4211.10 kW with only six operations, compared to 4199.90 kW and thirteen operations under MILP. Moreover, SLR-4 demonstrates a clear computational advantage, converging within 1100 seconds in all cases, whereas MILP consistently reaches the 3600-second time limit. Notably, for $(d_1, d_2) = (2, 4)$, MILP fails to find a feasible solution within the time limit. These results highlight the effectiveness of the proposed method in enhancing PV integration while maintaining computational efficiency and reducing device operations.

TABLE XI
TOTAL OPERATIONAL PV HC AND DEVICE OPERATION TIMES UNDER DIFFERENT d_1, d_2 ON 123-BUS SYSTEM

d_1, d_2	Total operational PV HC (KW)		Total device operation times		Solve time (s)	
	MILP	SLR-4	MILP	SLR-4	MILP	SLR-4
2, 4	-	**4059.54**	-	**7**	3600	**547**
4, 8	4118.05	**4152.22**	13	**9**	3600	**896**
8, 12	4199.90	**4211.10**	13	**6**	3600	**1057**
12, 16	4244.06	**4269.92**	10	**6**	3600	**870**

V. CONCLUSION

This paper presents an advanced operational hosting capacity (HC) optimization framework to address the limitations of traditional static HC methods and Dynamic Operating Envelopes (DOEs). We develop a stochastic HC model incorporating a novel voltage violation duration constraint and pro-

pose a Surrogate Lagrangian Relaxation (SLR)-based temporal decomposition scheme to improve computational efficiency. Compared with other decomposition schemes such as Benders Decomposition and Progressive Hedging, the SLR method relaxes coupling constraints using surrogate multipliers and updates them through an aggregated subgradient approach. This formulation requires fewer iterations and eliminates the need for solving a master problem. Numerical results show a significant speedup while maintaining operational HC, highlighting the proposed framework's efficiency and scalability. A limitation of the proposed SLR method is its sensitivity to parameter selection, particularly the choice of stepsize, which may require careful tuning in certain cases. A promising direction for future research is the development of adaptive schemes, such as surrogate Level-Based Lagrangian Relaxation (SLBLR), to enhance robustness and accelerate convergence. Another direction is to investigate decision-focused learning frameworks that jointly optimize GCRN and SLR, allowing uncertainty information to be directly propagated into the decision-making process. Furthermore, the current formulation of the voltage violation duration constraint may be relatively loose, suggesting the need for tighter and more effective formulations. Finally, the proposed framework could be extended to broader applications, including Transmission System Operator–Distribution System Operator coordination and VVC.

REFERENCES

[1] R. Torquato, D. Salles, C. O. Pereira, P. C. M. Meira, and W. Freitas, "A comprehensive assessment of PV hosting capacity on low-voltage distribution systems," *IEEE Trans. Power Del.*, vol. 33, no. 2, pp. 1002–1012, 2018.

[2] "Open Energy Networks Project - Energy Networks Australia Position Paper," ENA and AEMO, Melbourne, Australia, Tech. Rep., 2020. [Online]. Available: https://tinyurl.com/4bd9tr5n

[3] B. Liu and J. H. Braslavsky, "Robust dynamic operating envelopes for der integration in unbalanced distribution networks," *IEEE Trans. Power Syst.*, vol. 39, no. 2, pp. 3921–3936, 2023.

[4] "Managing Flexibility on the Distribution System with Dynamic Operating Envelopes," Electric Power Research Institute, Palo Alto, California, Tech. Rep. 3002025065, 2022. [Online]. Available: https://www.epri.com/research/programs/108271/results/3002025065

[5] K. Petrou, A. T. Procopiou, L. Gutierrez-Lagos, M. Z. Liu, L. F. Ochoa, T. Langstaff, and J. M. Theunissen, "Ensuring distribution network integrity using dynamic operating limits for prosumers," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 3877–3888, 2021.

[6] M. Z. Liu, L. F. Ochoa, P. K. Wong, and J. Theunissen, "Using OPF-based operating envelopes to facilitate residential DER services," *IEEE Trans. Smart Grid*, vol. 13, no. 6, pp. 4494–4504, 2022.

[7] M. Mahmoodi, L. Blackhall, S. M. N. RA, A. Attarha, B. Weise, and A. Bhardwaj, "Der capacity assessment of active distribution systems using dynamic operating envelopes," *IEEE Trans. Smart Grid*, vol. 15, no. 2, pp. 1778–1791, 2023.

[8] M. T. Islam, M. J. Hossain, M. A. Habib, and M. A. Zamee, "Adaptive hosting capacity forecasting in distribution networks with distributed energy resources," *Energies*, vol. 18, no. 2, p. 263, 2025.

[9] B. L. M. Santos, L. S. Barros, F. A. Moreira, and D. Barbosa, "Hosting capacity maximization of distributed energy resources through simultaneous optimized volt–var curve and network reconfiguration," *Electr. Power Syst. Res.*, vol. 234, p. 110743, 2024.

[10] P.-H. Trinh and I.-Y. Chung, "Integrated active and reactive power control methods for distributed energy resources in distribution systems for enhancing hosting capacity," *Energies*, vol. 17, no. 7, p. 1642, 2024.

[11] S. Zhang, Y. Fang, H. Zhang, H. Cheng, and X. Wang, "Maximum hosting capacity of photovoltaic generation in SOP-based power distribution network integrated with electric vehicles," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 8213–8224, 2022.

