

Ramping-aware Enhanced Flexibility Aggregation of Distributed Generation with Energy Storage in Power Distribution Networks

Hyeongon Park, Daniel K. Molzahn, and Rahul K. Gupta

Abstract—Power distribution networks are increasingly hosting controllable and flexible distributed energy resources (DERs) that, when aggregated, can provide ancillary support to transmission systems. However, existing aggregation schemes often ignore the ramping constraints of these DERs, which can render them impractical in real deployments. This work proposes a ramping-aware flexibility aggregation scheme, computed at the transmission–distribution boundary, that explicitly accounts for DER ramp limits and yields flexibility envelopes that are provably disaggregable. To further enhance the attainable flexibility region, we introduce a novel “pre-ramping” strategy, which proactively adjusts resource operating points to enlarge the aggregated flexibility envelope while preserving both network feasibility and disaggregation guarantees. The proposed method demonstrates a 5.2% to 19.2% improvement in flexibility relative to the baseline model, depending on system conditions. We validate the scheme on an IEEE-33 bus distribution system and provide formal proofs showing that both aggregation strategies are disaggregable for all feasible trajectories within the aggregate flexibility envelope.

Index Terms—Aggregate flexibility, ramping limits, distributed energy resources, flexibility envelopes, distribution networks

NOMENCLATURE

Notation: Superscripts \wedge and \vee denote variables associated with the upper and lower envelope trajectories, respectively. The superscript “ \bullet ” indicates quantities evaluated at the pre-ramped operating point.

Indices:

- $t \in \mathcal{T}$ Time index.
- $g \in \mathcal{G}$ Generator index.
- $e \in \mathcal{E}$ Energy storage system (ESS) index.
- $n \in \mathcal{N}$ Node index.
- $\ell \in \mathcal{L}$ Line index.
- s Envelope index, $s \in \{\wedge, \vee\}$.

Parameters:

- Δt Duration of a time step.
- P_g^{\min}, P_g^{\max} Active power limits of generator g .
- Q_g^{\min}, Q_g^{\max} Reactive power limits of generator g .
- $R_g^{\uparrow}, R_g^{\downarrow}$ Ramp-up/down limits of generator g per time step.
- P_e^{\max} Maximum charging/discharging power output of ESS e .
- E_e^{\min}, E_e^{\max} State-of-charge (SoC) limits of ESS e .
- E_e^{init} Initial SoC of ESS e .

- P_g^{init} Initial active power output of generator g .
- κ_e Energy efficiency factor of ESS e .
- $p_t^{\text{load}}, q_t^{\text{load}}$ Nodal active and reactive power demand.
- p_t^{PV} Nodal PV generation.
- \underline{u}, \bar{u} Lower and upper bounds on squared nodal voltages.

Decision Variables:

- $p_t^{\text{GCP},s}$ Net active power exchanged at the grid connection point (GCP) for envelope s .
- $p_t^{\text{GCP},s\bullet}$ Pre-ramped net active power at the GCP for envelope s .
- $p_{g,t}^s, q_{g,t}^s$ Active and reactive power outputs of generator g for envelope s .
- $p_{e,t}^s$ Active power output of ESS e for envelope s (positive for discharging).
- $e_{e,t}^s$ SoC of ESS e for envelope s .
- $p_{g,t}^{\text{pre},s}$ Generator pre-ramp variable for envelope s .
- $p_{e,t}^{\text{pre},s}$ ESS pre-ramp variable for envelope s .
- $e_{e,t}^{s\bullet}$ Pre-ramped SoC of ESS e for envelope s .
- u_t^s Nodal squared voltages for envelope s .
- $u_t^{s\bullet}$ Pre-ramped nodal squared voltages for envelope s .
- $p_t^{\text{inj},s}, q_t^{\text{inj},s}$ Nodal active and reactive power injections for envelope s .
- $p_t^{\text{inj},s\bullet}$ Pre-ramped nodal active power injections for envelope s .

I. INTRODUCTION

The increasing penetration of renewable energy resources has significantly heightened the need for operational flexibility in modern power systems to reliably balance variability and uncertainty. This stochastic nature often leads to frequency instability and acute shortages in short-term balancing, frequently triggering price spikes in real-time markets [1]. In response, system operators such as California Independent System Operator (CAISO), Midcontinent Independent System Operator (MISO), and Southwest Power Pool (SPP) have established and operated markets for Flexible Ramping Products (FRPs) [2]. Within the FRP market framework, generation resources are compensated for providing upward and downward ramping capability, enabling them to strategically offer flexibility in exchange for additional revenue. This market-based mechanism incentivizes active participation by generators while ensuring that sufficient ramping capability is available to maintain system reliability under real-time operational uncertainty.

While large-scale generators have traditionally provided these services, there is growing interest in leveraging distribu-

School of Systems Management and Safety Engineering, Pukyong National University, Busan, 48516, South Korea (e-mail: hyeongon@pknu.ac.kr)

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30313, USA (e-mail: molzahn@gatech.edu)

School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, 99163, USA (e-mail: rahul.k.gupta@wsu.edu)

TABLE I: Comparison of Existing Studies on Flexibility Aggregation

Reference	Objective	Grid Model	Explicit Ramp [†]	Pre-ramp	Disaggregation	Remarks
Wen et al. [5]	Exact derivation of temporally coupled aggregate flexibility	Linearized AC / shift-factor model	✗	✗	Guaranteed	Exact but relies on a complex projection framework that is hard to extend for additional constraints.
Chen et al. [6]	Maximization of inner-box flexibility	Linearized multiphase	✗	✗	Guaranteed	Inner-box construction ensures disaggregation.
Öztürk et al. [7]	Constructing an efficient inner approximation of the Minkowski sum	None	✗	✗	Guaranteed	Constructs a scalable inner approximation of storage flexibility using a vertex-based method.
Wen et al. [8]	Flexibility cost/value quantification for TSO-DSO coordination	LinDistFlow	✗	✗	Guaranteed	Uses inner-approximated flexibility to derive time-coupled cost models and marginal prices.
Huang et al. [16]	Minimization of infeasibility via boundary shrinkage	Linearized AC/DC model	✓	✗	Not addressed	Initializes a flexibility region and iteratively shrinks it by removing infeasible trajectories.
Ramping-aware baseline model (Section II)	Maximization of inner-box flexibility	LinDistFlow	✓	✗	Guaranteed	Time-coupled envelope under DER and grid constraints; no pre-ramping.
Pre-ramped enhanced model (Section III)	Maximization of inner-box flexibility	LinDistFlow	✓	✓	Guaranteed	Pre-ramping enlarges the feasible envelope while preserving DER and grid constraints.

[†] “Explicit ramp” denotes direct limits on power change rates (as in generator ramping). DER models with only state-dependent time coupling (e.g., ESS SoC) are state-coupled rather than explicitly ramp-limited.

tion systems as additional sources of flexibility. Distribution systems, once passive, are now emerging as vital flexibility providers due to the proliferation of distributed energy resources (DERs). This growing interest has spurred research into aggregation-based models that represent the collective capability of these diverse assets. Specifically, recent studies by [3], [4] demonstrate that active distribution networks can yield substantial flexibility for the upstream grid. By coordinating and aggregating heterogeneous DERs, including energy storage systems (ESS), electric vehicles (EV), and flexible loads, these networks act as aggregated virtual power plants that support the broader system’s operational needs.

From a mathematical perspective, power aggregation can be viewed as a projection of high-dimensional, device-level operational constraints onto a low-dimensional feasible region of net substation power injections. This projection characterizes the aggregate flexibility that a distribution system can reliably offer to the transmission grid [5]. However, as noted by [6], [7], explicitly computing this feasible region is computationally intractable at scale due to the large number of heterogeneous DERs, temporal coupling across time periods, and network security constraints.

To address this challenge, several recent studies focus on constructing conservative inner approximations of the aggregate feasible region [6]–[10]. Such approximations provide disaggregation guarantees, ensuring that any aggregate power trajectory within the approximated region can be reliably mapped back to a feasible dispatch of individual DERs. For clarity, we call an aggregate trajectory *disaggregable* if any power setpoint within the computed flexibility region can be disaggregated while satisfying all underlying DER-level and network-feasibility constraints over the horizon.

Representative studies include the work by Chen et al. [6], that introduces an inner-approximation framework for aggregating multi-period flexibility in unbalanced distribution systems with disaggregation guarantees. This work is later

extended in [9] using a two-stage adaptive robust optimization framework to provide stronger feasibility guarantees with reduced conservativeness. Building on this line of work on inner approximations, [10] proposes linear programming-based methods to compute inner approximations of the Minkowski sum of EV flexibility sets, while [11] further integrates aggregation with cost-minimization frameworks for non-industrial air-conditioning systems and EVs. As summarized in Table I, existing literature primarily focuses on state-coupled constraints but does not explicitly account for ramp-rate limits, making them difficult to apply to ramp-limited resources.

