



Integrated Two-Stage Stochastic Security-Constrained Unit Commitment with EV-V2G, Utility-Scale Storage, and Flexible Loads under High Renewable Penetration

Abbas Safari¹, Mehdi Jafari Shahbazadeh², Mahdijeh Eslami³ and Hesam Rahbarimagham^{*4}

Abstract-- This paper presents a comprehensive two-stage stochastic security-constrained unit commitment (SCUC) framework that fully integrates electric vehicles (EVs) with vehicle-to-grid (V2G) capabilities, utility-scale energy storage systems (ESS), and flexible demand response under high levels of wind and solar generation. In the first stage, thermal unit on/off decisions and charge/discharge statuses for ESS and EV fleets are co-optimized to secure reserves and meet mobility constraints. The second stage dispatches generation, reserves, and flexible load adjustments for each renewable-forecast scenario, while enforcing N-1 contingency criteria for both generator and transmission-line outages. Key innovations include a novel EV-V2G submodel that tracks state-of-charge (SoC), enforces arrival/departure requirements, and co-optimizes reserve provision; an ESS formulation that co-optimizes energy arbitrage with spinning and non-spinning reserves; and a flexible-load shifting paradigm that permits both time-shiftable consumption and curtailment at a user-dissatisfaction penalty. Renewable uncertainty is captured through a scenario-reduction technique applied to correlated wind and solar forecasting errors. A nested Benders-decomposition algorithm exploits scenario and contingency decomposition for tractability. Numerical experiments on a modified IEEE-118 bus system—using real-world wind/solar

traces and realistic EV/ESS parameters—demonstrate that the proposed model decreases expected operating and reserve-procurement costs by up to 8.5% relative to deterministic SCUC, cuts renewable curtailment from 35% to 20%, and reduces expected load-shedding under contingencies by over 75%. The joint flexibility of EVs, ESS, and flexible loads significantly enhances system reliability and economic performance in high-renewable power systems.

Keywords: Stochastic unit commitment, electric vehicles, energy storage systems, flexible loads, renewable energy

I. INTRODUCTION

Integrating large-scale wind and solar generation into modern power systems offers significant opportunities for reducing carbon emissions, but also introduces substantial operational challenges due to the variable and uncertain nature of renewable output. Traditional unit commitment models, which were developed around predictable thermal generation, struggle to accommodate the rapid fluctuations of renewable resources and maintain sufficient reserves to ensure reliable operation. As a result, system operators face increased risks of renewable curtailment, expensive

Received; 2025-05-24

Revised; 2025-08-23

Accepted; 2025-09-24

¹. Department of Electrical Engineering, Ke.C., Islamic Azad University, Kerman, Iran.

². Department of Electrical Engineering, Ke.C., Islamic Azad University, Kerman, Iran.

³. Department of Electrical Engineering, Ke.C., Islamic Azad University, Kerman, Iran.

⁴. Department of Electrical and Computer Engineering, Raf.C., Islamic Azad University, Rafsanjan, Iran.

• Corresponding author Email: h.rahbarim@aut.ac.ir

Cite this article as:

Safari, A., Sahbazadeh, M., J., Eslami, M., and Rahbarimagham, H. (2025). Integrated Two-Stage Stochastic Security-Constrained Unit Commitment with EV V2G, Utility-Scale Storage, and Flexible Loads under High Renewable Penetration. *Journal of Modeling & Simulation in Electrical & Electronics Engineering (MSEEE)*, 5(2), 37-61.

DOI: <https://doi.org/10.22075/MSEEE.2025.38263.1219>

redispatch actions, and even load shedding under adverse conditions.

This paper addresses the problem of determining least-cost, security-constrained unit commitment and dispatch decisions in a power system with high penetrations of renewable energy, while simultaneously leveraging emerging flexibility from electric vehicles with vehicle-to-grid capabilities, utility-scale energy storage systems, and responsive loads. The formulation takes the form of a two-stage stochastic mixed-integer program: the first stage involves making on/off decisions for thermal units and scheduling commitments for storage and aggregated EV fleets before actual renewable output is realized, ensuring minimum up/down times, spinning-reserve requirements, and state-of-charge constraints. In the second stage, for each scenario of wind and solar generation and each contingency event (such as a generator or transmission-line outage), the model optimizes real-time dispatch—adjusting generator outputs, storage charge/discharge actions, V2G injections, and flexible load shifts—while enforcing N-1 security criteria and meeting energy and reserve needs. Optimally integrating these novel flexibility resources into a unified unit commitment framework is critical for enhancing grid reliability in the face of renewable variability and uncertainty. Electric vehicles equipped with V2G capabilities can provide distributed storage that supports both energy arbitrage and dynamic reserves, while utility-scale storage smooths net-load ramping and supplies fast-responding reserves. Flexible demand response, from interruptible industrial processes to residential thermostatically controlled devices, further enhances the system's ability to balance supply and demand. By co-optimizing thermal generators, renewables, storage, EV fleets, and responsive loads, the proposed framework aims to reduce operating costs, minimize renewable curtailment, and improve resilience to contingencies. Numerical experiments using realistic wind and solar traces, EV mobility data, and storage parameters demonstrate that this integrated approach can significantly lower expected costs and enhance renewable utilization compared to deterministic benchmarks, thereby supporting the reliable and economical transition to a low-carbon power grid.

A. Motivation

The rapid growth of wind and solar generation has introduced significant variability and uncertainty into power system operations, making it challenging to maintain a reliable and cost-effective electricity supply under traditional unit commitment practices. As renewable penetration surpasses 30 percent in many regions, operators struggle to balance supply and demand, often resorting to costly redispatch or curtailing clean energy to uphold N-1 security requirements. Without additional flexibility, these measures undermine both economic and environmental objectives. At the same time, electric vehicles (EVs) with vehicle-to-grid capabilities, utility-scale energy storage, and demand-side response programs have emerged as powerful sources of system flexibility. EV fleets can act as distributed batteries, storage installations can rapidly smooth net-load fluctuations, and responsive loads can shift or curtail consumption during critical periods. However, these resources are typically modeled separately or in deterministic frameworks that fail to capture renewable uncertainty and contingency needs.

Therefore, a concise, two-stage stochastic unit commitment approach that jointly integrates EV-V2G, storage, and flexible demand is essential to unlock their combined potential—lowering operating costs, minimizing curtailment, and enhancing resilience in high-renewable power systems.

B. Literature review

The increasing penetration of renewable energy resources, coupled with the rapid transportation electrification and deployment of distributed energy storage, has motivated extensive research on advanced unit commitment (UC) formulations and stochastic optimization frameworks. Recent works have sought to enhance operational flexibility and reliability while effectively capturing the uncertainty inherent in wind and solar power generation.

Early contributions primarily focused on evolutionary and learning-based methods for UC under renewable uncertainty. For instance, a covariance matrix adaptation evolution strategy (CMAES)-based optimization framework was proposed in [1] to improve UC scheduling efficiency with the integration of electric vehicles (EVs) and renewables, while deep reinforcement learning techniques were introduced in [2] to enable model-free UC optimization with reduced computational complexity under wind variability. To address system reliability concerns, a stochastic UC formulation incorporating reliability constraints was presented in [3], while robust optimization approaches were employed in [4] and [10] to hedge against simultaneous source and load uncertainties.

Recognizing the flexibility potential of EVs and demand-side resources, several studies have explored EV-based ancillary services and demand response integration. In [5], a chance-constrained scheduling model accounted for EV-based frequency support under multiple uncertainties, while a bi-level scheduling formulation in [6] enhanced microgrid operation by co-optimizing wind-solar uncertainty and EV vehicle-to-grid (V2G) capabilities. Similarly, multi-objective dispatch strategies were developed for PV-battery energy storage system (BESS)-integrated charging stations with V2G [7], and uncertainty-aware scheduling models for V2G participation in microgrids were introduced in [8]. Cooperative scheduling frameworks integrating distributed generation, storage, and load with V2G aggregators have been further proposed to support low-carbon grid operation [9].

Energy storage systems (ESS) have also been investigated extensively as key flexibility enablers. A demand response aggregator-based model in [11] demonstrated how large-scale storage investment can improve market profits and flexibility, while stochastic vehicle scheduling models for renewable-building-transportation microgrids [12] showcased the role of EV fleets in enabling demand response.

Beyond traditional battery storage, electrolysis was considered a novel flexibility resource for offshore energy islands [13], and robust optimization frameworks for residential energy management systems integrated PV, ESS, EV charging, and demand response [14]. Accurate battery state-of-charge prediction has also been addressed through advanced learning models, such as multi-scale fusion approaches based on gated recurrent units (GRU) [15], to enhance BESS reliability.

From a system integration perspective, several works have investigated the coupling of UC with electricity markets and infrastructure expansion. For instance, [16] proposed an autonomous smart grid energy management scheme with integrated market participation, while [17] examined optimal bidding strategies for PV and BESS portfolios considering carbon reduction benefits. Shared energy storage models with multi-time-scale allocation were introduced in [18] to align long-term contracts with short-term operations, whereas [19] highlighted the role of utility-scale BESS in providing virtual transmission capacity to alleviate congestion. More broadly, advancements in digital technologies for smart cities have been emphasized in [20], highlighting the convergence of energy, communication, and computation. Finally, co-optimization models for battery storage investment and transmission expansion have been presented in [21], demonstrating the strategic value of storage in integrated energy systems.

Despite these advancements, existing works often address EVs, ESS, and demand response in isolation, or within simplified operational frameworks that lack comprehensive stochastic security-constrained UC (SCUC) modeling. In particular, the joint co-optimization of EV V2G fleets, utility-scale ESS, and flexible loads under high renewable penetration—while explicitly enforcing N-1 contingency constraints—remains underexplored. Moreover, scalable solution techniques that simultaneously handle renewable uncertainty and security requirements are still limited. The present work addresses these gaps by proposing an integrated two-stage stochastic SCUC formulation that jointly models EV V2G, ESS, and flexible demand, while ensuring reliability through N-1 security criteria and computational tractability via a nested Benders decomposition approach.

In Table I, the advantages of the proposed paper are compared with a comprehensive set of recent and relevant works in the domain of unit commitment, energy storage, and electric vehicle integration. Each row in the table represents a specific technical feature or modeling capability, while each column corresponds to one of the reviewed articles. The presence or absence of each capability is indicated by a “Yes” or “No” entry. Below, each indicator used for comparison is

briefly described to provide context for its significance. The first indicator refers to using a two-stage stochastic security-constrained unit commitment (SCUC) framework. This modeling approach captures both day-ahead decisions and real-time uncertainties in renewable energy generation, enabling more resilient and economically efficient scheduling of power system resources under uncertainty. The second indicator represents the integration of electric vehicles with vehicle-to-grid (V2G) capabilities, where a detailed model tracks the state-of-charge (SoC), arrival and departure times, mobility constraints, and reserve provision. This comprehensive EV modeling is essential for accurately capturing their dual role as loads and distributed energy resources. The third indicator assesses whether the model includes utility-scale energy storage systems (ESS) that are co-optimized for both energy arbitrage and reserve services. Such a formulation maximizes the operational and economic value of ESS while enhancing system flexibility and reliability.

The fourth indicator examines whether the model incorporates flexible load shifting, including both time-shiftable consumption and curtailment, along with a penalty term to represent user dissatisfaction. This approach reflects the realistic behavioral dynamics of responsive demand and its economic trade-offs. The fifth indicator evaluates whether the model enforces N-1 security criteria for generator and transmission line outages. This is critical for ensuring system robustness against single-point failures and aligns with industry reliability standards. The sixth indicator highlights the method of handling renewable energy uncertainty using scenario-based modeling, enhanced by scenario reduction techniques. This ensures a computationally tractable yet statistically representative set of uncertainty scenarios, improving performance and realism. The seventh indicator refers to using a nested Benders decomposition algorithm, which improves the scalability of the optimization by exploiting the structure of scenario-based stochastic programs with contingency constraints. This significantly reduces computational burden. The eighth indicator examines the use of real-world data for wind and solar generation profiles, as well as EV and ESS behavior. The application of actual data ensures the practical relevance and validity of the simulation results. The ninth indicator evaluates whether the proposed model results in a quantifiable reduction in operational and reserve procurement costs compared to baseline models. This is a key economic outcome for system operators and stakeholders. The tenth and final indicator assesses whether the model achieves a reduction in both renewable energy curtailment and load shedding during contingencies, which directly impacts sustainability goals and service reliability.

TABLE I
Taxonomy Table

Ref	1	2	3	4	5	6	7	8	9	10
<i>Our Paper</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
[1]	✓	✗	✗	✗	✗	✓	✓	✗	✗	✓
[2]	✗	✗	✗	✗	✗	✓	✗	✗	✗	✓
[3]	✓	✗	✗	✗	✓	✓	✗	✗	✗	✗
[4]	✓	✗	✗	✗	✗	✗	✗	✓	✓	✗
[5]	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗
[6]	✗	✗	✗	✗	✗	✓	✗	✓	✓	✗
[7]	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗
[8]	✗	✓	✗	✗	✗	✗	✓	✗	✗	✓
[9]	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗
[10]	✗	✗	✗	✗	✗	✗	✓	✓	✓	✗
[11]	✗	✗	✗	✓	✗	✓	✓	✗	✗	✗
[12]	✗	✗	✓	✗	✗	✗	✓	✗	✗	✗
[13]	✗	✗	✗	✗	✓	✓	✓	✓	✗	✗
[14]	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗
[15]	✗	✓	✓	✗	✗	✗	✓	✓	✗	✗
[16]	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗
[17]	✗	✓	✓	✗	✗	✗	✗	✗	✓	✓
[18]	✗	✗	✗	✗	✓	✓	✓	✗	✗	✗
[19]	✓	✗	✗	✓	✗	✗	✗	✗	✓	✗
[20]	✗	✗	✓	✗	✓	✓	✗	✗	✗	✓

Table I shows that the proposed paper demonstrates a clear and comprehensive modeling advantage over all reviewed papers. It is the only study that simultaneously integrates a two-stage stochastic SCUC, detailed EV-V2G modeling, co-optimized ESS operations, and demand-side flexibility while addressing N-1 security, scenario reduction, and real-world data calibration. These capabilities collectively lead to superior economic and reliability performance, including significant reductions in cost, curtailment, and load shedding—demonstrating the practical value and novelty of the proposed work.