[12] *Electric Power Systems and Equipment - Voltage Ratings (60 Hz)*, American National Standards Institute (ANSI) Std. C84.1-2016, 2016.

[13] A. K. Jain, K. Horowitz, F. Ding, K. S. Sedzro, B. Palmintier, B. Mather, and H. Jain, "Dynamic hosting capacity analysis for distributed photovoltaic resources—Framework and case study," *Appl. Energy*, vol. 280, p. 115633, 2020.

[14] *IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces*, IEEE Standards Association Std. IEEE 1547-2018, 2018.

[15] H. Wu, P. Dong, and M. Liu, "Distribution network reconfiguration for loss reduction and voltage stability with random fuzzy uncertainties of renewable energy generation and load," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 5655–5666, 2018.

[16] E. Kianmehr, S. Nikkhah, V. Vahidinasab, D. Giaouris, and P. C. Taylor, "A resilience-based architecture for joint distributed energy resources allocation and hourly network reconfiguration," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5444–5455, 2019.

[17] S. N. Salih and P. Chen, "On coordinated control of oltc and reactive power compensation for voltage regulation in distribution systems with wind power," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4026–4035, 2015.

[18] S. Hashemi, J. Østergaard, T. Degner, R. Brandl, and W. Heckmann, "Efficient control of active transformers for increasing the PV hosting capacity of LV grids," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 270–277, 2016.

[19] X. Chen, W. Wu, and B. Zhang, "Robust capacity assessment of distributed generation in unbalanced distribution networks incorporating ANM techniques," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 651–663, 2017.

[20] J. Qin, R. Yang, and N. Yu, "Physics-informed graph neural networks for collaborative dynamic reconfiguration and voltage regulation in unbalanced distribution systems," *IEEE Trans. Ind. Appl.*, pp. 1–11, 2025.

[21] D. Gebbran, S. Mhanna, A. C. Chapman, and G. Verbič, "Multiperiod DER coordination using ADMM-based three-block distributed AC optimal power flow considering inverter Volt-Var control," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 2874–2889, 2022.

[22] Y. Ju, Z. Zhang, W. Wu, W. Liu, and R. Zhang, "A bi-level consensus ADMM-based fully distributed inverter-based Volt/Var control method for active distribution networks," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 476–487, 2021.

[23] T. Xu and W. Wu, "Accelerated ADMM-based fully distributed inverter-based Volt/Var control strategy for active distribution networks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7532–7543, 2020.

[24] H. Al-Saadi, R. Zivanovic, and S. F. Al-Sarawi, "Probabilistic hosting capacity for active distribution networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 5, pp. 2519–2532, 2017.

[25] S. F. Santos, D. Z. Fitiwi, M. Shafie-Khah, A. W. Bizuayehu, C. M. Cabrita, and J. P. Catalão, "New multistage and stochastic mathematical model for maximizing RES hosting capacity—Part I: Problem formulation," *IEEE Trans. Sustain. Energy*, vol. 8, no. 1, pp. 304–319, 2016.

[26] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A review of graph neural networks and their applications in power systems," *J. Mod. Power Syst. Clean Energy.*, vol. 10, no. 2, pp. 345–360, 2021.

[27] M. A. Bragin, "Survey on Lagrangian relaxation for MILP: importance, challenges, historical review, recent advancements, and opportunities," *Ann. Oper. Res.*, vol. 333, no. 1, pp. 29–45, 2024.

[28] M. A. Bragin and E. L. Tucker, "Surrogate "level-based" Lagrangian relaxation for mixed-integer linear programming," *Sci. Rep.*, vol. 12, no. 1, p. 22417, 2022.

[29] C. Ben Hammouda, E. Rezvanova, E. von Schwerin, and R. Tempone, "Lagrangian relaxation for continuous-time optimal control of coupled hydrothermal power systems including storage capacity and a cascade of hydropower systems with time delays," *Optim. Control Appl. Methods*, vol. 45, no. 5, pp. 2279–2311, 2024.

[30] B. A. Robbins and A. D. Domínguez-García, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903–2913, 2015.

[31] B. A. Robbins, H. Zhu, and A. D. Domínguez-García, "Optimal tap setting of voltage regulation transformers in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 256–267, 2015.

[32] X. Chen, W. Wu, and B. Zhang, "Robust restoration method for active distribution networks," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4005–4015, 2015.

[33] M. A. Bragin, P. B. Luh, J. H. Yan, N. Yu, and G. A. Stern, "Convergence of the surrogate lagrangian relaxation method," *J. Optim. Theory Appl.*, vol. 164, pp. 173–201, 2015.

[34] T. Elgindy and K. Balasubramaniam, "Sample IEEE123 Bus system for OEDI SI," 09 2022. [Online]. Available: https://data.openei.org/submissions/5773

[35] F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, "A time-series distribution test system based on real utility data," in *North American Power Symposium (NAPS)*. IEEE, 2019, pp. 1–6.

[36] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2024. [Online]. Available: https://www.gurobi.com

[37] Y. Seo, M. Defferrard, P. Vandergheynst, and X. Bresson, "Structured sequence modeling with graph convolutional recurrent networks," in *Int. Conf. Neural Inf. Process.* Springer, 2018, pp. 362–373.

[38] Y. Yang, W. Wu, B. Wang, and M. Li, "Analytical reformulation for stochastic unit commitment considering wind power uncertainty with gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 2769–2782, 2019.