Meanwhile, a growing number of generation and load-side resources are being integrated into power systems, driven by the expansion of data centers, artificial intelligence workloads, and other electricity-intensive applications [12], [13]. Several of these resources, including smaller-scale or distributed deployments of fuel cells, combined heat and power (CHP) units [14], and small modular reactors (SMRs) [15], exhibit inherent ramp-rate limitations. Therefore, there is need of aggregation schemes, where the ramp-limits are explicitly accounted.

To address this gap, we first develop a *ramping-aware* flexibility aggregation framework that explicitly integrates ramp-rate constraints into the inner-box approximation. While this ensures disaggregation guarantees for ramp-limited resources, the integration of ramping constraints leads to an overly conservative flexibility envelope, as the feasible power range at each interval is constrained by the previous state. To overcome this conservativeness, we propose an *enhanced* formulation that introduces new decision variables, referred to as “pre-ramping” decisions at interval $t - 1$, which proactively reposition resource operating points to expand the feasible power range available at interval t . This approach helps partially decouple the current flexibility setpoint from the previous interval. The pre-ramping variables are optimized to enable delivering maximum ramping capability when required. As summarized in Table I, this mechanism effectively enlarges

the aggregate flexibility envelope while strictly preserving disaggregation guarantees and grid feasibility.

To the best of our knowledge, this is the first grid-aware aggregation framework that account for explicit ramping constraints with provable disaggregation guarantees. While existing work in Huang et al. [16] address ramping through iterative boundary-shrinkage, they do not guarantee disaggregation. Moreover, the proposed pre-ramping scheme enhances the flexibility region compared to the baseline model. This performance gain is quantitatively validated in Section IV.

The main contributions of this paper are listed below.

- 1) We propose a flexibility aggregation framework that explicitly accounts for ramp-rate constraints of resources and guarantees the existence of a feasible unit-level disaggregation for any aggregate trajectory within the constructed envelope. It is referred to as the **Ramping-aware baseline model**.
- 2) We introduce pre-ramping decisions that reposition resource operating points in advance to enlarge the aggregate flexibility envelope, while preserving disaggregation guarantees and network feasibility. It is referred to as the **Pre-ramped enhanced model**.
- 3) We apply the proposed aggregation framework to a FRP market setting and quantitatively evaluate its effectiveness in enhancing practically deliverable flexibility.

The paper is organized as follows: Section II presents the baseline flexibility envelope formulation without pre-ramping. Section III introduces the proposed pre-ramped flexibility envelope model and details its mathematical formulation. Section IV provides numerical case studies to validate the effectiveness of the proposed framework. Finally, Section V concludes the paper and discusses key findings.

II. RAMPING-AWARE FLEXIBILITY AGGREGATION: BASELINE MODEL

We consider a power distribution system with multiple distributed generators (DGs) and ESSs. The objective is to aggregate the combined flexibility from DGs and ESSs at the grid connection point (GCP), i.e., at the transmission-distribution interconnection. We next present an aggregation model that takes ramping constraints into account. It is referred to as baseline model and described below.

The baseline model constructs an inner-box approximation of the aggregate flexibility at the GCP. The model explicitly accounts for resource-level operational constraints and distribution-network constraints, while ensuring temporal feasibility through ramping and time-coupled constraints. Specifically, aggregate flexibility is represented by a pair of upper and lower feasible power trajectories at the GCP over the considered time horizon. These trajectories define a time-coupled flexibility envelope such that any aggregate power trajectory lying between them is disaggregable into a feasible dispatch of individual resources. The proposed formulation aims to determine the largest possible envelope by jointly optimizing the upper and lower trajectories subject to resource-level operating limits, inter-temporal ramping constraints, energy-coupling constraints, and distribution-network feasibility.

A. Optimization Problem

We next present the objective function as well as resource and network constraints.

1) *Objective Function*: The objective maximizes the aggregate flexibility area at the GCP over the considered time horizon. This is achieved by maximizing the difference between the upper and lower feasible power trajectories. Using the symbols from the nomenclature, it is

$$\max \sum_t \left(p_t^{\text{GCP},\wedge} - p_t^{\text{GCP},\vee} \right) \cdot \Delta t. \quad (1)$$

2) *Resources Constraints*: The resource constraints define feasible upper and lower operating trajectories for generators and ESSs. Generator power outputs are constrained within their operating limits for both the upper and lower envelope trajectories. These constraints are enforced as follows:

$$P_g^{\min} \leq p_{g,t}^s \leq P_g^{\max}, \quad \forall g \in \mathcal{G}, t \in \mathcal{T}, s \in \{\wedge, \vee\}, \quad (2a)$$

$$Q_g^{\min} \leq q_{g,t}^s \leq Q_g^{\max}, \quad \forall g \in \mathcal{G}, t \in \mathcal{T}, s \in \{\wedge, \vee\}, \quad (2b)$$

$$p_{g,t}^{\wedge} \geq p_{g,t}^{\vee}, \quad \forall g \in \mathcal{G}, t \in \mathcal{T}. \quad (2c)$$

As discussed earlier, we consider generator ramping constraints between the subsequent time periods are defined as:

$$p_{g,t}^s - p_{g,t-1}^s \leq R_g^{\uparrow}, \quad \forall g \in \mathcal{G}, t \geq 2, s \in \{\wedge, \vee\}, \quad (3a)$$

$$p_{g,t-1}^s - p_{g,t}^s \leq R_g^{\downarrow}, \quad \forall g \in \mathcal{G}, t \geq 2, s \in \{\wedge, \vee\}, \quad (3b)$$

$$p_{g,t}^{\wedge} - p_{g,t-1}^{\vee} \leq R_g^{\uparrow}, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (3c)$$

$$p_{g,t-1}^{\vee} - p_{g,t}^{\wedge} \leq R_g^{\downarrow}, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (3d)$$

$$p_{g,t}^{\vee} - p_{g,t-1}^{\wedge} \leq R_g^{\uparrow}, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (3e)$$

$$p_{g,t-1}^{\wedge} - p_{g,t}^{\vee} \leq R_g^{\downarrow}, \quad \forall g \in \mathcal{G}, t \geq 2. \quad (3f)$$

These ramp constraints are expressed for worst-case cross transitions, i.e., from lower to upper envelopes and vice-versa, as well as for upper and lower envelopes. These constraints ensure that transitions between the upper and lower envelopes remain achievable under finite ramp-rate limits. Specifically, the cross-corner constraints in (3) capture the most restrictive transitions, such as moving from the lower bound at $t-1$ to the upper bound at t . By bounding these worst-case power changes, the formulation guarantees that any interior trajectory can be disaggregated into a feasible resource-level schedule.

Similarly, we include constraints on ESS power outputs:

$$-P_e^{\max} \leq p_{e,t}^s \leq P_e^{\max}, \quad \forall e \in \mathcal{E}, t \in \mathcal{T}, s \in \{\wedge, \vee\}, \quad (4a)$$

$$p_{e,t}^{\wedge} \geq p_{e,t}^{\vee}, \quad \forall e \in \mathcal{E}, t \in \mathcal{T}. \quad (4b)$$

Then, the state-of-charge (SoC) dynamics and the constraints on the SoC are expressed as

$$e_{e,t}^s = \kappa_e \cdot e_{e,t-1}^s - p_{e,t-1}^s \cdot \Delta t, \quad \forall e \in \mathcal{E}, t \geq 2, s \in \{\wedge, \vee\}, \quad (4c)$$

$$e_{e,t}^{\wedge} \geq E_e^{\text{init}} - \Delta t \sum_{\tau=1}^{t-1} p_{e,\tau}^{\vee}, \quad \forall e \in \mathcal{E}, t \geq 2, \quad (4d)$$

$$e_{e,t}^{\vee} \leq E_e^{\text{init}} - \Delta t \sum_{\tau=1}^{t-1} p_{e,\tau}^{\wedge}, \quad \forall e \in \mathcal{E}, t \geq 2, \quad (4e)$$

$$E_e^{\min} \leq e_{e,t}^s \leq E_e^{\max}, \quad \forall e \in \mathcal{E}, t \in \mathcal{T}, s \in \{\wedge, \vee\}. \quad (4f)$$

In (4c), κ_e denotes the storage efficiency factor that models energy loss over time. We assume ideal charging and discharging efficiency,¹ i.e., $\kappa_e = 1$, as in [6], [9], to simplify the disaggregation analysis. Constraints (4d) and (4e) ensure that, even if the ESS follows the most energy-accumulating (for $e_{e,t}^\wedge$) or energy-depleting (for $e_{e,t}^\vee$) trajectory up to time t , the SoC remains within energy limits determined by its initial energy state. Equation (4f) ensures that the SoC stays within its admissible energy limits.