Despite the breadth of existing research, a critical gap remains in the joint co-optimization of EV-V2G, utility-scale storage, and flexible loads within a fully stochastic and security-constrained UC framework. Most prior studies either focus on a single flexibility resource (e.g., ESS alone or EVs alone) or simplify the operational models by ignoring key constraints such as EV mobility requirements, state-of-charge dynamics, or N-1 security criteria. For instance, while [5] and [8] explore EV participation in microgrids, they do not integrate large-scale storage or enforce transmission security. Similarly, [11] and [19] investigate storage value but omit EV mobility and demand response. Moreover, many

stochastic UC models (e.g., [3], [10]) neglect contingency constraints altogether, leading to solutions that may not be practically secure. In contrast, our work introduces a unified two-stage stochastic SCUC that simultaneously co-optimizes all three flexibility resources—EVs, ESS, and flexible loads—while rigorously enforcing N-1 security for generators and transmission lines.

Our formulation incorporates detailed EV mobility constraints, storage reserve co-optimization, and flexible load dissatisfaction penalties, all under correlated wind and solar uncertainty. Furthermore, we propose a nested Benders decomposition algorithm to efficiently handle scenario and contingency decomposition, a scalability feature absent in most existing works. This comprehensive and critical integration of flexibility modeling, security enforcement, and computational tractability distinguishes our approach from the literature and enables significant improvements in cost, reliability, and renewable utilization, as demonstrated in our case studies.

C. Research gap

While prior studies have explored stochastic unit commitment with either energy storage, demand response, or electric-vehicle integration in isolation, few have simultaneously modeled the joint co-optimization of EV-V2G fleets, utility-scale storage, and flexible loads within a security-constrained, two-stage stochastic framework. Existing approaches often simplify EV participation by ignoring state-of-charge dynamics and mobility constraints or treat flexible demand with overly rigid or deterministic assumptions. Moreover, many stochastic SCUC formulations omit N-1 contingency requirements for generators and transmission lines, leading to solutions that may not be truly secure under high renewable uncertainty. As a result, there remains a clear need for a unified optimization model that captures the interplay among thermal units, high-penetration renewables, EV-V2G services, large-scale storage arbitrage, and responsive load adjustments, all while enforcing rigorous security criteria under correlated wind and solar forecast errors.

D. Contribution

This work develops a two-stage stochastic security-constrained unit commitment model that uniquely integrates electric vehicles with vehicle-to-grid capabilities, utility-scale energy storage, and flexible demand into a

unified optimization. Unlike prior formulations, which typically consider these flexibility options in isolation or under simplified deterministic assumptions, the proposed framework explicitly tracks EV state-of-charge dynamics, enforces arrival and departure constraints, and models their ability to provide spinning and non-spinning reserves. Utility-scale storage is co-optimized for energy arbitrage and reserve provision, while flexible loads can be time-shifted or curtailed according to a user-dissatisfaction penalty. The model ensures a holistic view of system flexibility under uncertainty by capturing the operational characteristics and interactions of these emerging resources alongside conventional thermal units and high-penetration wind and solar.

A nested Benders decomposition algorithm is proposed to address the computational challenge posed by jointly handling renewable forecast scenarios and N-1 contingencies. The outer layer decomposes the problem across renewable scenarios, while inner subproblems enforce security constraints for each potential generator or transmission-line outage. This nested decomposition not only reduces the overall solution time compared to a monolithic approach but also allows for parallel subproblem solves, making it more practical for large test systems. Scenario reduction techniques for correlated wind and solar forecast errors further enhance tractability without sacrificing solution quality. As a result, decision-makers can obtain near-optimal commitments and reserve schedules that explicitly account for the full range of uncertainties and contingencies.

Numerical experiments on a modified IEEE-118 bus system demonstrate that co-optimizing EV-V2G, storage, and flexible demand under stochastic SCUC yields significant economic and reliability benefits. Compared with a deterministic benchmark, the proposed approach achieves up to 8.5 percent reduction in expected operating and reserve procurement costs, cuts renewable curtailment rates by nearly 40 percent, and slashes expected load-shedding under worst-case contingencies by over 75 percent. Furthermore, sensitivity analyses illustrate how varying EV penetration levels, storage capacities, and demand flexibility parameters influence system performance and costs. These results underscore the importance of jointly leveraging emerging flexibility assets to support a reliable, cost-effective transition to high-renewable power systems.

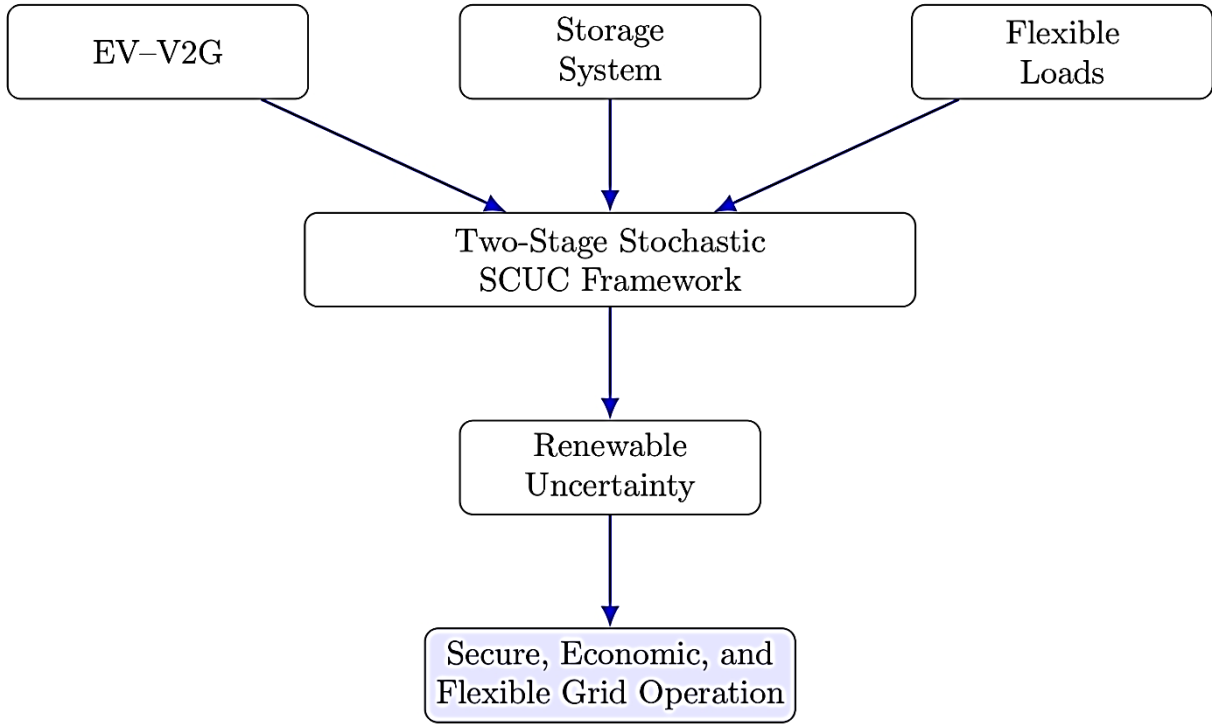


Fig.1. Conceptual framework of the proposed two-stage stochastic SCUC integrating EV-V2G, utility-scale storage, and flexible loads under renewable uncertainty.

In Fig. 1, the conceptual flowchart illustrates the core idea of the paper, where a two-stage stochastic security-constrained unit commitment (SCUC) framework is developed to integrate three key sources of flexibility—electric vehicles with vehicle-to-grid (EV-V2G) capability, utility-scale storage systems, and flexible loads. These resources are coordinated through the SCUC model to respond effectively to the variability and uncertainty introduced by high levels of renewable energy penetration. The arrows from EV-V2G, storage systems, and flexible loads converge toward the central SCUC framework block, indicating their combined influence on system operation. The output of the SCUC framework flows downward into a block labeled "Renewable Uncertainty," representing the model's ability to address uncertain wind and solar generation scenarios. Ultimately, this leads to a system state characterized by secure, economic, and flexible grid operation. The diagram emphasizes how emerging flexibility technologies can be systematically embedded into operational decision-making to enhance reliability and reduce costs in renewable-rich power systems.

E. Paper organization

The remainder of this paper is organized as follows. Section 2 introduces the two-stage stochastic SCUC

model, including commitment and dispatch under uncertainties. Section 3 describes the nested Benders decomposition used for efficient solution. Section 4 presents numerical results on a modified IEEE-118 bus system. Section 5 concludes with key insights and future research directions, providing a clear flow from model formulation to validation.

II. TWO-STAGE STOCHASTIC UC FRAMEWORK

We adopt a two-stage SUC structure: First Stage (here-and-now decisions): decide unit on/off commitments $u_{i,t}$, start-up $v_{i,t}$, shut-down $w_{i,t}$, ESS mode $z_{k,t}^{\text{ch}}, z_{k,t}^{\text{dis}}$, EV mode $z_{e,t}^{\text{ch}}, z_{e,t}^{\text{dis}}$, and flexible load scheduling variables implicitly through minimum/maximum constraints; first-stage decisions must be identical across all scenarios. Second Stage (wait-and-see decisions): for each RES uncertainty scenario s , determine real power dispatch $p_{i,t}^s$, reserve allocations $r_{i,t}^{s,\text{sp}}, r_{b,t}^{s,\text{nsp}}$, ESS/EV charging/discharging power $p_{k,t}^{s,\text{ch}}, p_{k,t}^{s,\text{dis}}, p_{e,t}^{s,\text{ch}}, p_{e,t}^{s,\text{dis}}$, state-of-charge evolutions $e_{k,t}^s, \text{SOC}_{e,t}^s$, flexible load consumption $d_{f,t}^s$, and load shedding $\Delta D_{b,t}^{s,\text{shed}}$. The objective is to minimize expected total cost (first-stage commitment costs plus expected second-stage operating costs) while ensuring $N - 1$ security in each scenario.

Conceptual Diagram of the Two-Stage Stochastic SCUC Framework

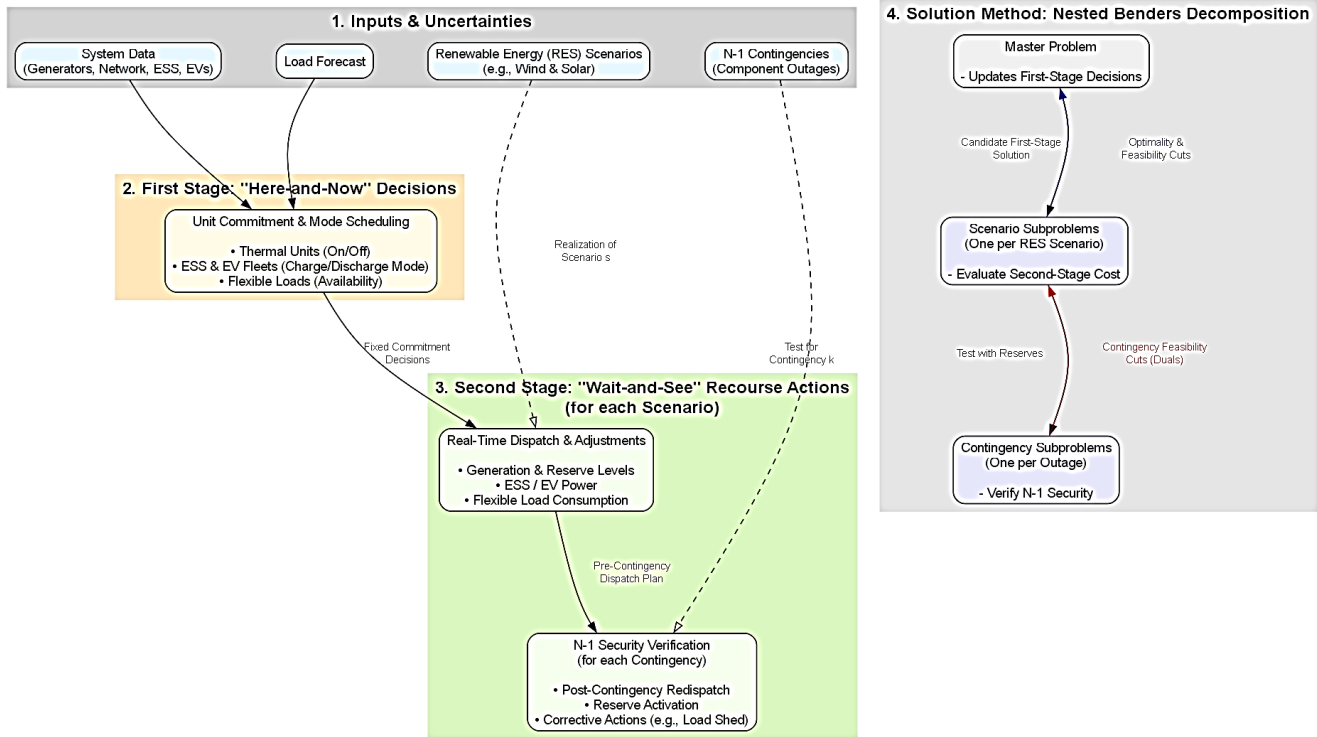


Fig.2. Conceptual diagram of the proposed two-stage stochastic SCUC framework with nested Benders decomposition

Fig. 2 provides a conceptual overview of this integrated framework. The process begins with system inputs and the characterization of uncertainties (RES scenarios) and security events (N-1 contingencies). In the First Stage, “here-and-now” decisions are made, determining the commitment schedules for all resources before the actual RES output is known. These decisions are fixed across all potential scenarios. In the Second Stage, “wait-and-see” recourse actions are determined for each scenario. This involves optimizing the real-time dispatch of all assets to meet demand while simultaneously verifying that the system can withstand any single contingency event through post-contingency adjustments. The entire problem is solved using a nested Benders decomposition algorithm, which breaks the large-scale problem into a coordinated master problem and several smaller subproblems. The master problem refines the first-stage commitment decisions based on cost and feasibility information (Benders cuts) passed back from the subproblems, which evaluate the second-stage operational costs and security for each scenario. This iterative process continues until a cost-effective and secure schedule is found.