Finally, the initial conditions for the upper and lower trajectories are given by:

$$p_{g,0}^\wedge = p_{g,0}^\vee = P_g^{\text{init}}, \quad \forall g \in \mathcal{G}, \quad (5a)$$

$$e_{e,0}^\wedge = e_{e,0}^\vee = E_e^{\text{init}}, \quad \forall e \in \mathcal{E}, \quad (5b)$$

$$p_{e,0}^\wedge = p_{e,0}^\vee = 0, \quad \forall e \in \mathcal{E}. \quad (5c)$$

3) *Network Constraints:* We model the network constraints using the linearized DistFlow model (LinDistFlow) [18].

Let $B \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{L}|}$ denote the node-branch incidence matrix, $\mathbf{r}, \mathbf{x} \in \mathbb{R}^{|\mathcal{L}|}$ collect line resistances and reactances, and define $R := \text{diag}(\mathbf{r})$, $X := \text{diag}(\mathbf{x})$. Let A be the network incidence matrix and define matrices $H_P = A^{-1}RA^{-\top}$ and $H_Q = A^{-1}XA^{-\top}$. Then, the voltage magnitudes are given as

$$\mathbf{u}_t^s = \mathbf{1} + H_P \mathbf{p}_t^{\text{inj},s} + H_Q \mathbf{q}_t^{\text{inj},s}, \quad t \in \mathcal{T}, s \in \{\wedge, \vee\} \quad (6a)$$

$$\underline{\mathbf{u}} \leq \mathbf{u}_t^s \leq \bar{\mathbf{u}}, \quad t \in \mathcal{T}, s \in \{\wedge, \vee\}. \quad (6b)$$

Constraint (6b) enforces lower and upper bounds on the squared nodal voltages, ensuring that voltage magnitudes remain within their admissible limits for all time periods and for both the upper and lower flexibility envelope trajectories.

The power injections and branch flows for all $t \in \mathcal{T}$ and $s \in \{\wedge, \vee\}$ are given as

$$\mathbf{p}_t^{\text{inj},s} = \mathbf{p}_{g,t}^s + \mathbf{p}_{e,t}^s + \mathbf{p}_t^{\text{PV}} - \mathbf{p}_t^{\text{load}}, \quad (7a)$$

$$\mathbf{q}_t^{\text{inj},s} = \mathbf{q}_{g,t}^s - \mathbf{q}_t^{\text{load}}, \quad (7b)$$

$$B \mathbf{P}_t^s = \mathbf{p}_t^{\text{inj},s}, \quad (7c)$$

$$B \mathbf{Q}_t^s = \mathbf{q}_t^{\text{inj},s}. \quad (7d)$$

where $\mathbf{p}_{g,t}^s$ and $\mathbf{p}_{e,t}^s$ define the generator and ESS injection vectors for each trajectory s and time t . The terms $\mathbf{p}_{g,t}^s$ and $\mathbf{p}_{e,t}^s$ are given as $\sum_{g \in G(n)} p_{g,t}^s$ and $\sum_{e \in E(n)} p_{e,t}^s$, respectively, where $G(n)$ and $E(n)$ denote the sets of generators and ESS units located at node n . Here, branch power-flows are denoted by \mathbf{P}_t^s and \mathbf{Q}_t^s , and B is matrix relating the injections to the branch flows. Constraints (7a)–(7b) define these nodal injections as the aggregation of generator and ESS outputs at each node minus the exogenous nodal load, thereby linking resource-level decisions to the network power-flow model.

Finally, at the GCP, the following constraints are imposed:

$$p_t^{\text{GCP},s} = \mathbf{1}^\top \mathbf{p}_t^{\text{inj},s}, \quad \forall t \in \mathcal{T}, \forall s \in \{\wedge, \vee\}, \quad (8a)$$

$$p_t^{\text{GCP},\wedge} \geq p_t^{\text{GCP},\vee}, \quad \forall t \in \mathcal{T}. \quad (8b)$$

Constraint (8a) defines the net active power exchanged at the GCP as the aggregation of nodal active-power injections,

where $\mathbf{1}$ is a vector of ones. Constraint (8b) enforces the ordering between the upper and lower aggregate trajectories, ensuring a valid flexibility envelope.

4) *Final formulation:* The optimization problem in (1)–(8) is linear and can be written in the following standard form.

$$\max \mathbf{c}^\top \mathbf{x} \quad (9a)$$

$$\text{subject to: } \mathbf{A}\mathbf{x} \leq \mathbf{b} \quad (9b)$$

$$\mathbf{C}\mathbf{x} \leq \mathbf{d}, \quad (9c)$$

where \mathbf{c} , \mathbf{b} , \mathbf{d} , and \mathbf{A} , \mathbf{C} are appropriate vectors and matrices. The variable \mathbf{x} contains all the variables, i.e., $[\mathbf{p}_{g,t}^s, \mathbf{q}_{g,t}^s, \mathbf{p}_{e,t}^s, \mathbf{e}_{e,t}^s, \mathbf{u}_{e,t}^s, \forall t, s \in \{\wedge, \vee\}]$.

B. Conservativeness of Ramping-aware Formulation

As will be demonstrated via the numerical validations in Section IV, the above formulation with ramping constraints results in a substantially smaller flexibility region than ones without. This is due to the worst-case transitions from upper to lower envelopes and vice-versa as included in constraints (3c)–(3f). To tackle this issue, we next propose a new model.

III. RAMPING-AWARE ENHANCED FLEXIBILITY AGGREGATION MODEL USING PRE-RAMPING STRATEGY

In order to address the conservativeness identified in the baseline formulation, we extend the model in Section II by introducing pre-ramping variables that enable resources to pre-set their outputs, thereby expanding the available ramping flexibility in subsequent time intervals.

The conceptual example shown in Figure 1 illustrates the impact of pre-ramping on upward flexibility. In the *first* scenario (left panel), no pre-ramping is applied. The generator can increase its output only up to its ramp-rate limit, while the ESS can contribute at most its available discharge capacity. As a result, the total output at time $t+1$ is limited to $p_g + R^\uparrow + p_e$.

In the *second* scenario (right panel), pre-ramping is enabled. By pre-increasing the generator output and simultaneously pre-decreasing the ESS output at time t , the system maintains the same net output at the GCP. This re-positioning allows both units to fully exploit their upward ramping capability in the next interval, leading to a higher achievable aggregate output at time $t+1$, i.e., $p_g + R^\uparrow + p_e + p^{\text{pre}}$, where the additional term p^{pre} captures the additional ramp-up capability provided by pre-ramping.

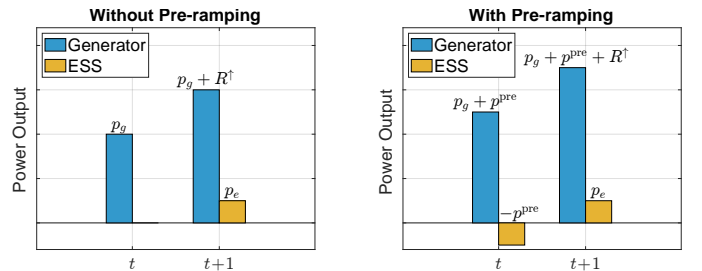


Fig. 1: Comparison of upward flexibility with and without pre-ramping. As shown, the upward flexibility increases from $p_g + R^\uparrow + p_e$ to $p_g + R^\uparrow + p_e + p^{\text{pre}}$, where p^{pre} captures the additional ramp-up capability provided by pre-ramping.

¹The losses can be taken into account by adding an equivalent resistance in series with the ideal ESS as implemented in [17], which can be embedded in the LinDistFlow model.

We next formally describe the enhanced model with pre-ramping variables. These variables are defined separately for generators and ESSs, and for the upper (\wedge) and lower (\vee) flexibility envelopes, as summarized in Table II. All pre-ramping variables are defined as non-negative, with their physical interpretations enforced through the constraints.

TABLE II: Interpretation of pre-ramping variables

Direction	Resource	Pre-Ramping Action	Variable
Upward	Generator	Increase output	$p_{g,t}^{\text{pre},\vee}$
	ESS	Decrease output (charge more)	$p_{e,t}^{\text{pre},\vee}$
Downward	Generator	Decrease output	$p_{g,t}^{\text{pre},\wedge}$
	ESS	Increase output (discharge more)	$p_{e,t}^{\text{pre},\wedge}$

Note: Each pre-ramping variable is associated with the envelope opposite to the direction of ramping. That is, ramp-up flexibility is prepared along the lower envelope (\vee), and ramp-down flexibility along the upper envelope (\wedge).

A. Modified Constraints

We define pre-ramped variables for generators and ESSs as $p_{g,t}^{\vee,\bullet} = p_{g,t}^{\vee} + p_{g,t}^{\text{pre},\vee}$, $p_{g,t}^{\wedge,\bullet} = p_{g,t}^{\wedge} - p_{g,t}^{\text{pre},\wedge}$, and $p_{e,t}^{\vee,\bullet} = p_{e,t}^{\vee} - p_{e,t}^{\text{pre},\vee}$, $p_{e,t}^{\wedge,\bullet} = p_{e,t}^{\wedge} + p_{e,t}^{\text{pre},\wedge}$. The modified constraint are same as (3c)–(3f) but the generator and ESS variables are replaced by pre-ramped variables. For generators, ramping feasibility is enforced by bounding all corner-to-corner transitions between the pre-ramped operating points at time $t-1$ and the envelope bounds at time t , which results in a set of eight linear ramping constraints; the complete set is reported in Appendix B. For ESS units, no ramp-rate limits are imposed; the baseline power and energy constraints retain the same structure and are evaluated using the pre-ramped variables.

B. Additional Constraints to Ensure Feasibility and Coordination

To ensure that the pre-ramped operating points remain physically realizable and compatible with both resource and network limits, additional feasibility and consistency constraints are introduced. These constraints guarantee that all upper and lower envelope trajectories, together with their associated pre-ramping actions, remain within the operational bounds of the generators, ESS units, and the distribution network. Moreover, these constraints are formulated to ensure that any trajectory contained within the resulting flexibility envelopes is disaggregable into individual resource schedules, as formally established in Appendix B.

Along with preventing excessive pre-ramping actions that violate the energy boundaries of the ESS, the total pre-ramping energy over time must also respect the initial SoC. Specifically, the following constraints ensure that the cumulative upward and downward pre-ramping energy does not exceed the remaining chargeable or dischargeable energy, respectively:

$$\sum_t p_{e,t}^{\text{pre},\wedge} \cdot \Delta t \leq E_e^{\text{init}} - E_e^{\text{min}}, \quad \forall e \in \mathcal{E}, \quad (10a)$$

$$\sum_t p_{e,t}^{\text{pre},\vee} \cdot \Delta t \leq E_e^{\text{max}} - E_e^{\text{init}}, \quad \forall e \in \mathcal{E}. \quad (10b)$$

It is worth noting that $p_{e,t}^{\text{pre},\vee}$ serves to support upward flexibility by reducing the ESS output in advance. This effectively increases charging, which is only possible if sufficient room remains below the maximum energy capacity. Conversely, $p_{e,t}^{\text{pre},\wedge}$ increases ESS output (discharging) to support downward flexibility, and is limited by the energy stored.

Pre-ramping is used to prepare future ramping capability and should not alter the net power injected at the GCP. In other words, from the perspective of the upstream grid, the GCP power must remain the same before and after applying the pre-ramp adjustments. This can be enforced by

$$p_t^{\text{GCP},\wedge} = p_t^{\text{GCP},\wedge,\bullet}, \quad \forall t \in \mathcal{T}, \quad (11a)$$

$$p_t^{\text{GCP},\vee} = p_t^{\text{GCP},\vee,\bullet}, \quad \forall t \in \mathcal{T}. \quad (11b)$$

Although the GCP invariance constraints (11) already imply the coordination between generators and ESSs during pre-ramping, we explicitly impose the following balancing constraints to clearly reflect the physical coordination mechanism between resource types:

$$\sum_{g \in \mathcal{G}} p_{g,t}^{\text{pre},\wedge} = \sum_{e \in \mathcal{E}} p_{e,t}^{\text{pre},\wedge}, \quad \forall t \in \mathcal{T}, \quad (12a)$$

$$\sum_{g \in \mathcal{G}} p_{g,t}^{\text{pre},\vee} = \sum_{e \in \mathcal{E}} p_{e,t}^{\text{pre},\vee}, \quad \forall t \in \mathcal{T}. \quad (12b)$$

C. Final Formulation of the Pre-Ramped Model

The final formulation of the proposed pre-ramped flexibility envelope model can be compactly expressed as follows. The key idea is that flexibility at the GCP is evaluated with respect to a baseline operating point, while pre-ramping actions are introduced to expand the aggregate flexibility envelope. Importantly, pre-ramping actions are defined on top of the baseline operating constraints rather than replacing them. Accordingly, both the baseline constraints and the associated pre-ramping constraints are jointly enforced to guarantee the feasibility of the expanded flexibility region.

1) *Final formulation:* Similar to the baseline model, the optimization problem with pre-ramping variables in (1)–(12) is linear and can be written in the following standard form.

$$\max \mathbf{c}^\top \mathbf{x} \quad (13a)$$

$$\text{subject to: } \mathbf{Ax} \leq \mathbf{b} \quad (13b)$$

$$\mathbf{Cx} = \mathbf{d} \quad (13c)$$

$$\mathbf{A}^\bullet \mathbf{x}^\bullet \leq \mathbf{b}^\bullet \quad (13d)$$

$$\mathbf{C}^\bullet \mathbf{x}^\bullet = \mathbf{d}^\bullet \quad (13e)$$

$$(10), (11), (12). \quad (13f)$$

where \mathbf{c}^\bullet , \mathbf{b}^\bullet , \mathbf{d}^\bullet and \mathbf{A}^\bullet , \mathbf{C}^\bullet are appropriate vectors and matrices. The variable \mathbf{x}^\bullet contains all the variables, i.e., $[p_{g,t}^{s,\bullet}, q_{g,t}^{s,\bullet}, p_{e,t}^{s,\bullet}, p_{g,t}^{\text{pre},s}, p_{e,t}^{\text{pre},s}, e_{e,t}^{s,\bullet}, u_{e,t}^{s,\bullet}, \forall t, s \in \{\wedge, \vee\}]$.

The baseline model can be interpreted as a special case of (13) by setting all pre-ramping variables to zero and removing the corresponding coordination constraints. It can be also observed that the optimization model in (13) is a linear program (LP) and can be solved efficiently by any LP solver.

IV. NUMERICAL VALIDATION

A. System Description and Simulation Setup

The proposed method is evaluated on the IEEE 33-bus radial distribution system (case33bw). We modify the system to include a ramp-constrained conventional generator, multiple ESSs, PV units, and time-varying loads. The detailed configurations and parameters of this system are provided in [18]. The conventional generator is located at bus 5, while ESS units are installed at buses 10, 13, 14, and 24; PV units and loads are assumed to be distributed across multiple buses in the system.

The test system is modeled with a base power of 10 MVA and a base voltage of 12.66 kV. To maintain power quality, node voltage magnitudes are strictly limited to the range of $[0.95, 1.05]$ p.u. The dispatchable resources are characterized by a maximum capacity of 215 kW, a minimum of 80 kW, and a ramp-rate limit of 100 kW/h [14]. ESS units are assumed to have a power rating of 12.5 kW and an energy capacity of 50 kWh, with the initial state of charge set to 50% of the rated capacity. Time-series profiles for PV generation and nodal loads are adopted from the smart grid test setup in [4].

Three case studies are considered.

- Case I presents results on aggregate flexibility envelopes via area maximization.
- Case II presents cost-optimal flexibility envelopes considering ramping products through net operational cost minimization under FRP markets.
- Case III presents a robust formulation accounting for uncertainty of PV and load forecasts.

For the economic evaluation in Case II, we consider hourly energy prices based on 2024 average CAISO locational marginal prices (LMPs), which ranges from \$5.8/MWh to \$66.7/MWh, with an average of \$32.4/MWh [19]. Reserve capacity prices are set to \$20/MW for both upward and downward reserves, while compensation for providing FRPs is set to \$5.44/MW [20], [21]. The marginal generation cost of dispatchable units is assumed to be \$14.5/MWh [14].

The simulation horizon spans 24 hours with an hourly resolution, but the proposed formulation is not tied to a specific dispatch interval. Under tighter real-time operational windows (e.g., 5 or 15 minutes), ramp-rate constraints become increasingly binding, which can substantially limit the deployable flexibility of ramp-constrained distributed energy resources.

B. Case I: Aggregate Flexibility Envelopes via Area Maximization

We examine the effectiveness of the proposed ramping-aware formulations in terms of feasibility with ramp constraints and in area within the aggregate flexibility envelopes. The optimized aggregated lower and upper envelopes are shown in Figure 2 for the model without ramping constraint (i.e., without eq. (3)), the ramping-aware baseline model, and the enhanced model with pre-ramping decisions.