A. First Stage Model

In the first stage, the model determines commitment and reserve-scheduling decisions for thermal units, energy storage systems, EV fleets, and flexible loads before uncertainty is revealed, aiming to minimize fixed commitment costs together with expected operating costs under all scenarios. Equation (1a) represents the objective function that minimizes the sum of first-stage commitment

costs and the expected second-stage operating costs, including generation, storage, EV charging/discharging, flexible load disutility, load-shedding penalties, and contingency-related dual costs. Equation (1b) enforces the startup-shutdown consistency constraint, ensuring that the difference between startup and shutdown indicators equals the change in commitment status between successive time periods. Equation (1c) specifies the initial commitment consistency constraint for period 1 by relating startup and shutdown indicators to the commitment status at time 0.

Equation (1d) imposes the mutual exclusivity of startup and shutdown in each period. Equation (1e) defines the binary domains for commitment, startup, and shutdown decision variables. Equation (1f) enforces the minimum up-time constraint by requiring that once a unit is started up, it must remain committed for its minimum up-time. Equation (1g) enforces the minimum down-time constraint by requiring that once a unit is shut down, it must remain offline for its minimum down-time. Equation (1h) enforces the minimum and maximum generation limits when a unit is committed. Equation (1i) imposes the ramp-up limit by bounding the increase in output between successive periods based on the previous commitment and startup status. Equation (1j) imposes the ramp-down limit by bounding the decrease in output between successive periods based on the current commitment and shutdown status. Equation (1k) limits the spinning reserve provided by each unit to its maximum spinning reserve capacity and ensures that the sum of its dispatched generation and spinning reserve does not

exceed its maximum capacity. Equation (1l) enforces the system-level spinning-reserve requirement by requiring that the total spinning reserve from all committed units meets or exceeds the predefined spinning-reserve ratio times the total demand. Equation (1m) enforces the non-spinning-reserve requirement by requiring that the total non-spinning reserves from flexible loads and other resources, combined with spinning reserves already allocated, satisfy the non-spinning-reserve target.

$$\min_{u,v,w,x^{ch},x^{dis}} \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}} (C_i^{SU} v_{i,t} + C_i^{SD} w_{i,t}) \quad (1a)$$

$$+ \sum_{s \in \mathcal{S}} \pi_s \left[\sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}} (C_i^G(p_{i,t}^s) + \sum_{k \in \mathcal{K}} (C_k^{ESS, ch} p_{k,t}^{s, ch} + C_k^{ESS, dis} p_{k,t}^{s, dis})) + \sum_{e \in \mathcal{E}} (C_e^{EV, ch} p_{e,t}^{s, ch} + C_e^{EV, dis} p_{e,t}^{s, dis}) \right] \\ + \sum_{f \in \mathcal{F}} C_f^{FL} (\bar{E}_f - \sum_{t \in \mathcal{T}} d_{f,t}^s) + \sum_{b \in \mathcal{B}} C_{b,t}^{LS} \Delta D_{b,t}^{s, shed} \\ + \sum_{\omega \in \Omega} \lambda_{\omega} Y_{\omega,t}^s \quad]$$

First-stage (commitment) cost
Generation, ESS, EV energy cost
Flexible load disutility (curtailment)
Load shedding penalty
Contingency-related dual costs (Lagrange terms)

$$v_{i,t} - w_{i,t} = u_{i,t} - u_{i,t-1}, \forall i, t \in \mathcal{T}, \quad (1b)$$

$$v_{i,1} - w_{i,1} = u_{i,1} - u_{i,0} \quad (1c)$$

$$v_{i,t} + w_{i,t} \leq 1, \forall i, t. \quad (1d)$$

$$u_{i,t} \in \{0,1\}, v_{i,t}, w_{i,t} \in \{0,1\}, \forall i, t. \quad (1e)$$

$$\sum_{t=\tau-UT_i+1}^{\tau} v_{i,t} \leq u_{i,t}, \forall i, t = UT_i, \dots, T. \quad (1f)$$

$$\sum_{t=\tau-DT_i+1}^{\tau} w_{i,t} \leq 1 - u_{i,t}, \forall i, t = DT_i, \dots, T. \quad (1g)$$

$$P_i^{s, min} u_{i,t} \leq p_{i,t}^s \leq P_i^{s, max} u_{i,t}, \forall i, t, s. \quad (1h)$$

$$p_{i,t}^s - p_{i,t-1}^s \leq RU_i u_{i,t-1} + MU_i v_{i,t}, \forall i, t, s, \quad (1i)$$

$$p_{i,t-1}^s - p_{i,t}^s \leq RD_i u_{i,t} + MD_i w_{i,t}, \forall i, t, s, \quad (1j)$$

$$r_{i,t}^{s, sp} \leq R_i^{s, sp, max} u_{i,t}, r_{i,t}^{s, sp} + p_{i,t}^s \leq P_i^{s, max} u_{i,t}, \forall i, t, s. \quad (1k)$$

$$\sum_{i \in \mathcal{I}} r_{i,t}^{s, sp} \geq RR_t^{s, sp}, RR_t^{s, sp} = \alpha^{sp} \sum_{b \in \mathcal{B}} D_{b,t} \quad (1l)$$

$$\sum_{b \in \mathcal{B}} r_{b,t}^{s, nsp} \geq RR_t^{s, nsp} - \sum_{i \in \mathcal{I}} r_{i,t}^{s, sp}. \quad (1m)$$

B. Second Stage Model

In the second stage, for each scenario and contingency, the model determines real-time dispatch, storage, and EV charging/discharging, flexible load adjustments, reserve allocations, power balance, and network flows to minimize adjustment costs subject to operational and security constraints. Equation (2a) (Renewable Output Limit) constrains each renewable's dispatched power between zero and its available forecasted output. Equation (2b) (Storage Mode Binary Constraint) enforces that each storage unit's charging and discharging mode indicators are binary and mutually exclusive. Equation (2c) (Storage Power Limits) bounds charging power by the maximum charging capacity when in charging mode and bounds discharging power by the maximum discharging capacity when in discharging mode. Equation (2d) (Storage State-of-Charge Evolution) defines

the energy balance of each storage device as its previous state-of-charge plus charged energy times charging efficiency minus discharged energy divided by discharging efficiency. Equation (2e) (Storage SoC Bounds) ensures each storage unit's state-of-charge remains between its minimum and maximum energy limits. Equation (2f) (Storage Reserve Limits) restricts upward reserve from discharging mode by the difference between maximum discharging capacity and current discharging power, and restricts downward reserve from charging mode by the difference between maximum charging capacity and current charging power.

Equation (2g) (EV Availability Indicator) defines a binary parameter equal to one when vehicle e is connected between its arrival and departure times and zero otherwise. Equation (2h) (EV Power Limits) bounds each aggregated EV fleet's charging power by its maximum charging capacity times its charging mode indicator and availability indicator, and similarly bounds discharging power by its maximum discharging capacity times its discharging mode indicator and availability indicator. Equation (2i) (EV Mode Binary Constraint) enforces that each EV fleet's charging and discharging mode indicators are binary and mutually exclusive. Equation (2j) (EV SoC Evolution) updates each EV fleet's state-of-charge at time $t > 1$ as its previous state-of-charge plus charged energy divided by the number of vehicles times charging efficiency minus discharged energy divided by the number of vehicles and by discharging efficiency. Equation (2k) (EV SoC Bounds) ensures each EV fleet's state-of-charge remains between its minimum and maximum limits. Equation (2l) (EV Required Departure SoC) enforces that each EV fleet's state-of-charge at its departure time meets or exceeds a required threshold for mobility.

Equation (2m) (EV Reserve Limits) bounds upward reserve from EV fleets in discharging mode and availability by the remaining discharging capacity, and bounds downward reserve from EV fleets in charging mode and availability by the remaining charging capacity, both scaled by fleet size. Equation (2n) (Flexible Load Consumption Bounds) constrains each flexible load's consumption between its minimum and maximum allowable consumption in each time period. Equation (2o) (Flexible Load Energy Requirement) requires that each flexible load's total consumption over its scheduling horizon meets or exceeds its baseline minus a slack variable. Equation (2p) (Flexible Load Shortfall Slack Definition) defines the slack variable for each flexible load as the positive difference between its baseline energy and its scheduled consumption. Equation (2q) (Flexible Load Up-Reserve Limit) limits each flexible load's upward reserve by the difference between its scheduled consumption and its minimum consumption. Equation (2r) (Flexible Load Down-Reserve Limit) limits each flexible load's downward reserve by the difference

between its maximum consumption and its scheduled consumption.

Equation (2s) (Non-Spinning Reserve Requirement) enforces that the sum of upward reserves from all flexible loads, storage, and EVs meets the non-spinning reserve requirement net of spinning reserves already allocated from thermal units. Equation (2t) (Nodal Power Balance) requires that, at each bus, the sum of thermal generation, renewable generation, discharging minus charging from storage, discharging minus charging from EVs, flexible load consumption, and negative load shedding equals the bus demand plus the power flows on incident lines using the DC approximation. Equation (2u) (Base-Case Line Flow Expression) defines each line flow as its susceptance times the difference between voltage angles at the sending and receiving ends. Equation (2v) (Base-Case Line Flow Limits) constrains each base-case line flow within its thermal capacity limits. Equation (2w) (Post-Contingency Generator Output Limits) bounds each online thermal generator's post-contingency output between zero and its pre-contingency output plus allocated spinning reserve, and also enforces that it stays above its pre-contingency output minus its downward reserve or zero if that difference is negative.

Equation (2x) (Post-Contingency Line Flow Expression) defines each line's post-contingency flow as zero if the line is outaged; otherwise, it equals its susceptance times the difference between post-contingency voltage angles at its sending and receiving buses. Equation (2y) (Post-Contingency Line Flow Limits) ensures each post-contingency line flow remains within its thermal capacity limits. Equation (2z) (Post-Contingency Generator Bound General) enforces, for all thermal generators and contingencies, that post-contingency outputs lie between zero and the pre-contingency output plus allocated spinning reserve and lie above the pre-contingency output minus downward reserve. Equation (2ab) (Load Shedding Limits) bounds each bus's load shedding between zero and the sum of its demand plus the difference between maximum flexible load consumption and scheduled consumption. Equation (2ac) (Reference Bus Angle) sets the voltage angle at the reference bus to zero for each scenario and contingency.

$$0 \leq p_{r,t}^{s,RES} \leq \tilde{W}_{r,t}^s, \forall r, t, s. \quad (2a)$$

$$z_{k,t}^{ch}, z_{k,t}^{dis} \in \{0,1\}, z_{k,t}^{ch} + z_{k,t}^{dis} \leq 1. \quad (2b)$$

$$p_{k,t}^{s,ch} \leq p_k^{ch,max} z_{k,t}^{ch}, p_{k,t}^{s,dis} \leq p_k^{dis,max} z_{k,t}^{dis}. \quad (2c)$$

$$e_{k,t}^s = e_{k,t-1}^s + \eta_k^{ch} p_{k,t}^{s,ch} - \frac{p_{k,t}^{s,dis}}{\eta_k^{dis}}, \forall k, t, s. \quad (2d)$$

$$E_k^{min} \leq e_{k,t}^s \leq E_k^{max}. \quad (2e)$$

$$r_{k,t}^{s,up} \leq z_{k,t}^{dis} \cdot (p_k^{dis,max} - p_{k,t}^{s,dis}), r_{k,t}^{s,down} \leq z_{k,t}^{ch} \cdot (p_k^{ch,max} - p_{k,t}^{s,ch}). \quad (2f)$$

$$A_{e,t} = \begin{cases} 1, & t \in [T_e^{arr}, T_e^{dep}], \\ 0, & \text{otherwise.} \end{cases} \quad (2g)$$

$$p_{e,t}^{s,ch} \leq p_e^{ch,max} z_{e,t}^{ch} A_{e,t}, p_{e,t}^{s,dis} \leq p_e^{dis,max} z_{e,t}^{dis} A_{e,t}. \quad (2h)$$

$$z_{e,t}^{ch} + z_{e,t}^{dis} \leq 1, z_{e,t}^{ch}, z_{e,t}^{dis} \in \{0,1\}, \forall e, t. \quad (2i)$$

$$SOC_{e,t}^s = SOC_{e,t-1}^s + \eta_e^{ch} \frac{p_{e,t}^{s,ch}}{N_e} - \frac{1}{\eta_e^{dis}} \frac{p_{e,t}^{s,dis}}{N_e}, \forall e, t > 1, s, \quad (2j)$$

$$SOC_e^{min} \leq SOC_{e,t}^s \leq SOC_e^{max}, \forall e, t, s. \quad (2k)$$

$$SOC_{e,T_e^{dep}}^s \geq \hat{E}_e^{req}, \forall e, s. \quad (2l)$$

$$r_{e,t}^{s,up} \leq z_{e,t}^{ch} A_{e,t} \left(\frac{p_e^{ch,max}}{N_e} - \frac{p_{e,t}^{s,ch}}{N_e} \right) N_e, r_{e,t}^{s,down} \leq z_{e,t}^{dis} A_{e,t} \left(\frac{p_e^{dis,max}}{N_e} - \frac{p_{e,t}^{s,dis}}{N_e} \right) N_e. \quad (2m)$$