Figures 2(a)-(c) show optimized upper and lower envelopes in solid and dashed blue. The trajectories in between the upper and lower envelopes are also included to show whether they can be disaggregated or not. The disaggregation feasibility

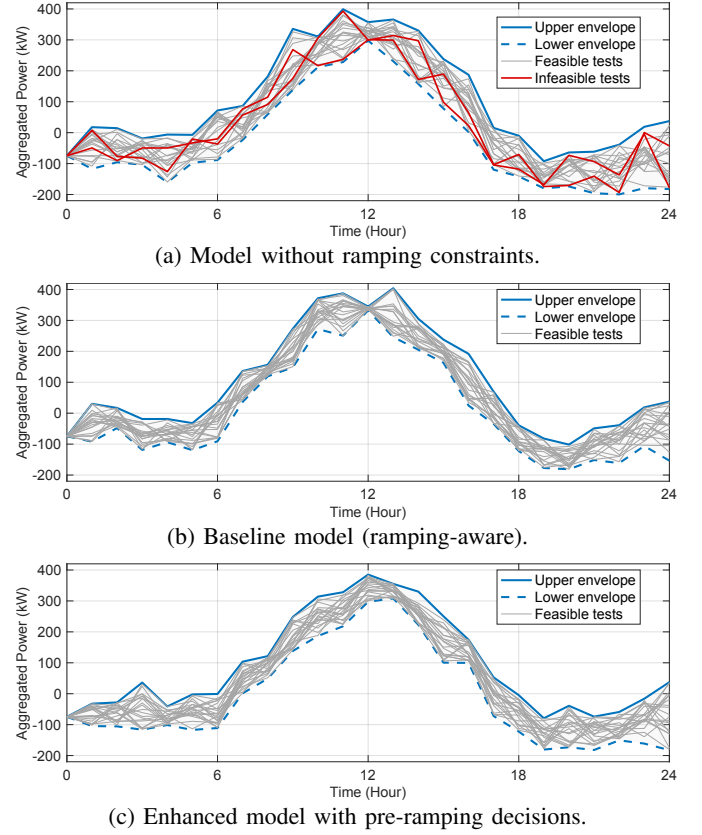


Fig. 2: Comparison of disaggregation feasibility for various GCP flexibility envelope models. Positive values represent power export from the distribution system to the upstream grid. (a) envelope neglecting ramp-rate constraints where red trajectories indicate infeasible disaggregation, (b) the baseline model considering ramping limits, and (c) the proposed enhanced model with pre-ramping decisions.

of the constructed envelopes is verified through Monte Carlo simulations following the methodology established in [6]. Specifically, 1,000 vertex trajectories and 4,000 randomly sampled trajectories within the envelope are tested for each case. The trajectories in Figure 2 correspond to a representative subset of these Monte Carlo samples, selected solely for clarity of visualization. As shown in Figure 2a, when ramping constraints are neglected, several sampled trajectories inside the aggregate GCP envelopes cannot be disaggregated into feasible device-level schedules, as indicated by the infeasible trajectories (shown in red). This occurs because inter-temporal ramp-rate limits of dispatchable resources restrict the magnitude of feasible aggregate power changes between consecutive time intervals, even if the aggregate power remains within the envelope bounds. By explicitly accounting for these ramping constraints, the ramping-aware baseline model in Figure 2b eliminates such infeasible trajectories, ensuring that all tested trajectories within the envelope are deliverable. Moreover, the enhanced model with pre-ramping decisions in Figure 2c further expands the aggregate flexibility envelope by strategically re-positioning resources in advance, thereby enabling larger feasible power variations at the GCP. Importantly, all tested trajectories are found to be feasible, confirming that every aggregate power trajectory within the proposed envelopes can be disaggregated into device-level schedules.

TABLE III: Impact of ESS Capacity and Placement on Aggregate Flexibility Envelope Area (per-unit ESS ratings; four total ESS units).

ESS Rating (kW / kWh)	w/o Ramping	Envelope Area (kWh)	
		Baseline	Pre-ramp
12.5 / 50	3324.90	2490.96	2619.90 (+5.2%)
12.5 / 50 [†]	3329.71	2500.00	2635.00 (+5.4%)
25 / 100	3391.52	2557.57	2786.52 (+9.0%)
37.5 / 150	3457.62	2624.13	2952.62 (+12.5%)
62.5 / 250	3589.58	2756.09	3284.58 (+19.2%)

[†] ESS units are relocated to buses 5–8, near the generator.

Table III compares the aggregate flexibility envelope area at the GCP under different ESS capacity and placement scenarios. As shown in Table III, the flexibility gains enabled by pre-ramping grow significantly as ESS capacity increases, highlighting the key role of coordinated pre-ramping actions between the generator and ESS units. By proactively adjusting their operating points, the proposed formulation relaxes inter-temporal ramping constraints and unlocks additional aggregate flexibility. In contrast, the baseline formulation exhibits only modest improvements with increasing ESS capacity, as ramp-rate limits remain binding without pre-ramping.

The impact of resource placement is also examined. For the same ESS capacity, placing storage units on buses closer to the generator expands the flexibility envelope, as shown in the last row of Table III. This suggests that reduced electrical distances between dispatchable units can mitigate network-induced limitations and better facilitate the coordinated ramping required for envelope enlargement.

C. Case II: Cost-Optimal Operation under FRP Markets

This case study investigates the economic implications of the proposed pre-ramping formulation when the distribution system participates in energy, reserve, and FRP markets. In contrast to Case I, which focuses on maximizing the aggregate flexibility envelope, this case also optimizes a reference (base) trajectory at the GCP, representing the scheduled operating point of the distribution system.

Accordingly, three aggregate power trajectories are optimized simultaneously: an upper trajectory $p_t^{GCP,\wedge}$, a lower trajectory $p_t^{GCP,\vee}$, and a base trajectory $p_t^{GCP,-}$. The base trajectory is constrained to lie within the flexibility envelope,

$$p_t^{GCP,\vee} \leq p_t^{GCP,-} \leq p_t^{GCP,\wedge}, \quad \forall t \in \mathcal{T}, \quad (14)$$

ensuring that both upward and downward flexibility are deliverable relative to the scheduled dispatch.

Based on these trajectories, the available upward and downward reserve capacities are defined as

$$R_t^{cap,\uparrow} = p_t^{GCP,\wedge} - p_t^{GCP,-}, \quad R_t^{cap,\downarrow} = p_t^{GCP,-} - p_t^{GCP,\vee}. \quad (15)$$

In addition, FRP quantities capture the deliverable ramping capability between consecutive time periods:

$$R_t^{frp,\uparrow} = p_{t+1}^{GCP,\wedge} - p_t^{GCP,-}, \quad R_t^{frp,\downarrow} = p_t^{GCP,-} - p_{t+1}^{GCP,\vee}. \quad (16)$$

These definitions, applied for all $t \in \mathcal{T}$, ensure that reserve and ramping services are consistent with inter-temporal ramp-rate constraints.

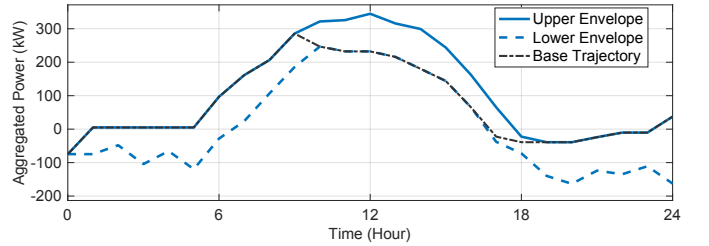
The objective function minimizes the net operating cost of the distribution system over the scheduling horizon:

$$\min \sum_t \left(C_t^{\text{energy}} - R_t^{\text{reserve}} - R_t^{\text{FRP}} \right). \quad (17)$$

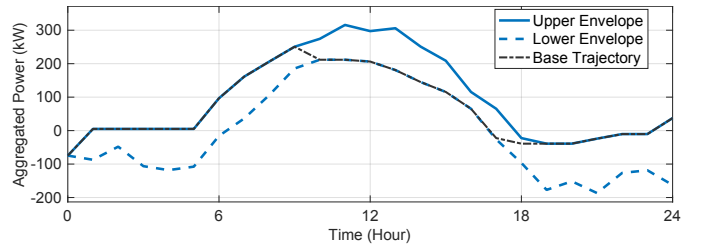
Here, C_t^{energy} corresponds to the operating cost of the base trajectory and includes the energy procurement or export cost at the grid connection point, valued at energy prices, together with the generation cost of dispatchable generators.

The revenue terms consist of compensation for reserve capacity and flexible ramping services. In particular, R_t^{reserve} is determined by applying the reserve capacity price to the available upward and downward reserve margins, while R_t^{FRP} is computed by valuing the deliverable ramping capacities using the FRP price defined in the market setup. This formulation directly links the economic incentives to the deliverable flexibility characterized by the aggregate power trajectories.

All resource-level operating constraints, ramping limits, and distribution-network constraints remain consistent with the flexibility envelope formulation in Section II and Section III. By jointly optimizing the base dispatch and the flexibility envelope, the proposed model enables the distribution system to economically schedule its resources while simultaneously maximizing the value of reserve and ramping services.



(a) Baseline model (ramping-aware)



(b) Enhanced model with pre-ramping decisions

Fig. 3: Aggregate power trajectories at the GCP in Case II. The black line denotes the base (reference) dispatch.

Figure 3 shows the results for the cost-optimal operation under FRP markets. It shows the aggregated lower and upper envelopes along with the GCP base trajectory. Observe that the base (reference) trajectory is contained between the upper and lower envelopes. Figure 4 shows the generator output trajectories corresponding to the aggregate GCP trajectories. With pre-ramping enabled (Figure 4(b)), the relaxed ramping constraints are clearly reflected in a larger separation between the upper and lower envelopes. Under the baseline formulation (Figure 4(a)), the separation remains fixed at 100 kW across all time periods due to binding ramp-rate constraints. The generator's base trajectory exhibits similar patterns in both cases and always satisfies physical ramping limits. The base

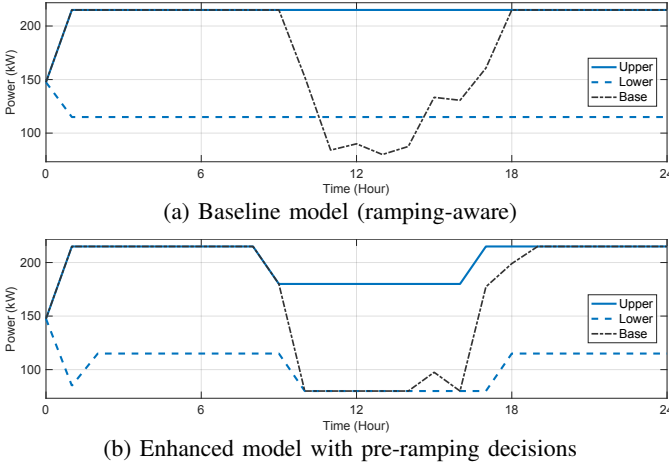


Fig. 4: Generator output trajectories in Case II corresponding to the aggregate trajectories at the GCP. The black line denotes the base (reference) dispatch.

TABLE IV: Economic Results: Comparison of Baseline and Pre-ramp Cases (per-unit ESS ratings; four ESS units in total)

ESS Rating (kW / kWh)	Case	C^{energy} (\$)	$R^{\text{res+frp}}$ (\$)	Objective (\$)
12.5 / 50	Baseline	42.52	63.60	-21.08
	Pre-ramp	41.23	64.36	-23.14
	Diff.	-1.29	+0.76	-2.06
37.5 / 150	Baseline	9.44	67.37	-57.94
	Pre-ramp	8.19	68.26	-60.07
	Diff.	-1.25	+0.89	-2.14
62.5 / 250	Baseline	-20.87	70.02	-90.89
	Pre-ramp	-22.09	70.91	-93.00
	Diff.	-1.22	+0.89	-2.11
125 / 500	Baseline	-71.76	73.81	-145.57
	Pre-ramp	-71.01	83.63	-154.64
	Diff.	+0.75	+9.82	-9.07

trajectory is not necessarily confined within the generator-level upper and lower bounds, as these bounds are auxiliary constructs for defining the aggregate GCP flexibility envelope rather than strict limits on the base dispatch.

The quantitative comparison of the costs are summarized in Table IV. A negative energy (C^{energy}) value represents exports. Regardless of the installed ESS capacity, the introduction of the proposed pre-ramping consistently improves economic performance. Specifically, pre-ramping yields both higher revenues from reserve and FRP provision and lower operating costs. These gains arise from the expansion of the feasible operating region, which enables the distribution system to exploit flexibility more effectively without requiring additional physical resources, thereby enhancing economic efficiency through flexibility reallocation rather than capacity expansion.

Across all tested capacities, the pre-ramping strategy consistently improves the objective value compared to the baseline. For smaller ESS ratings, the improvement remains relatively modest, as the storage is mainly used for energy cost reduction. However, as the ESS capacity increases, the economic gain grows significantly, driven by a sharp rise in reserve and FRP revenues ($R^{\text{res+frp}}$). This shows that while pre-ramping is effective even at small scales, its true value is realized in larger

systems where the additional capacity can be fully shifted from energy arbitrage to providing high-value flexibility services.

Overall, the results show that the proposed pre-ramping formulation enlarges the feasible flexibility region and improves economic performance in energy, reserve, and FRP markets.

D. Case III: Robust Envelope under Forecast Errors

The above formulations assume that the PV and load forecasts can be obtained with 100% certainty which is not realistic. Thus, we next consider a robust version of the proposed scheme accounting for uncertain PV and load injections. Due to the linear model, robust constraints can efficiently added in the previous formulation as described below.

To ensure the disaggregation of the flexibility envelope under net-load forecast uncertainty, we incorporate a robust optimization framework. Specifically, the nodal net-load realizations are modeled by box uncertainty bounds in both load demand ξ_t^L and PV generation ξ_t^{PV} :

$$p_t^{\text{net}} = (p_t^{\text{load}} + \xi_t^L) - (p_t^{\text{PV}} + \xi_t^{PV}), \quad (18)$$

where p_t^{load} and p_t^{PV} are the forecasted values. The uncertainty terms reside within a box uncertainty set \mathcal{U}_t :

$$\mathcal{U}_t = \{(\xi_t^L, \xi_t^{PV}) : |\xi_t^L| \leq \bar{\xi}_t^L, |\xi_t^{PV}| \leq \bar{\xi}_t^{PV}\}. \quad (19)$$

The maximum anticipated deviations are defined as $\bar{\xi}_t^L = \alpha |p_t^{\text{load}}|$ and $\bar{\xi}_t^{PV} = \beta |p_t^{\text{PV}}|$, where α and β represent the forecast error percentages for load and PV, respectively.

Since the voltage model (6a) is affine, the uncertainty impacts nodal voltages through the term $-H_P(\xi_t^L - \xi_t^{PV})$.² To account for worst-case voltage deviations within \mathcal{U}_t , the robust counterpart of the nominal voltage constraints (6b) is:

$$\max_{(\xi_t^L, \xi_t^{PV}) \in \mathcal{U}_t} \{u_t^s - H_P(\xi_t^L - \xi_t^{PV})\} \leq \bar{u}, \quad (20a)$$

$$\min_{(\xi_t^L, \xi_t^{PV}) \in \mathcal{U}_t} \{u_t^s - H_P(\xi_t^L - \xi_t^{PV})\} \geq \underline{u}. \quad (20b)$$

By applying the analytical robust reformulation for box uncertainty, the deterministic tightened margins are derived as $\Delta u_t^{\text{unc}} := |H_P|(\bar{\xi}_t^L + \bar{\xi}_t^{PV})$. Consequently, the robust model is obtained by replacing the nominal constraints (6b) and with the following tightened versions:

$$\underline{u} + \Delta u_t^{\text{unc}} \leq u_t^s \leq \bar{u} - \Delta u_t^{\text{unc}}, \quad (21a)$$

$$\underline{u} + \Delta u_t^{\text{unc}} \leq u_t^{s^*} \leq \bar{u} - \Delta u_t^{\text{unc}}, \quad (21b)$$

for all $t \in \mathcal{T}$ and $s \in \{\wedge, \vee\}$. This formulation ensures that the flexibility envelope remains feasible even under simultaneous worst-case fluctuations of load and renewable generation.

To investigate the impact of net-load forecast uncertainty on aggregated flexibility, the proposed robust formulation is evaluated under different forecast error levels. For simplicity, identical error bounds are assumed for load and PV generation ($\alpha = \beta$), and varied from 0% to 10%. For each case, flexibility envelopes are computed for both the baseline and pre-ramped models using the robust voltage constraints.

²Reactive power uncertainty can be handled analogously using the sensitivity matrix H_Q .

Table V summarizes the resulting aggregate envelope areas at the GCP. As the forecast error increases, the envelope area gradually decreases, reflecting the proactive reservation of voltage margins. Although pre-ramping consistently outperforms the baseline, its relative improvement declines from 5.2% to 3.7% at $\alpha = \beta = 10\%$, as the tightened voltage limits reduce the available dispatch range for pre-ramping actions.