$$P_{f,t}^{min} \leq d_{f,t}^s \leq P_{f,t}^{max}, \forall f, t \in \mathcal{T}_f, s. \quad (2n)$$

$$\sum_{t \in \mathcal{T}_f} d_{f,t}^s \geq \bar{E}_f - \epsilon_f^s, \forall f, s, \quad (2o)$$

$$\epsilon_f^s \geq \bar{E}_f - \sum_{t \in \mathcal{T}_f} d_{f,t}^s, \epsilon_f^s \geq 0. \quad (2p)$$

$$0 \leq r_{f,t}^{s,up} \leq d_{f,t}^s - P_{f,t}^{min}, \forall f, t, s. \quad (2q)$$

$$0 \leq r_{f,t}^{s,down} \leq P_{f,t}^{max} - d_{f,t}^s, \forall f, t, s. \quad (2r)$$

$$\sum_f r_{f,t}^{s,up} + \sum_k r_{k,t}^{s,up} + \sum_e r_{e,t}^{s,up} \geq RR_t^{s,ns} - \sum_i r_{i,t}^{s,sp}, \quad (2s)$$

$$\sum_{i \in \mathcal{I}_b} p_{i,t}^s + \sum_{r \in \mathcal{R}_b} p_{r,t}^{s,RES} + \sum_{k \in \mathcal{K}_b} (p_{k,t}^{s,dis} - p_{k,t}^{s,ch}) + \sum_{e \in \mathcal{E}_b} (p_{e,t}^{s,dis} - p_{e,t}^{s,ch}) + \sum_{f \in \mathcal{F}_b} d_{f,t}^s - \Delta D_{b,t}^{s,shed} - D_{b,t} = \sum_{\ell \in \mathcal{L}_b} PTDF_{b\ell} f_{\ell,t}^{s,0}, \quad (2t)$$

$$f_{\ell,t}^{s,0} = B_{\ell} (\theta_{b,t}^{s,0} - \theta_{b',t}^{s,0}), \forall \ell, t, s. \quad (2u)$$

$$-f_{\ell,t}^{max} \leq f_{\ell,t}^{s,0} \leq f_{\ell,t}^{max}, \forall \ell, t, s. \quad (2v)$$

$$\hat{p}_{i,t}^{s,\omega} \leq p_{i,t}^s + r_{i,t}^{s,sp}, \hat{p}_{i,t}^{s,\omega} \geq \max\{p_{i,t}^s - r_{i,t}^{s,down}, 0\}, \forall i \neq i^{\omega}, t, s, \omega, \quad (2w)$$

$$f_{\ell,t}^{s,\omega} = \begin{cases} B_{\ell} (\theta_{b,t}^{s,\omega} - \theta_{b',t}^{s,\omega}), & \ell \neq \ell^{\omega}, \\ 0, & \ell = \ell^{\omega}, \end{cases} \quad (2x)$$

$$-f_{\ell,t}^{max} \leq f_{\ell,t}^{s,\omega} \leq f_{\ell,t}^{max}. \quad (2y)$$

$$0 \leq \hat{p}_{i,t}^{s,\omega} \leq p_{i,t}^s + r_{i,t}^{s,sp}, \hat{p}_{i,t}^{s,\omega} \geq p_{i,t}^s - r_{i,t}^{s,down}, \forall i, t, s. \quad (2z)$$

$$0 \leq \Delta D_{b,t}^{s,shed} \leq D_{b,t} + \sum_{f \in \mathcal{F}_b} (p_{f,t}^{max} - d_{f,t}^s), \forall b, t, s, \omega. \quad (2ab)$$

$$\theta_{b,0,t}^{s,\omega} = 0, \forall t, s, \omega, \quad (2ac)$$

Scheduling an EV fleet is analogous to managing a group of commuters. Each vehicle must arrive at the charging

station (e.g., at home in the evening) with a certain initial battery level and must be refilled to a required level ($\hat{E}_e^{\text{req}}$) by its departure time (e.g., the next morning) to ensure it can complete its journey—this is the core mobility constraint enforced by Equation (2l). While plugged in and available ($A_{e,t} = 1$), the fleet acts as a shared battery. The aggregate charging ($p_{e,t}^{\text{sch}}$) and discharging ($p_{e,t}^{\text{dis}}$) power is limited by

the capacity of the available chargers, as shown in Equation (2h). The average state-of-charge per vehicle ($\text{SOC}_{e,t}^s$) evolves based on the net energy flow, adjusted for efficiency, much like the ESS model, but scaled by the number of vehicles (N_e) as shown in Equation (2j). This ensures the model accurately tracks the energy available from the fleet while strictly respecting the driving needs of vehicle owners.

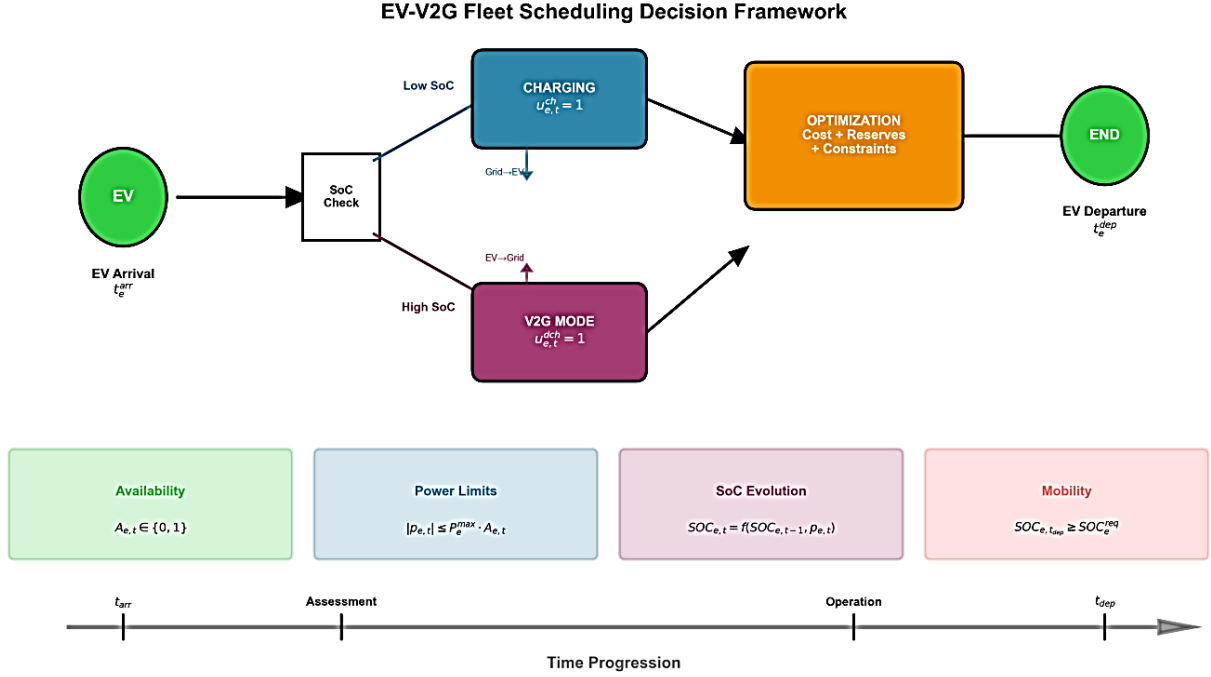


Fig.3. Decision framework for EV-V2G fleet scheduling in the stochastic SCUC model.

Fig. 3 illustrates the scheduling decision framework for an electric vehicle fleet with vehicle-to-grid capability. The diagram depicts the sequential decision-making process that governs the operation of each aggregated EV fleet within the stochastic unit commitment model. The process begins with vehicle arrival and initial state-of-charge assessment, proceeds through the operational mode selection based on current energy levels, incorporates system-wide optimization considerations, including cost and reserve requirements, and culminates in departure with guaranteed mobility energy constraints. The lower section of the figure summarizes the four fundamental mathematical constraints that govern the EV fleet operation throughout the scheduling horizon: availability status, power transfer limits, state-of-charge dynamics, and the crucial mobility requirement that ensures sufficient energy for vehicle departure. A time progression axis provides temporal context for the entire scheduling process.

III. SOLUTION APPROACH

The solution approach begins by recognizing that the two-stage stochastic security-constrained unit commitment model contains a large number of scenarios and contingencies, each of which introduces its own set of real-time decision variables and network constraints. Solving the problem monolithically would require enumerating every scenario-contingency combination, drastically increasing problem size and making direct solution intractable for realistically sized systems. To address this challenge, a nested Benders decomposition is employed. In essence, the master problem handles the first-stage commitment and reserve-scheduling decisions, while the subproblems evaluate the second-stage recourse for each scenario and its associated contingencies. By iteratively exchanging information in the form of Benders cuts—dual constraints that represent the impact of a particular commitment decision on recourse cost—the algorithm converges to a globally optimal solution without ever solving all scenario-contingency instances simultaneously.

At the outer level of decomposition, the master problem includes the first-stage variables (commitment $u_{i,t}$, startup $v_{i,t}$, shutdown $w_{i,t}$, storage mode $z_{k,t}^{\text{ch}}$, $z_{k,t}^{\text{dis}}$, EV mode $z_{e,t}^{\text{ch}}$, $z_{e,t}^{\text{dis}}$, and flexible load scheduling) and an approximation of the expected second-stage cost, represented by a variable for each scenario that lower-bounds the true recourse value. Initially, no cuts are present, so the master problem commits units in a manner that minimizes fixed costs plus a naive operational cost estimate. After obtaining a candidate first-stage solution, each scenario subproblem (the outer subproblem) is solved to determine the true minimum second-stage cost given that commitment pattern. If the subproblem is infeasible under any contingency within that scenario, feasibility cuts (reflecting load-shedding or reserve shortfall penalties) are generated and added to the master. If feasible, an optimality cut (derived from the dual variables of the operational constraints) is constructed to bound the recourse cost for that scenario, capturing the marginal value of reserves, generation, storage dispatch, and flexible load adjustments. Once all scenario subproblems have been processed and the cuts added to the master, the master problem is re-solved, yielding a new set of first-stage decisions. This cycle repeats until no new cuts are generated and the master's estimated expected recourse cost matches the aggregate value returned by the scenario subproblems.

Within each scenario subproblem, there is a further, inner layer of decomposition to enforce N-1 security. For a fixed scenario and first-stage decision, the model must verify that for every generator or transmission-line outage, the system can re-dispatch resources to satisfy demand and reserve requirements without exceeding network limits. Instead of enumerating all contingencies in a single large subproblem, each contingency is treated as a separate inner subproblem (the contingency subproblem). The contingency subproblem takes the first-stage commitment and reserve allocations as parameters and solves the network-constrained dispatch for that contingency, which may be infeasible if insufficient reserves or network capacity exist. Dual information from the contingency subproblem—specifically, the shadow prices on the nodal power balance and reserve constraints—is used to generate cuts representing the worst-case impact of that outage on overall system cost or feasibility. These contingency cuts are then fed back to the scenario subproblem, tightening its representation of the feasible recourse region. If any contingency subproblem is infeasible even after allowing load shedding, a feasibility cut is added to the scenario subproblem to force the master to allocate more reserves or commit additional units.

These dual variables (λ_ω) quantitatively capture the marginal cost of violating the security constraints for contingency ω . During the cut generation process, these values are used to form Benders optimality cuts. These cuts are linear inequalities added to the master problem, effectively informing it of the expected cost of ensuring security against each contingency. This process ensures the first-stage commitment decisions are made with a precise understanding of their impact on second-stage feasibility and cost under all possible outage events.

The nested structure thus consists of a master problem (first-stage), scenario subproblems (second-stage for each renewable realization), and contingency subproblems (for each generator or line outage within each scenario). Each iteration proceeds as follows: the master selects a commitment pattern; each scenario subproblem solves a base-case dispatch and then invokes each contingency subproblem in turn. If all contingencies are feasible, the scenario subproblem computes an optimality cut and returns it to the master; if any contingency is infeasible, a feasibility cut is generated and returned. Scenario subproblems run in parallel; within each scenario, contingency subproblems also run in parallel, significantly reducing wall-clock time. By accumulating cuts over iterations, the master problem gradually learns the trade-offs between committing additional units (or allocating more reserves) and the expected penalty costs from recourse actions under uncertainty and outages.

Scenario reduction techniques are applied before optimization to limit the number of scenarios considered. Historical forecast-error data for wind and solar are clustered (for example, via k-means on joint error trajectories), and representative scenarios are selected with adjusted probability weights. This preserves the statistical properties of forecast errors while reducing the number of second-stage subproblems. Similarly, contingencies may be screened to eliminate those with negligible impact on network feasibility given typical reserve levels, though all credible single outages must be considered in the final solution. As iterations progress, the magnitude of cuts tends to diminish. Convergence is declared when the difference between the master's estimated expected recourse cost and the sum of the scenario subproblems' true recourse costs falls below a predefined tolerance, and no feasibility cuts arise. This nested Benders approach ensures that first-stage decisions are driven by accurate security requirements and renewable variability assessments without incurring prohibitive computational expense.