TABLE V: Impact of Forecast Errors on Robust Flexibility Envelopes

Forecast Error (%)	Envelope Area (kWh)		Voltage Extremes (pu)	
	Baseline	Pre-ramp	max $ V $	min $ V $
0	2490.96	2619.90	1.0500	0.9500
3	2488.77	2610.41	1.0468	0.9511
5	2487.31	2601.31	1.0447	0.9518
10	2473.72	2564.41	1.0403	0.9533

From a feasibility perspective, all constructed envelopes satisfy the voltage limits. Increasing uncertainty pushes the nodal voltage extrema ($\max |V|, \min |V|$) further away from the operational bounds (1.05 and 0.95 p.u.), reflecting increasingly conservative robust margins. The robustness of the envelopes is further validated via Monte Carlo simulations, which confirm that randomly sampled aggregate trajectories remain disaggregable and voltage-feasible.

Overall, these results highlight a trade-off between robustness and flexibility: while uncertainty necessitates conservative voltage margins that reduce the envelope size, the proposed pre-ramping strategy consistently preserves additional deliverable flexibility under worst-case conditions.

V. CONCLUSIONS

This paper presented a ramping-aware flexibility aggregation framework for power distribution networks leveraging distributed energy resources (DERs). Since ramp constraints inherently reduce the available flexibility region, we introduced a pre-ramping strategy to mitigate this limitation. The proposed approach proactively coordinates ramp-constrained resources (e.g., generators) with ramp-unconstrained resources (e.g., energy storage systems), thereby expanding the flexibility region despite ramping limitations.

The framework was validated on the IEEE 33-bus benchmark network with multiple distributed generators and storage units. Simulations show that the pre-ramping model significantly enhances the flexibility region through proactive actions of energy storage units. Additionally, a cost-optimal formulation was developed to enable participation in multiple markets, including energy, reserve capacity, and flexible ramping products. Comparative analysis shows that the pre-ramping model yields higher economic benefits than the baseline approach. Finally, the formulation was extended to incorporate uncertainties in load and PV forecasts, confirming its robustness and effectiveness under uncertain operating conditions.

APPENDIX A

PROOF OF DISAGGREGATION GUARANTEE FOR THE BASELINE MODEL

Consider the baseline envelope solutions $\{p_{g,t}^\wedge, p_{g,t}^\vee\}$ and $\{p_{e,t}^\wedge, p_{e,t}^\vee\}$ that satisfy the baseline constraints in Section II-A,

including (2)–(4). Let the corresponding GCP envelopes be denoted by $\{p_t^{\text{GCP},\wedge}, p_t^{\text{GCP},\vee}\}$.

For any aggregate trajectory $\{p_t^{\text{GCP},o}\}$ such that $p_t^{\text{GCP},\vee} \leq p_t^{\text{GCP},o} \leq p_t^{\text{GCP},\wedge}$ for all $t \in \mathcal{T}$, define the auxiliary coefficient $\lambda_t \in [0, 1]$ as

$$\lambda_t := \begin{cases} \frac{p_t^{\text{GCP},\wedge} - p_t^{\text{GCP},o}}{p_t^{\text{GCP},\wedge} - p_t^{\text{GCP},\vee}}, & p_t^{\text{GCP},\wedge} > p_t^{\text{GCP},\vee}, \\ 0, & \text{otherwise,} \end{cases}$$

$$p_t^{\text{GCP},o} = \lambda_t p_t^{\text{GCP},\vee} + (1 - \lambda_t) p_t^{\text{GCP},\wedge}. \quad (\text{A.1})$$

We construct a disaggregation by convex interpolation:

$$p_{g,t}^o := \lambda_t p_{g,t}^\vee + (1 - \lambda_t) p_{g,t}^\wedge, \quad p_{e,t}^o := \lambda_t p_{e,t}^\vee + (1 - \lambda_t) p_{e,t}^\wedge. \quad (\text{A.2})$$

By linearity of aggregation, (A.1)–(A.2) and the definition of GCP power in (8a) imply

$$p_t^{\text{GCP},o} = \mathbf{1}^\top (p_{g,t}^o + p_{e,t}^o + p_t^{\text{PV}} - p_t^{\text{load}}), \quad \forall t \in \mathcal{T}. \quad (\text{A.3})$$

Moreover, since generator/ESS power limits are interval constraints, $p_{g,t}^o$ and $p_{e,t}^o$ remain feasible for all t as convex combinations of feasible endpoints.

Next, we verify ramp feasibility. For each g and $t \geq 2$, the baseline model imposes the cross-corner bounds (3), which ensure that all four “corners” $(p_{g,t-1}^a, p_{g,t}^b)$ with $a, b \in \{\vee, \wedge\}$ satisfy the ramp limits. Since the ramp constraints define a convex set in $(p_{g,t-1}, p_{g,t})$, any point within the rectangle $[p_{g,t-1}^\vee, p_{g,t-1}^\wedge] \times [p_{g,t}^\vee, p_{g,t}^\wedge]$ also satisfies the ramp limits. Since $(p_{g,t-1}^o, p_{g,t}^o)$ lies in this rectangle by construction, the realized trajectory $\{p_{g,t}^o\}$ satisfies the ramp-rate limits.

For ESS energy feasibility, define the realized SoC dynamics $e_{e,1}^o = E_e^{\text{init}}$ and $e_{e,t}^o = e_{e,t-1}^o - \Delta t p_{e,t-1}^o$ for $t \geq 2$:

$$e_{e,t}^o = E_e^{\text{init}} - \Delta t \sum_{\tau=1}^{t-1} p_{e,\tau}^o. \quad (\text{A.4})$$

Since $p_{e,\tau}^\vee \leq p_{e,\tau}^o \leq p_{e,\tau}^\wedge$ for all τ , (A.4) yields the sandwich bound

$$E_e^{\text{init}} - \Delta t \sum_{\tau=1}^{t-1} p_{e,\tau}^\wedge \leq e_{e,t}^o \leq E_e^{\text{init}} - \Delta t \sum_{\tau=1}^{t-1} p_{e,\tau}^\vee. \quad (\text{A.5})$$

The baseline energy-coupling constraints (4d)–(4e) (and bounds (4f)) ensure that the left-hand side of (A.5) is no smaller than E_e^{min} and the right-hand side is no larger than E_e^{max} , thus $E_e^{\text{min}} \leq e_{e,t}^o \leq E_e^{\text{max}}$.

Therefore, for any aggregate trajectory inside the baseline envelope, the construction (A.2)–(A.4) produces a feasible disaggregation that respects generator power/ramp constraints and ESS power/SoC constraints, while exactly matching the aggregate trajectory at the GCP.

Finally, the disaggregation guarantee directly extends to LinDistFlow network constraints. Since nodal injections, branch flows, and squared voltages are affine functions of device-level injections under LinDistFlow, and voltage limits define a convex (feasible) set, any convex interpolation of two feasible envelope solutions at time t remains feasible. Thus, all interior aggregate trajectories preserve network feasibility.

APPENDIX B

PROOF OF DISAGGREGATION GUARANTEE FOR THE
PRE-RAMPED MODEL

Consider the pre-ramped envelope solutions $\{p_{g,t}^\wedge, p_{g,t}^\vee\}$, $\{p_{e,t}^\wedge, p_{e,t}^\vee\}$ and the associated pre-ramping variables $\{p_{g,t}^{\text{pre},\wedge}, p_{g,t}^{\text{pre},\vee}\}$ and $\{p_{e,t}^{\text{pre},\wedge}, p_{e,t}^{\text{pre},\vee}\}$ obtained from Section III. These variables satisfy all pre-ramped feasibility constraints, including (10)–(12). Let the corresponding GCP envelopes be denoted by $\{p_t^{\text{GCP},\wedge}, p_t^{\text{GCP},\vee}\}$. By (11), pre-ramping does not change the net power at the GCP, i.e., $p_t^{\text{GCP},s} = p_t^{\text{GCP},s\bullet}$ for all $t \in \mathcal{T}$ and $s \in \{\wedge, \vee\}$.