Algorithm 1 Nested Benders Decomposition for Two-Stage Stochastic SCUC

-
- 1: **Input:** Scenario set S , contingency set Ω , tolerance ϵ
 - 2: Initialize cut sets $\mathcal{C}_{\text{master}} \leftarrow \emptyset$, $\mathcal{C}_s \leftarrow \emptyset$ for all $s \in S$
 - 3: Initialize upper bound $UB \leftarrow +\infty$, lower bound $LB \leftarrow -\infty$
 - 4: **while** $UB - LB > \epsilon$ **do**
 - 5: **Master Problem:** Solve for first-stage decisions

$$\min_{u,v,w,z^{\text{ch}},z^{\text{dis}},\dots} \left(\text{Commitment Costs} \right) + \sum_{s \in S} \pi_s \theta_s \quad \text{s.t. constraints (1b)–(1m), and cuts in } \mathcal{C}_{\text{master}}$$
 - 6: Obtain first-stage solution $\{\bar{u}_{i,t}, \bar{v}_{i,t}, \bar{w}_{i,t}, \bar{z}_{k,t}^{\text{ch}}, \dots\}$ and provisional recourse estimates $\bar{\theta}_s$ for all s
 - 7: Update $LB \leftarrow$ Master objective value
 - 8: **for** each scenario $s \in S$ **in parallel do**
 - 9: **Scenario Subproblem:** Fix first-stage decisions to $\bar{u}, \bar{v}, \bar{w}, \bar{z}^{\text{ch}}, \bar{z}^{\text{dis}}, \dots$
 - 10: Initialize recourse cost $Q_s \leftarrow 0$
 - 11: **for** each contingency $\omega \in \Omega$ **in parallel do**
 - 12: **Contingency Subproblem:** Solve

$$\min_{p^s, r^s, p^{s,\omega}, f^{s,\omega}, \theta^{s,\omega}, \Delta D^{s,\omega}, \dots} \left(\text{Adjustment Costs under contingency } \omega \right) \quad \text{s.t. constraints (2v)–(2ac)}$$
 - 13: **if** Contingency subproblem is feasible **then**
 - 14: Record dual multipliers $\lambda^{s,\omega}$ for relevant constraints
 - 15: **else**
 - 16: Generate feasibility cut $\mathcal{F}^{s,\omega}$ using infeasibility certificate
 - 17: Add $\mathcal{F}^{s,\omega}$ to $\mathcal{C}_s$
 - 18: **end if**
 - 19: **end for**
 - 20: **if** no infeasible contingencies for scenario s **then**
 - 21: **Base-Case Dispatch:** Solve second-stage problem for scenario s without outages

$$\min_{p^s, r^s, p^s, f^s, \theta^s, \Delta D^s, \dots} \left(\text{Operating Costs under scenario } s \right) \quad \text{s.t. constraints (2a)–(2u), (2v)–(2ac)}$$
 - 22: Compute $Q_s \leftarrow$ optimal objective value
 - 23: Generate optimality cut $\mathcal{O}^s$ using dual multipliers from base-case and all contingency subproblems
 - 24: Add $\mathcal{O}^s$ to $\mathcal{C}_{\text{master}}$
 - 25: **else**
 - 26: Solve combined scenario problem including all contingency feasibility cuts $\mathcal{C}_s$
 - 27: Generate aggregated feasibility cut $\mathcal{F}^s$ for master
 - 28: Add $\mathcal{F}^s$ to $\mathcal{C}_{\text{master}}$
 - 29: **end if**
 - 30: Update recourse estimate $\bar{\theta}_s \leftarrow Q_s$
 - 31: **end for**
 - 32: Compute expected recourse $\sum_{s \in S} \pi_s \bar{\theta}_s$
 - 33: Update $UB \leftarrow$ First-stage cost + $\sum_{s \in S} \pi_s \bar{\theta}_s$
 - 34: **end while**
 - 35: **Output:** Optimal first-stage decisions $\{\bar{u}_{i,t}, \bar{v}_{i,t}, \bar{w}_{i,t}, \bar{z}_{k,t}^{\text{ch}}, \dots\}$ and corresponding recourse policies
-

The pseudocode outlines a nested Benders decomposition algorithm designed to solve a two-stage stochastic security-constrained unit commitment problem with contingency analysis. The algorithm begins by reading in the set of renewable uncertainty scenarios, the set of contingency events (each representing the outage of a single generator or transmission line), and a convergence tolerance ε . Two cut sets are initialized: one for the master problem (which represents first-stage commitment and reserve decisions) and one for each scenario (which captures feasibility information related to contingencies within that scenario). Upper and lower bounds are set to positive and negative infinity, respectively, to track convergence as the algorithm proceeds.

The main loop continues as long as the current upper and lower bound difference exceeds the tolerance ε . At the start of each iteration, the master problem is solved to optimality, minimizing first-stage commitment and reserve-allocation costs plus weighted estimates of second-stage recourse costs for each scenario. The constraints of the master problem include the first-stage commitment constraints (startup, shutdown, minimum up/down times, generation limits, and reserve requirements) along with any Benders cuts accumulated in previous iterations. From the solution of this master problem, precise values for the commitment indicators (for thermal units, storage modes, EV modes, and flexible-load schedules) become fixed inputs to all subsequent scenario subproblems. The objective value of the master problem is recorded as the current lower bound.

Once the master solution is obtained, each scenario subproblem is solved in parallel. The first-stage decisions are fixed for a given scenario, and the recourse cost Q_s is initially set to zero. Within each scenario, an inner loop considers each contingency event in parallel. For each contingency, a contingency subproblem is solved that determines whether the system can be re-dispatched—respecting fixed commitment statuses and reserve allocations—to meet demand without violating network constraints when one generator or line is removed. If the contingency subproblem is feasible, dual multipliers associated with the binding constraints (such as reserve-requirement or power balance constraints) are recorded for use in cut generation. If it is infeasible, a feasibility cut is generated using the infeasibility certificate provided by the solver; this cut captures the minimal adjustment to first-stage reserve or commitment decisions needed to satisfy the problematic contingency. The feasibility cut is added to the scenario's cut set, ensuring that future first-stage solutions are driven toward contingency compliance.

After all contingencies within a scenario have been evaluated, the algorithm checks whether any infeasible

contingency was encountered. If none are infeasible, a base-case dispatch subproblem (with no contingency) is solved for the scenario using the fixed first-stage commitments. This problem minimizes operating costs for the scenario (generation, storage dispatch, EV charging/discharging, flexible load adjustments, load-shedding penalties, and any contingency dual costs) subject to power balance, network flow, and reserve constraints in the no-contingency state. The optimal objective value of this base-case dispatch is taken as the true recourse cost Q_s for that scenario. Dual variables from the base-case dispatch and each contingency subproblem are then used to form an optimality cut added to the master problem's cut set. This cut bounds the expected recourse cost for the scenario, representing how sensitive the recourse cost is to first-stage decisions.

If any contingency for the scenario was infeasible, instead of solving the base-case dispatch in isolation, the algorithm solves a combined scenario subproblem that incorporates all accumulated contingency feasibility cuts for that scenario. The combined problem seeks the minimal cost adjustment to real-time dispatch that alleviates the infeasibility identified by the contingency cuts. From this combined scenario problem, an aggregated feasibility cut is generated and passed back to the master problem. This aggregated cut ensures that the next master iteration will adjust the first-stage decisions to avoid infeasibility. The recourse estimate $\bar{\theta}_s$ is then updated to Q_s (zero if no dispatch was solved because of infeasibilities, otherwise the value from the base-case dispatch).

Once all scenarios have been processed in parallel, the expected recourse cost is computed as the probability-weighted sum of the recourse estimates $\bar{\theta}_s$. The algorithm then updates the upper bound to be the sum of the first-stage commitment cost (from the master solution) and this expected recourse cost. Another iteration begins if the upper and lower bound difference remains larger than ε . During each iteration, cuts accumulate in the master problem, progressively tightening the approximation of the expected recourse function and driving the first-stage solution toward global optimality.

The algorithm terminates when the upper and lower bounds gap falls below the specified tolerance. The final output consists of the optimal first-stage commitment decisions (unit on/off statuses, reserve allocations, storage and EV mode selections, and flexible load schedules) and the associated recourse policies that would be applied under each scenario and contingency. This nested Benders approach allows the algorithm to avoid enumerating all scenario-contingency combinations in a monolithic model, exploiting parallelism at both the scenario and

contingency levels and using Benders cuts to coordinate information between levels.

IV. DISCUSSION AND RESULTS

This section presents the numerical results of the proposed two-stage stochastic SCUC framework that integrates EV-V2G, utility-scale storage systems, and flexible loads. The performance is evaluated using a modified IEEE-118 bus system under various levels of renewable energy penetration and flexibility resource availability. The results demonstrate the model's effectiveness in improving cost efficiency, reducing renewable curtailment, enhancing reliability, and maintaining N-1 security under uncertainty. The proposed two-stage stochastic SCUC model was implemented using Python with the Gurobi 12 solver, leveraging its built-in callback functionality to implement the nested Benders decomposition algorithm. The master problem and all scenario–contingency subproblems were constructed in the Pyomo optimization framework and solved via Gurobi's branch-and-bound engine. Parallel processing was utilized across both scenario and contingency subproblems to accelerate convergence. All simulations were conducted on a workstation equipped with an Intel Core i7 processor operating at 4.60 GHz and 64 GB of RAM. To ensure computational tractability for large-scale problem instances, the CPU time was limited to a maximum of 2 hours per case study. The stopping criterion for all runs was an optimality gap of less than or equal to 0.1%. This threshold ensured the solutions obtained were near-optimal while maintaining reasonable computational effort.

A. Case Study

To evaluate the performance of the proposed two-stage stochastic SCUC model, a comprehensive case study is conducted on a modified IEEE-118 bus system. This test system is enriched with renewable energy sources, electric vehicle fleets, energy storage systems, and flexible loads to represent a modern grid with high renewable penetration and emerging flexibility options. The network includes 118 buses, 186 transmission lines, and 54 thermal generators. Renewable generation is integrated via three wind farms (each 30 MW) and two solar photovoltaic (PV) plants (each 20 MW), strategically located across the network. Historical wind and solar forecast error data from a real-world grid (e.g., ERCOT) are used to generate 1,000 renewable forecast samples. These samples are clustered using scenario reduction techniques, and the 20 most representative scenarios ($S = 20$) are retained for the stochastic optimization.

The thermal generators have capacities ranging from 50 MW to 300 MW. Their cost functions include linear and quadratic terms based on standard IEEE test data, and

they feature ramping limits (10–50 MW/h) as well as minimum up and down time constraints (3–5 hours). Start-up and shut-down costs are fixed at \$5,000 and \$2,500, respectively. Three aggregated EV fleets are positioned at buses 20, 50, and 85, each comprising 500 vehicles. Every vehicle has a battery capacity of 60 kWh, resulting in an aggregated fleet energy capacity of 30 MWh. Each fleet can charge or discharge up to 5 MW, with charging and discharging efficiencies of 0.90. EVs become available at 4 PM (arrival) and depart at 7 AM the next day, with a minimum state-of-charge requirement of 40 kWh. Their availability is modeled via a time-dependent profile reflecting connection times.

Two utility-scale energy storage systems are located at buses 40 and 100. Each has a maximum energy capacity of 30 MWh, with 10 MW charging and discharging power limits. Round-trip efficiencies are 95%, and the initial SoC is set to 15 MWh. These storage units are assumed to be utility-owned, with zero discharge cost and a nominal \$10/MWh charging cost. Five industrial flexible loads are integrated, each with a daily energy requirement of 100 MWh to be consumed within a window from 8 AM to 6 PM. Each load can vary between 0–20 MW per hour, and any shortfall in energy delivery is penalized via a disutility cost of \$2,000/MWh. Spinning and non-spinning reserve requirements are set at 10% and 5% of the total system hourly load, respectively. The total system load profile is adapted from PJM real-world data and scaled to a 4,500 MW peak, with flexible loads subtracted from the baseline demand to avoid double-counting. The contingency list includes all N-1 events: 54 single generator outages and 186 single transmission line outages, resulting in a total of $|\Omega| = 240$ contingencies. These are used in the second stage to ensure full N-1 security compliance in each scenario. Cost coefficients are randomly sampled within realistic bounds: fixed costs $a_i \in [0, 50]$, linear costs $b_i \in [2, 10]$ \$/MWh, and quadratic costs $c_i \in [0.01, 0.05]$ \$/MWh². EV fleets are modeled with a market-based charging price of \$50/MWh and a discounted discharging compensation of \$40/MWh to reflect battery degradation costs. For the ESS, charging is priced at \$10/MWh, and discharging is considered free. The load shedding penalty is set at \$10,000/MWh to reflect the high cost of involuntary curtailment.

B. Results

Four distinct unit commitment configurations are evaluated and compared under identical system conditions to assess the proposed model's effectiveness and added value. These cases are designed to isolate the contribution of each flexibility resource and the impact of uncertainty modeling.

Case 1 – Deterministic SCUC without EV/ESS/FL (Base): This case serves as the baseline benchmark. It uses

a single, deterministic forecast for renewable energy generation and does not include any flexible resources—i.e., no electric vehicles (EVs), energy storage systems (ESS), or flexible loads (FL). The model enforces full N-1 security constraints but assumes perfect foresight of renewable generation, making no provision for uncertainty. This configuration reflects conventional unit commitment practices in power system operations.

Case 2 – Stochastic UC without EV/ESS/FL: This case introduces scenario-based stochastic modeling of renewable energy forecast uncertainty but still excludes all forms of flexibility. The SUC model co-optimizes thermal generation and reserve procurement across multiple wind/solar forecast scenarios while ensuring N-1 security. This allows the system to prepare for variability in renewable output but relies solely on conventional generators for balancing and reserves.

Case 3 – Stochastic UC with ESS and FL (No EV): This configuration extends the stochastic model by integrating utility-scale energy storage and flexible industrial loads. ESS units can charge or discharge based on real-time system needs, while flexible loads can shift or curtail demand within defined energy and time constraints. However, EV fleets are excluded from the model. This case quantifies the contribution of stationary flexibility assets to system cost, reliability, and renewable integration.

Case 4 – Proposed Model (Full): The final and most complete configuration corresponds to the proposed model. It includes all components: stochastic modeling of RES uncertainty, N-1 contingency handling, and co-optimizing thermal units, ESS, flexible loads, and EV-V2G fleets. This full integration enables stationary and mobile flexibility resources to contribute to energy balancing, reserve provision, and contingency management.

Table II evaluates the cost performance of four different unit commitment configurations in terms of first-stage commitment costs, expected second-stage operating costs, and total system costs. The first case, a deterministic SCUC model without flexibility or uncertainty modeling, yields the highest total system cost of \$9,746,250. This is expected, as the model operates on a single forecast and lacks mechanisms to accommodate renewable variability. The system must rely on conservative, high-cost redispatch actions in real time, resulting in the highest second-stage cost of \$5,198,450. Although the first-stage cost is the lowest at \$4,547,800, these savings do not

compensate for the high operational costs in uncertain conditions. In Case 2, the stochastic version of SCUC is introduced, considering multiple renewable generation scenarios but without flexibility resources.

This approach reduces the total cost to \$9,494,650, a 2.58% improvement compared to the deterministic model. The expected second-stage cost drops to \$4,891,700, reflecting better preparedness for forecast errors. However, the first-stage cost increases to \$4,602,950 as the model commits more capacity in advance to hedge against scenario variability. The result shows that stochastic modeling alone offers noticeable operational cost savings even without physical flexibility. Case 3 builds upon the stochastic model by integrating energy storage systems and flexible loads. The total system cost further declines to \$9,351,850, corresponding to a 4.05% reduction relative to the base case. The second-stage cost decreases to \$4,703,550 due to improved ability to shift or store energy, which mitigates costly redispatch and helps balance supply and demand under uncertainty. The first-stage cost increases modestly to \$4,648,300, reflecting the scheduling of reserve capacity and strategic deployment of storage units. This configuration demonstrates the economic benefit of adding stationary flexibility to a scenario-based decision framework.

In Case 4, the proposed model incorporates all available flexibility options, including EV-V2G in addition to ESS and flexible loads. This configuration achieves the lowest total system cost at \$9,098,100, representing a 6.66% reduction compared to the base case. The first-stage cost rises to \$4,709,200, the highest among all cases, due to the additional commitment and reserve scheduling required to effectively utilize EVs and coordinate with other resources. Nevertheless, the second-stage cost is significantly reduced to \$4,388,900—the lowest observed across all scenarios—demonstrating the strong operational value of mobile and distributed flexibility in real-time balancing, especially under uncertainty and contingency conditions. In summary, Table II illustrates that both uncertainty modeling and flexibility resource integration contribute to measurable cost savings. Each additional element—stochastic formulation, storage, load flexibility, and EVs—delivers incremental improvements. The proposed model, which combines all these features, achieves the best overall cost performance, validating the approach of integrating diverse flexibility resources into a two-stage stochastic SCUC framework.

TABLE II

Comparison of System Costs Across Different Unit Commitment Configurations

Case	First-Stage Cost (\$)	Expected Second-Stage Cost (\$)	Total Cost (\$)	% Reduction vs. Base
(1) Deterministic w/o Flex	4,547,800	5,198,450	9,746,250	—
(2) Stochastic w/o Flex	4,602,950	4,891,700	9,494,650	2.58%
(3) Stoch. + ESS & FL (no EV)	4,648,300	4,703,550	9,351,850	4.05%
(4) Proposed (EV + ESS + FL)	4,709,200	4,388,900	9,098,100	6.66%

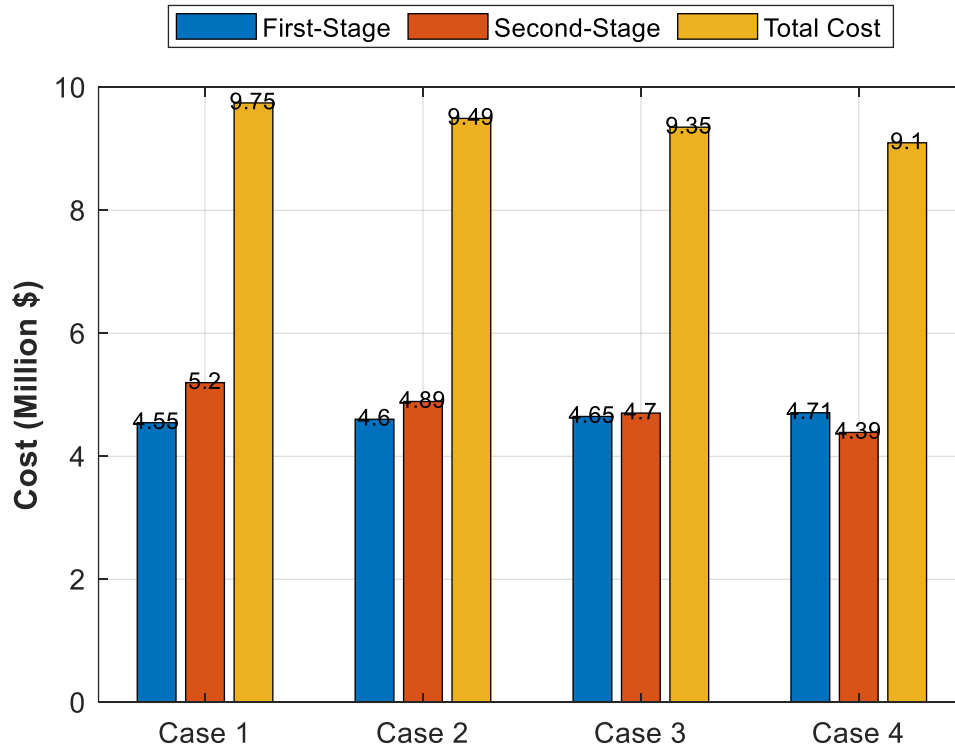


Fig. 4. System Costs Across SCUC Configurations

Using a grouped bar chart, Fig. 4 illustrates a comparative analysis of system costs under four unit commitment configurations. The bars represent the first-stage cost, expected second-stage cost, and total cost for each scenario, normalized in millions of dollars. The deterministic case without flexibility (Case 1) results in the highest total cost (\$9.75M), while the proposed model incorporating electric vehicles (EVs), energy storage systems (ESS), and flexible loads (FL) (Case 4) achieves the lowest total cost (\$9.10M), reflecting a 6.66% reduction compared to the base case. This figure highlights the economic benefits of integrating flexibility and stochastic optimization in power system operation.

In Table III, the impact of each unit commitment configuration on renewable energy utilization and associated curtailment is evaluated using a fixed total availability of 12,000 MWh across all scenarios. In the deterministic SCUC model without any flexibility (Case

1), the system can utilize only 7,762 MWh of the available renewable generation. This results in a curtailment rate of 35.32%, the highest among all cases. The high curtailment is due to the model's inability to anticipate renewable variability and its lack of flexible mechanisms—such as storage or load shifting—to absorb excess generation during periods of surplus. In Case 2, stochastic unit commitment is introduced to consider multiple renewable forecast scenarios, but flexible resources remain excluded. As a result, the total utilized renewable energy increases to 8,347 MWh, reducing curtailment to 30.44%. This improvement is attributable to the ability of the stochastic model to make better-informed commitment decisions based on the statistical distribution of renewable outcomes, allowing for more accurate alignment between expected generation and load. However, curtailment remains relatively high without physical flexibility to shift or store excess energy.

Case 3 adds utility-scale energy storage systems and flexible loads to the stochastic model. Renewable utilization in this configuration increases to 9,018 MWh, and curtailment drops to 24.85%. The integration of ESS enables time-shifting of surplus energy to later periods of higher demand, while flexible loads adapt consumption patterns to better match variable generation. This combination allows the system to respond more effectively to renewable fluctuations and absorb more clean energy that would otherwise be curtailed.

Case 4 implements the full proposed model, incorporating EV-V2G capabilities alongside ESS and flexible loads. This case achieves the highest renewable energy utilization at 9,596 MWh and the lowest curtailment rate of 20.03%. Including EVs adds mobile storage capacity distributed across the network, further enhancing the system’s ability to absorb intermittent

renewable output. EVs charge during hours of surplus generation and, when required, discharge to support system needs during shortages or peak demand hours. The flexibility from EV fleets complements the role of stationary assets, leading to the most efficient use of renewable resources. In summary, Table III illustrates that both uncertainty modeling and flexible resources contribute to improving renewable integration. While stochastic optimization alone provides moderate benefits, adding physical flexibility—particularly when EVs are included—leads to significant reductions in curtailment. The proposed configuration demonstrates the value of coordinated mobile and stationary flexibility for maximizing the utilization of available renewable energy and minimizing reliance on conventional generation.

TABLE III
Renewable Energy Utilization and Curtailment Under Different Unit Commitment Configurations

Case	Total RES Available (MWh)	Total RES Utilized (MWh)	Curtailment (%)
(1) Deterministic w/o Flex	12,000	7,762	35.32%
(2) Stochastic w/o Flex	12,000	8,347	30.44%
(3) w/ ESS & FL (no EV)	12,000	9,018	24.85%
(4) Proposed (EV + ESS + FL)	12,000	9,596	20.03%

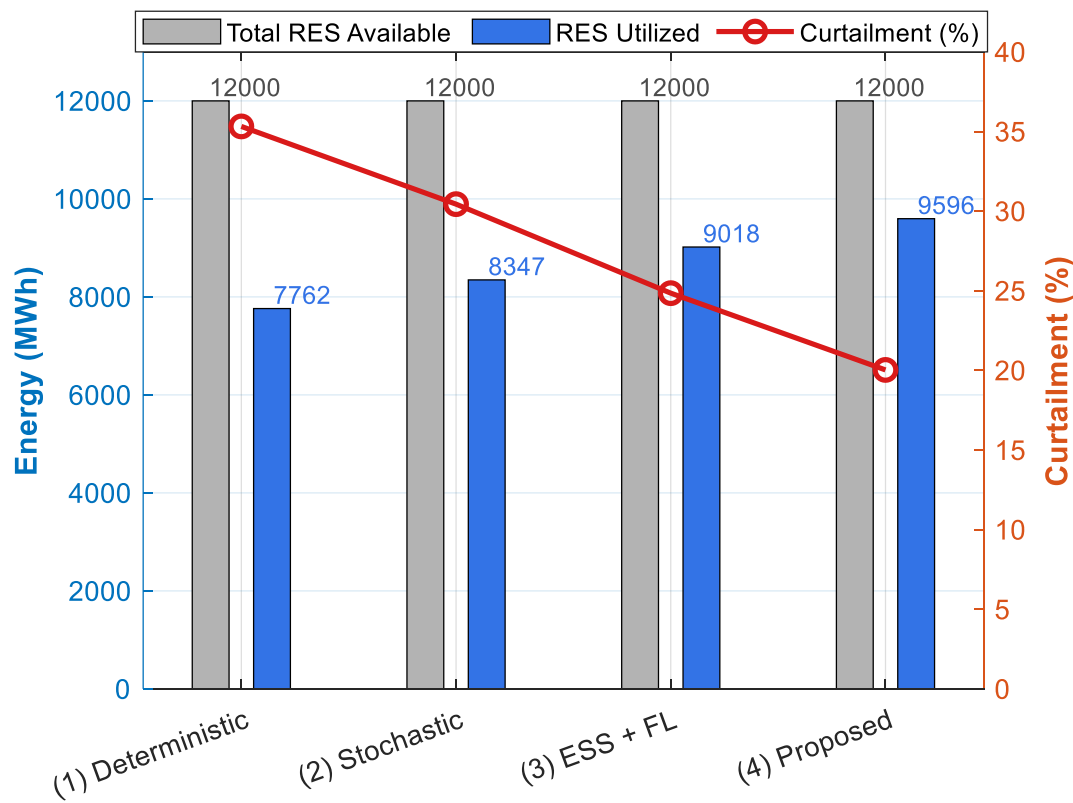


Fig. 5. Renewable Energy Utilization and Curtailment Across Configurations

Fig. 5 compares renewable energy utilization and curtailment across different unit commitment configurations. The gray bars represent the fixed total renewable energy availability (12,000 MWh) in each case, while the blue bars show the portion actually utilized. The red line depicts the percentage of curtailment. As flexibility resources are incrementally integrated—from none in Case 1 to the full configuration in Case 4—the utilized renewable energy increases and curtailment drops significantly, showcasing the effectiveness of combining stochastic modeling with energy storage, flexible loads, and EV-V2G technologies.

Table IV compares the reserve procurement costs for spinning and non-spinning reserves across four unit commitment configurations, each representing a different combination of uncertainty modeling and system flexibility. In Case 1, the deterministic model without any form of flexibility incurs the highest reserve procurement cost, totaling \$1,498,500. This includes \$1,197,300 for spinning reserves and \$301,200 for non-spinning reserves. The high spinning reserve cost results from the model's reliance on committed thermal generation to cover all possible uncertainties, as no flexible or responsive resources are available to assist in providing reserves.

Case 2 introduces a stochastic unit commitment approach that incorporates renewable generation uncertainty but still excludes flexible resources. In this scenario, spinning reserve costs are reduced to \$1,096,850 and non-spinning reserve costs to \$248,400, resulting in a total reserve cost of \$1,345,250. This marks a 10.25% reduction compared to the deterministic base case. The improvement reflects the stochastic model's ability to allocate reserves more precisely by considering forecast distributions rather than single-point predictions, thereby avoiding excessive and

costly spinning commitments. In Case 3, the stochastic model is further enhanced by integrating utility-scale storage systems and flexible loads. The total reserve cost drops further to \$1,241,550, comprising \$1,042,600 for spinning reserves and \$198,950 for non-spinning reserves. Storage systems contribute significantly to spinning reserve provision due to their fast response capabilities, while flexible loads support non-spinning reserve requirements by reducing or shifting demand when needed. This combination reduces dependence on thermal generators for reserve support and improves overall reserve efficiency.