For any aggregate trajectory $\{p_t^{\text{GCP},o}\}$ such that $p_t^{\text{GCP},\vee} \leq p_t^{\text{GCP},o} \leq p_t^{\text{GCP},\wedge}$ for all $t \in \mathcal{T}$, define $\lambda_t \in [0, 1]$ as in (A.1). We construct a disaggregation by convex interpolation:

$$p_{g,t}^o := \lambda_t p_{g,t}^\vee + (1 - \lambda_t) p_{g,t}^\wedge, \quad p_{e,t}^o := \lambda_t p_{e,t}^\vee + (1 - \lambda_t) p_{e,t}^\wedge, \quad (\text{B.1})$$

and similarly for the pre-ramped operating points,

$$p_{g,t}^{o\bullet} := \lambda_t p_{g,t}^{\vee\bullet} + (1 - \lambda_t) p_{g,t}^{\wedge\bullet}, \quad p_{e,t}^{o\bullet} := \lambda_t p_{e,t}^{\vee\bullet} + (1 - \lambda_t) p_{e,t}^{\wedge\bullet}. \quad (\text{B.2})$$

By linearity of aggregation, (A.1) and (B.1) with (8a) imply

$$p_t^{\text{GCP},o} = \mathbf{1}^\top (p_{g,t}^o + p_{e,t}^o + p_t^{\text{PV}} - p_t^{\text{load}}), \quad \forall t \in \mathcal{T}. \quad (\text{B.3})$$

Moreover, since generator/ESS power limits are interval constraints, $p_{g,t}^o$ and $p_{e,t}^o$ remain feasible for all t as convex combinations of feasible endpoints. The same argument applies to $p_{g,t}^{o\bullet}$ and $p_{e,t}^{o\bullet}$ because the pre-ramped operating points are also enforced to be feasible in the pre-ramped model.

Next, we verify generator ramp feasibility. In the pre-ramped model, the physical transition from $t-1$ to t is evaluated between the pre-ramped operating point at $t-1$ and the envelope dispatch at t , i.e., between $(p_{g,t-1}^\bullet, p_{g,t})$. Ramping feasibility is enforced by bounding all corner-to-corner transitions between $\{p_{g,t-1}^{\vee\bullet}, p_{g,t-1}^{\wedge\bullet}\}$ and $\{p_{g,t}^\vee, p_{g,t}^\wedge\}$, which yields the following eight linear constraints:

$$p_{g,t}^\wedge - (p_{g,t-1}^\vee + p_{g,t-1}^{\text{pre},\vee}) \leq R_g^\uparrow, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (\text{B.4a})$$

$$(p_{g,t-1}^\vee + p_{g,t-1}^{\text{pre},\vee}) - p_{g,t}^\wedge \leq R_g^\downarrow, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (\text{B.4b})$$

$$p_{g,t}^\vee - (p_{g,t-1}^\wedge + p_{g,t-1}^{\text{pre},\wedge}) \leq R_g^\uparrow, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (\text{B.4c})$$

$$(p_{g,t-1}^\wedge + p_{g,t-1}^{\text{pre},\wedge}) - p_{g,t}^\vee \leq R_g^\downarrow, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (\text{B.4d})$$

$$p_{g,t}^\wedge - (p_{g,t-1}^\wedge + p_{g,t-1}^{\text{pre},\wedge}) \leq R_g^\uparrow, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (\text{B.4e})$$

$$p_{g,t}^\vee - (p_{g,t-1}^\vee + p_{g,t-1}^{\text{pre},\vee}) \leq R_g^\uparrow, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (\text{B.4f})$$

$$(p_{g,t-1}^\vee + p_{g,t-1}^{\text{pre},\vee}) - p_{g,t}^\wedge \leq R_g^\downarrow, \quad \forall g \in \mathcal{G}, t \geq 2, \quad (\text{B.4g})$$

$$(p_{g,t-1}^\wedge + p_{g,t-1}^{\text{pre},\wedge}) - p_{g,t}^\vee \leq R_g^\downarrow, \quad \forall g \in \mathcal{G}, t \geq 2. \quad (\text{B.4h})$$

Since these inequalities bound the ramp for all four corner pairs, the feasible set in $(p_{g,t-1}^\bullet, p_{g,t})$ is convex. Hence, the pair $(p_{g,t-1}^{o\bullet}, p_{g,t}^o)$ also satisfies the generator ramp-rate limits.

For ESS units, the energy-feasibility argument follows the same structure as in Appendix A. The only difference is that, in the pre-ramped model, the SoC dynamics are driven by the pre-ramped net power $p_{e,t}^o$. Since the pre-ramped endpoint SoC trajectories corresponding to $\{\vee\bullet, \wedge\bullet\}$ are enforced to be feasible by the ESS energy and budget constraints, the same sandwich argument applies. Network feasibility under LinDistFlow is preserved by convex interpolation and therefore follows directly from Appendix A.

REFERENCES

- [1] O. Yurdakul, E. Ela, and F. Billimoria, "Flexible ramping product procurement in day-ahead markets," *IEEE Transactions on Energy Markets, Policy and Regulation*, vol. 3, no. 1, pp. 13–31, 2024.
- [2] R. Luo, S. Li, X. Liu, X. Han, and Z. Li, "Uncertainty-aware demand estimation and market clearing of flexible ramping products in power systems dominated by renewable energy," *Renewable Energy*, p. 124019, 2025.
- [3] H. Li, B. Qin, S. Wang, T. Ding, J. Liu, and H. Wang, "Aggregate power flexibility of multi-energy systems supported by dynamic networks," *Applied Energy*, vol. 377, p. 124565, 2025.
- [4] R. K. Gupta, S. Fahmy, M. Chevron, R. Vasapollo, E. Figini, and M. Paolone, "Grid-aware scheduling and control of electric vehicle charging stations for dispatching active distribution networks: Theory and experimental validation," *IEEE Transactions on Smart Grid*, 2025.
- [5] Y. Wen, Z. Hu, and L. Liu, "Aggregate temporally coupled power flexibility of DERs considering distribution system security constraints," *IEEE Transactions on Power Systems*, vol. 38, no. 4, pp. 3884–3896, 2022.
- [6] X. Chen, E. Dall'Anese, C. Zhao, and N. Li, "Aggregate power flexibility in unbalanced distribution systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 258–269, 2019.
- [7] E. Öztürk, T. Faulwasser, K. Worthmann, M. Preißinger, and K. Rheinberger, "Alleviating the curse of dimensionality in Minkowski sum approximations of storage flexibility," *IEEE Transactions on Smart Grid*, vol. 15, no. 6, pp. 5733–5743, 2024.
- [8] Y. Wen, Y. Guo, Z. Hu, and G. Hug, "Quantifying and optimizing the time-coupled flexibilities at the distribution-level for TSO-DSO coordination," *IEEE Transactions on Power Systems*, vol. 40, no. 6, pp. 5071–5085, 2025.
- [9] X. Chen and N. Li, "Leveraging two-stage adaptive robust optimization for power flexibility aggregation," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 3954–3965, 2021.
- [10] F. Al Taha, T. Vincent, and E. Bitar, "An efficient method for quantifying the aggregate flexibility of plug-in electric vehicle populations," *IEEE Transactions on Smart Grid*, vol. 16, no. 4, pp. 3142–3154, 2025.
- [11] L. Na, G. Song, Z. Xiaoyu, and C. Xuri, "Aggregating and scheduling flexibility of multiple flexible resources for distribution network: A feasible region-embedded EMPC method," in *IEEE PES 16th Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 2024.
- [12] M. Gooding, "Microsoft and Caterpillar power data center for 48 hours using hydrogen fuel cells," <https://www.datacenterdynamics.com/en/news/microsoft-caterpillar-hydrogen-fuel-cell-data-center/>, Jan. 2024, accessed: Dec. 21, 2025.
- [13] S. Cha, "The potential role of small modular reactors (SMRs) in addressing the increasing power demand of the artificial intelligence industry: A scenario-based analysis," *Nuclear Engineering and Technology*, vol. 57, no. 4, p. 103314, 2025.
- [14] M. Alipour, K. Zare, H. Seyedi, and M. Jalali, "Real-time price-based demand response model for combined heat and power systems," *Energy*, vol. 168, pp. 1119–1127, 2019.
- [15] D. Michaelson and J. Jiang, "Review of integration of small modular reactors in renewable energy microgrids," *Renewable and Sustainable Energy Reviews*, vol. 152, p. 111638, 2021.
- [16] X. Huang, P. Dong, X. He, M. Liu, and W. Deng, "Research on aggregation flexibility method of AC/DC distribution network considering flexibility balance and its application," *IEEE Systems Journal*, vol. 17, no. 3, pp. 3635–3645, 2023.
- [17] R. Gupta and F. Sossan, "Optimal sizing and siting of energy storage systems considering curtailable photovoltaic generation in power distribution networks," *Applied Energy*, vol. 339, p. 120955, 2023.
- [18] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, 1989.
- [19] California ISO. (2026) OASIS: Open Access Same-time Information System. Accessed: Jan. 15, 2026. [Online]. Available: <https://oasis.caiso.com/>
- [20] C. Feng, Q. Chen, Y. Wang, J. Ma, and X. Wang, "Frequency regulation service provision for virtual power plants through 5G RAN slicing," *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4943–4956, 2022.
- [21] B. Zhang and M. Kezunovic, "Impact on power system flexibility by electric vehicle participation in ramp market," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1285–1294, 2015.