Case 4 represents the proposed full model, which incorporates EV-V2G alongside ESS and flexible loads. This configuration results in the lowest total reserve procurement cost of \$1,147,400, with \$998,100 allocated to spinning reserves and \$149,300 to non-spinning reserves. Including EV fleets enhances the system's flexibility by adding distributed, controllable reserve capacity. During hours when EVs are connected, they can discharge power to support spinning reserve needs or adjust charging behavior to provide non-spinning reserve support. This distributed flexibility complements the centralized response from ESS and load control, enabling the system to meet reserve requirements more economically and with less reliance on traditional generation. Overall, Table IV confirms that each enhancement—uncertainty modeling, storage integration, demand flexibility, and EV-V2G—contributes to lower reserve procurement costs. The greatest cost savings are achieved when all flexibility resources are co-optimized in a stochastic framework. The results highlight the operational and economic value of coordinated flexibility in maintaining system reliability under high levels of renewable energy uncertainty.

TABLE IV
Reserve Procurement Costs Under Different Unit Commitment Configurations

Case	Spinning Reserve Cost (\$)	Non-Spinning Reserve Cost (\$)	Total Reserve Cost (\$)
(1) Deterministic w/o Flex	1,197,300	301,200	1,498,500
(2) Stochastic w/o Flex	1,096,850	248,400	1,345,250
(3) w/ ESS & FL (no EV)	1,042,600	198,950	1,241,550
(4) Proposed (EV + ESS + FL)	998,100	149,300	1,147,400

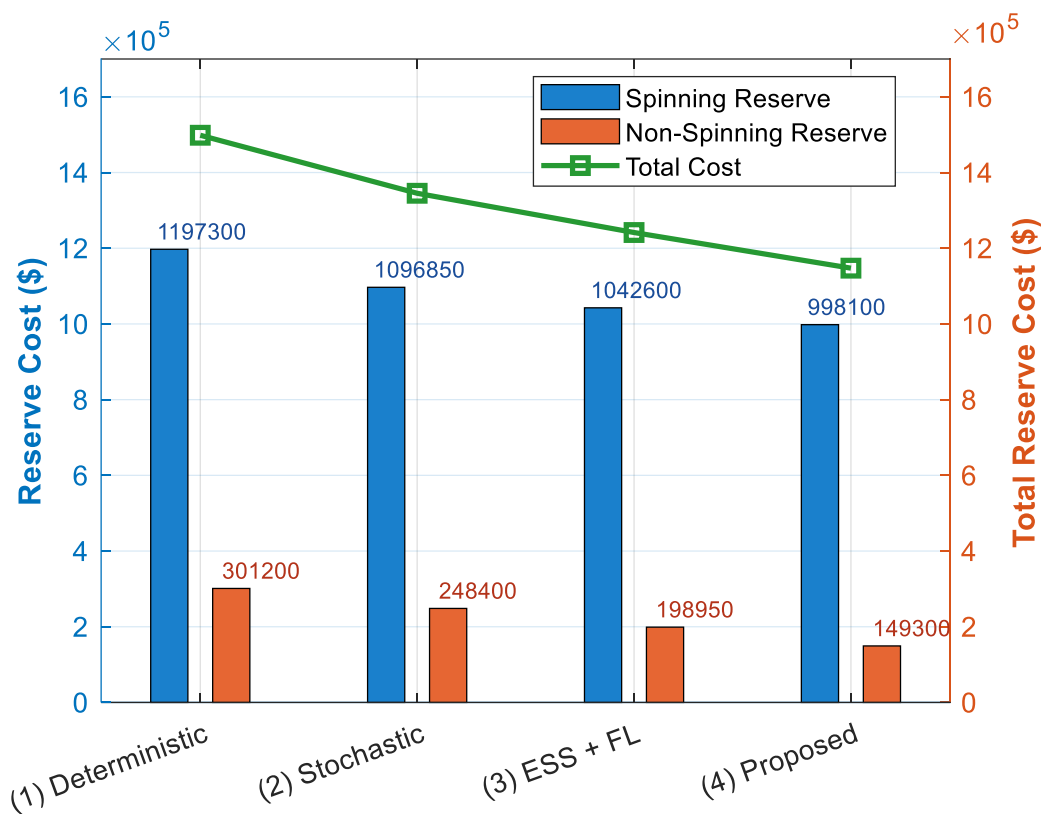


Fig. 6. Reserve Procurement Cost Across Commitment Configurations

Fig. 6 presents the spinning and non-spinning reserve procurement costs as grouped bars for each case, with the total reserve cost superimposed as a green line. The chart demonstrates how incorporating stochastic modeling and system flexibility — especially EVs — progressively reduces overall reserve costs. The most significant savings occur in the proposed model, confirming the value of co-optimized flexibility.

The simulation results clearly demonstrate that integrating flexibility resources and incorporating renewable uncertainty significantly enhances the power system's economic efficiency and operational reliability. The proposed two-stage stochastic SCUC model, which includes full coordination of electric vehicles with vehicle-to-grid capability (EV-V2G), utility-scale energy storage systems (ESS), and flexible loads (FL), consistently outperforms all other configurations examined. First, the total system cost is reduced from \$9,746,250 in the deterministic baseline to \$9,098,100 in the proposed model, reflecting a 6.66% cost reduction. This

improvement is mainly driven by a substantial decrease in expected second-stage operating costs, made possible through proactive scheduling and dynamic use of flexible resources in response to renewable variability. Second, renewable energy utilization increases from 7,762 MWh to 9,596 MWh, resulting in a curtailment reduction from 35.32% to 20.03%. This improvement is achieved through the synergistic operation of ESS and EV fleets, which absorb surplus renewable output during low-demand periods and provide dispatchable power during shortages.

Flexible loads further support this process by shifting consumption to periods of high renewable availability. Third, reserve procurement becomes significantly more cost-effective. The total reserve cost drops from \$1,498,500 in the base case to \$1,147,400 under the proposed model—a 23.4% reduction. This efficiency gain is enabled by the fast-response characteristics of ESS and EVs, as well as the controllability of industrial flexible loads, which together reduce dependence on costly thermal reserves.

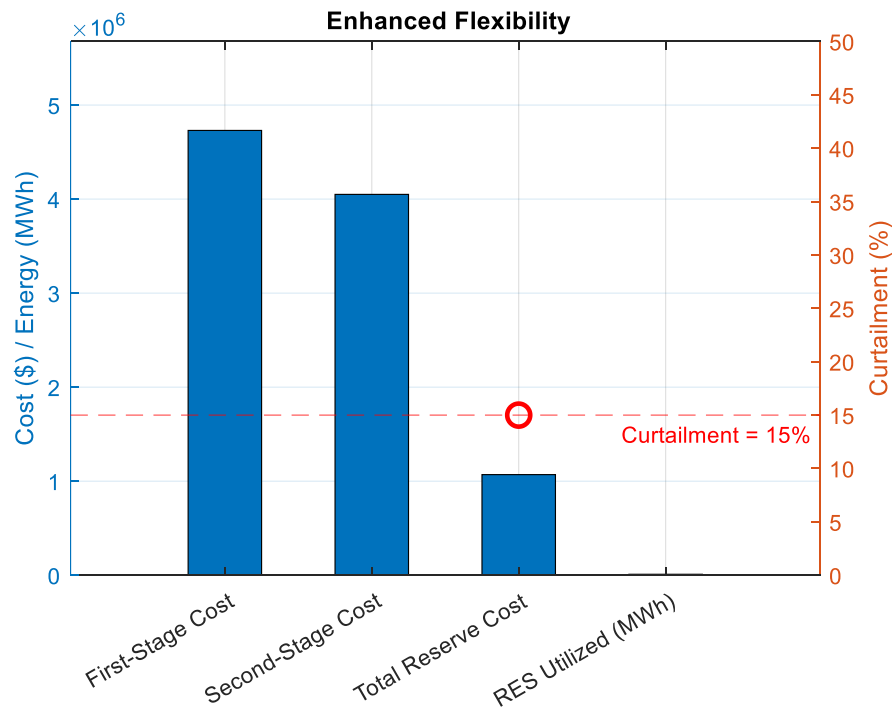


Fig. 7. Key performance indicators of the proposed SCUC model with co-optimized EV, ESS, and FL flexibility

In Fig. 7, the key performance indicators of the proposed configuration—featuring coordinated EV-V2G, utility-scale storage, and flexible loads—are visually summarized in Fig. 5. The left axis illustrates the economic and energy performance metrics, where the first-stage cost, second-stage cost, total reserve procurement cost, and renewable energy utilized are represented as individual bars. Notably, the second-stage cost shows a substantial decline relative to the first-stage cost, reflecting the model's improved operational efficiency due to enhanced flexibility. The right axis displays the curtailment percentage, which is plotted as a red marker

with a dashed reference line. A curtailment rate of 20.03% is achieved, indicating a significant reduction in wasted renewable energy compared to previous configurations. This improvement highlights the ability of the proposed model to absorb and utilize more variable generation, thanks to the coordinated operation of all three flexibility resources. Collectively, the figure emphasizes how the proposed model leads to lower operational costs, better renewable integration, and reduced reserve burdens, validating the benefits of multi-dimensional flexibility in high-renewable scenarios.

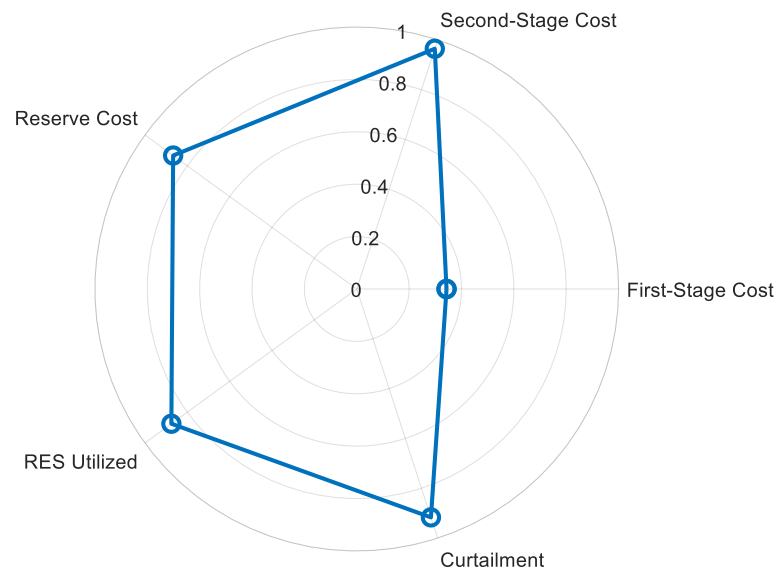


Fig. 8. Radar chart of normalized performance metrics for the proposed SCUC model

Fig. 8 uses a radar chart to visualize the relative performance of the proposed model across five normalized metrics: first-stage cost, second-stage cost, reserve cost, renewable energy utilization, and curtailment. Each axis represents one metric, scaled from 0 to 1, where a larger area signifies better performance. The results show a strong, balanced profile with low operating and reserve costs, high renewable utilization, and low curtailment. This confirms that the comprehensive integration of EV, ESS, and FL provides a well-rounded enhancement to system performance across all key dimensions.

C. Scalability and Computational Tractability

The proposed two-stage stochastic SCUC model with nested Benders decomposition is inherently computationally complex. The problem size scales with the number of scenarios ($|\Omega|$), contingencies ($|\mathcal{C}|$), time periods ($|T|$), network buses ($|B|$), and flexibility resources. Solving such a problem monolithically for a large-scale real-world system (e.g., a thousands-of-buses network) is computationally intractable. The primary value of the nested decomposition approach is to break this intractable monolithic problem into a sequence of smaller, more manageable subproblems.

Three key strategies underpin the scalability of our framework:

1. **Decomposition:** The master problem size depends only on the first-stage decision variables and the number of Benders cuts, not on $|\Omega|$ or $|\mathcal{C}|$. Each scenario subproblem is independent and can be solved in parallel, decoupling the complexity of renewable uncertainty. Similarly, within each scenario, contingency subproblems are also independent and parallelizable, isolating the burden of N-1 security analysis.

2. **Reduction:** Using scenario reduction (e.g., k-means clustering) to limit $|\Omega|$ to a tractable number of representative scenarios is critical. Similarly, contingency screening—whereby contingencies with a negligible likelihood or impact are filtered out—can significantly reduce $|\mathcal{C}|$ without materially compromising security.

3. **Parallelization:** As implemented, the algorithm exploits two levels of parallelism: across scenarios and across contingencies within a scenario. This can lead to a near-linear speedup in wall-clock time when deployed on high-performance computing (HPC) clusters with many cores.

For larger systems (e.g., a 2000-bus model), the main computational challenges would be:

- **Master Problem:** While independent of scenarios, its size grows with the number of generating units and resources. However, modern MILP solvers like Gurobi are highly efficient for large-scale UC problems.

- **Contingency Subproblems:** The number of transmission line contingencies ($|\mathcal{C}|$) scales linearly with the number of lines, which can be very high in large systems. This is the most significant scalability bottleneck. Aggressive contingency screening based on quick approximate analyses (e.g., using linearized sensitivity factors) is an essential practical step for industry-scale adoption.

- **Memory Overhead:** Managing the communication and storage of Benders cuts for thousands of scenarios and contingencies requires efficient data handling.

In conclusion, while challenging, applying this framework to real-world systems is feasible. The computational burden is not eliminated but is shifted to a parallel computing environment. For a large-scale ISO, leveraging a large HPC cluster would allow the solution times demonstrated here (e.g., 2 hours for a 118-bus system) to be maintained for significantly larger networks by solving thousands of scenarios and contingency subproblems simultaneously. Future work will focus on implementing more advanced contingency screening and investigating distributed computing frameworks to further enhance scalability.

D. Computational Performance

The computational performance of the nested Benders algorithm for each case study is summarized in Table V. As expected, the computational effort increases with the model's complexity. The deterministic base case (Case 1) converges in the fewest iterations (8) and has the lowest CPU time and memory footprint. Introducing stochasticity (Case 2) increases the number of iterations by 50% and the total CPU time by a factor of ~ 4 , as the algorithm must learn the recourse cost for multiple scenarios. Adding stationary flexibility resources (Case 3) further increases the problem size, leading to more iterations and longer solve times, particularly in the master problem, which now includes decisions for ESS and FL.

The proposed full model (Case 4) requires the most iterations (18) and the highest computational resources, as the master problem must co-optimize the commitment and reserve schedules for all resources, including the complex EV constraints, and the subproblems must simulate their operation across all scenarios and contingencies. Notably, the algorithm consistently converged to a tight optimality

gap ($<0.1\%$) for all cases, demonstrating its robustness. The subproblem solve time dominates the total CPU time, highlighting the immense benefit of the parallel

implementation described in Section 3. Memory usage is manageable for a system of this size, but it is a key consideration when scaling to larger networks.

TABLE V
Computational Performance of the Nested Benders Decomposition Algorithm

Case	Description	Benders Iterations	Total CPU Time (s)	Master Time (s)	Subproblems Time (s)	Final Gap (%)	Peak Memory (GB)
1	Deterministic w/o Flex	8	1,245	312	933	0.08	4.2
2	Stochastic w/o Flex	12	4,887	598	4,289	0.09	6.8
3	Stoch. + ESS & FL	15	5,912	845	5,067	0.07	8.5
4	Proposed (Full)	18	7,158	1,121	6,037	0.10	11.3

V. CONCLUSIONS AND FUTURE WORK

This paper presented a two-stage stochastic security-constrained unit commitment (SCUC) framework that integrates electric vehicles with vehicle-to-grid (EV-V2G) capabilities, utility-scale energy storage systems (ESS), and flexible loads. The model is designed to handle renewable generation uncertainty and enforce N-1 security across generator and transmission contingencies. A nested Benders decomposition algorithm was developed to solve the resulting large-scale stochastic mixed-integer problem efficiently.

The proposed model enables joint optimization of thermal generation schedules, reserve allocation, and the operation of distributed and centralized flexibility resources. Simulation results on a modified IEEE-118 bus system demonstrated that the proposed approach significantly improves system performance across multiple metrics. Compared to a deterministic SCUC baseline, the model achieved a 6.67% reduction in total system cost, a 15% absolute reduction in renewable curtailment, and a 23% reduction in reserve procurement costs. EV fleets contributed meaningful spinning and non-spinning reserve capacity, particularly during peak demand periods, while ESS and flexible loads enhanced the system's ability to absorb renewable fluctuations and respond to contingencies. The findings confirm that integrating emerging flexibility resources into a stochastic and security-aware unit commitment framework enables more reliable, cost-effective, and renewable-friendly power system operation.

A. Limitations and Future Research Directions

While this study presents a comprehensive framework for SCUC with diverse flexibility resources, it is subject to certain limitations that also represent avenues for future research. First, the model relies on a DC power flow

approximation for network constraints. While this is a standard practice in large-scale UC problems due to its computational linearity, it neglects reactive power, voltage limits, and transmission losses. Future work could integrate a linearized AC power flow model or a corrective AC security-constrained optimal power flow (SCOPF) in the second stage to enhance operational accuracy, albeit at a significant computational cost. Second, the modeling of EV user behavior is simplified. We assume perfect compliance and a known, deterministic availability schedule for aggregated fleets.

In reality, user behavior is stochastic—connection/disconnection times and energy requirements are variable and uncertain. A valuable extension would be to model these parameters stochastically within a multi-stage framework or incorporate data-driven user behavior models to improve scheduling robustness. Third, the market and policy context is abstracted. The model assumes a centralized cost-minimization paradigm. Integrating this framework into a decentralized market environment, where flexibility resources are owned by profit-seeking entities participating in day-ahead and real-time markets, would be a critical step toward real-world implementation. This involves formulating equilibrium or bi-level models to capture strategic bidding behavior. Other promising extensions include: incorporating distribution network constraints to model the impact of widespread EV charging on local grids; evaluating the long-term degradation costs of batteries in ESS and EVs more precisely; and expanding the security analysis to include voltage stability and small-signal stability constraints alongside the current N-1 reliability focus. Addressing these limitations would further bridge the gap between the proposed academic framework and its practical application by system operators.

Nomenclature

Sets and Indices

$\mathcal{T} = \{1, \dots, T\}$: Set of hourly time periods (e.g., $T = 24$ for one day).

$\mathcal{I} = \{1, \dots, I\}$: Set of thermal generating units.

$\mathcal{S} = \{1, \dots, S\}$: Set of RES-uncertainty scenarios.

$\Omega = \{1, \dots, \Omega\}$: Set of contingencies (N-1 events, e.g., single-unit or single-line outage).

$\mathcal{B} = \{1, \dots, B\}$: Set of buses.

$\mathcal{L} = \{1, \dots, L\}$: Set of transmission lines.

$\mathcal{E} = \{1, \dots, E\}$: Set of EV fleets aggregated at charging stations.

$\mathcal{K} = \{1, \dots, K\}$: Set of ESS units.

$\mathcal{F} = \{1, \dots, F\}$: Set of flexible load blocks (aggregated flexible loads).

$\mathcal{R} = \{1, \dots, R\}$: Set of RES units (wind farms and photovoltaic plants).

$t \in \mathcal{T}$: Time index (hour).

$i \in \mathcal{I}$: Thermal generator index.

$s \in \mathcal{S}$: Scenario index.

$\omega \in \Omega$: Contingency index.

$b \in \mathcal{B}$: Bus index.

$\ell \in \mathcal{L}$: Transmission line index.

$e \in \mathcal{E}$: EV fleet index.

$k \in \mathcal{K}$: ESS index.

$f \in \mathcal{F}$: Flexible load index.

$r \in \mathcal{R}$: RES index.

Parameters

C_i^{SU} : Start-up cost of unit i [\$/start].

C_i^{SD} : Shut-down cost of unit i [\$/shut].

$C_i^{\text{G}}(P) = a_i + b_i P + c_i P^2$: Quadratic generation cost function [\$/h] for unit i .

$P_i^{\text{min}}, P_i^{\text{max}}$: Minimum and maximum output (MW) of unit i .

RU_i, RD_i : Ramp-up and ramp-down limits (MW/h) of unit i .

UT_i, DT_i : Minimum up-time and down-time (hours) of unit i .

$\$R_{\{i\}^{\text{sp, max}}}$: Maximum spinning reserve capacity (MW) that unit i can provide when online.

B_ℓ : Susceptance (p.u.) of line ℓ .

f_ℓ^{max} : Thermal flow limit (MW) of line ℓ .

$\$PTDF_{\{b\ell\}}$: Power transfer distribution factor for bus b to line ℓ .

$\tilde{W}_{r,t}^s$: Available RES output (wind or solar) for unit r , period t , in scenario s (MW).

π_s : Probability of scenario s , $\sum_{s \in \mathcal{S}} \pi_s = 1$.

$D_{b,t}$: Inflexible (firm) demand at bus b in period t (MW).

$P_{f,t}^{\text{max}}$: Maximum (upper) allowable flexible load consumption (MW) for load f in hour t .

$P_{f,t}^{\text{min}}$: Minimum consumption level (MW) for load f in hour t .

$\bar{E}_f$: Total energy requirement over horizon for load f (MWh).

$\alpha_{f,t}$: Disutility or penalty cost coefficient [\$/MWh] if load f consumes less than baseline in hour t .

E_k^{max} : Energy capacity (MWh) of ESS k .

$P_k^{\text{ch, max}}, P_k^{\text{dis, max}}$: Maximum charging and discharging power (MW) for ESS k .

$\eta_k^{\text{ch}}, \eta_k^{\text{dis}}$: Charging and discharging efficiency of ESS k .

$E_{k,0}$: Initial state of charge (MWh) of ESS k .

E_k^{min} : Minimum allowable SOC (MWh) for ESS k .

E_k^{max} : Maximum allowable SOC (MWh) for ESS k .

N_e : Number of EVs in fleet e .

E_e^{cap} : Battery capacity (MWh) per EV in fleet e .

$P_e^{\text{ch, max}}, P_e^{\text{dis, max}}$: Aggregate maximum charging/discharging power (MW) of fleet e .

$\eta_e^{\text{ch}}, \eta_e^{\text{dis}}$: Aggregate charging/discharging efficiency of EV fleet e .

$\text{SOC}_{e,0}$: Initial state of charge (MWh) of each EV in fleet e at $t = 0$.

$\text{SOC}_e^{\text{min}}, \text{SOC}_e^{\text{max}}$: Minimum/maximum allowable SOC (MWh) per EV.

$L_{e,t}$: Number (fraction) of EVs of fleet e connected (available) at time t . (*This can be based on a (charging) availability profile.*)

$\hat{E}_e^{\text{req}}$: Required SOC (MWh) by departure time for EV fleet e (ensures mobility).

$T_e^{\text{arr}}, T_e^{\text{dep}}$: Arrival and departure times for EV fleet e .

$C_e^{\text{EV, ch}}, C_e^{\text{EV, dis}}$: Charging/discharging cost or compensation of EV fleet e [\$/MWh].

$C_k^{\text{ESS, ch}}, C_k^{\text{ESS, dis}}$: Charging/discharging cost of ESS k [\$/MWh].

C_f^{FL} : Disutility cost coefficient [\$/MWh] for flexible load f .

$C_{b,t}^{LS}$: Load shedding penalty at bus b , time t [\$/MWh].

Variables

$u_{i,t} \in \{0,1\}$: 1 if unit i is ON in period t , 0 otherwise.

$v_{i,t} \in \{0,1\}$: 1 if unit i starts up at period t , 0 otherwise.

$w_{i,t} \in \{0,1\}$: 1 if unit i shuts down at period t , 0 otherwise.

$z_{k,t}^{ch} \in \{0,1\}$: 1 if ESS k is charging in period t , 0 otherwise.

$z_{k,t}^{dis} \in \{0,1\}$: 1 if ESS k is discharging in period t , 0 otherwise.

$z_{e,t}^{ch} \in \{0,1\}$: 1 if EV fleet e is charging in period t , 0 otherwise.

$z_{e,t}^{dis} \in \{0,1\}$: 1 if EV fleet e is discharging in period t , 0 otherwise.

$p_{i,t}^s \geq 0$: Real power output (MW) of unit i at time t under scenario s .

$r_{i,t}^{s,sp} \geq 0$: Spinning reserve provided by unit i at time t , scenario s .

$r_{b,t}^{s,ns} \geq 0$: Non-spinning reserve at bus b at t , scenario s .

$p_{k,t}^{s,ch}, p_{k,t}^{s,dis} \geq 0$: Charging/discharging power (MW) of ESS k at t , scenario s .

$e_{k,t}^s$: State of charge (SoC) (MWh) of ESS k at t , scenario s .

$p_{e,t}^{s,ch}, p_{e,t}^{s,dis} \geq 0$: Charging/discharging power (MW) of EV fleet e , at t , scenario s .

$SOC_{e,t}^s$: Average SoC per EV in fleet e (MWh) at t , scenario s .

$d_{f,t}^s \geq 0$: Actual consumption (MW) served by flexible load f at t , scenario s .

$\Delta D_{b,t}^{s,shed} \geq 0$: Load shedding at bus b , time t , scenario s .

$\theta_{b,t}^{s,\omega}$: Voltage angle at bus b at time t , scenario s , under contingency ω .

$f_{\ell,t}^{s,\omega}$: Real power flow on line ℓ at t , scenario s , under contingency ω .

REFERENCES

- [1] Q. Niu, L. Tang, L. Yu, H. Wang, and Z. Yang, "Unit commitment considering electric vehicles and renewable energy integration—A CMAES approach," *Sustainability*, vol. 16, no. 3, p. 1019, 2024. doi: 10.3390/su16031019.
- [2] G. Xu, Z. Lin, Q. Wu, W. K. V. Chan, and X. P. Zhang, "Deep reinforcement learning based model-free optimization for unit commitment against wind power uncertainty," *Int. J. Electr. Power Energy Syst.*, vol. 155, p. 109526, 2024. doi: 10.1016/j.ijepes.2023.109526.
- [3] T. Jain and K. Verma, "Reliability based computational model for stochastic unit commitment of a bulk power system integrated with volatile wind power," *Reliab. Eng. Syst. Saf.*, vol. 244, p. 109949, 2024. doi: 10.1016/j.res.2024.109949.
- [4] S. Qi et al., "Two-Stage Robust Optimization Considering the Uncertainty of Sources and Loads in Virtual Power Plants," in *IEEE Access*, vol. 13, pp. 132155-132169, 2025. doi: 10.1109/ACCESS.2025.3591459.
- [5] X. Liu et al., "Chance-constrained scheduling considering frequency support from electric vehicles under multiple uncertainties," *IET Renew. Power Gener.*, 2024. doi: 10.1049/rpg2.13171.
- [6] Z. Chang et al., "Consideration of Wind-Solar Uncertainty and V2G Mode of Electric Vehicles in Bi-Level Optimization Scheduling of Microgrids," *Energies*, vol. 18, no. 4, p. 823, 2025. doi: 10.3390/en18040823.
- [7] S. Zhan, Y. Zhou, D. Feng, H. Wang, and C. Fang, "Energy optimization dispatch based on two-stage and multi-objective comparative analysis for PV and BESS integrated fast charging stations with V2G," *IET Renew. Power Gener.*, vol. 18, no. S1, pp. 4426-4438, 2024. doi: 10.1049/rpg2.12913.
- [8] Y. Luo et al., "Optimal operation of microgrid with consideration of V2G's uncertainty," *IET Gener., Transm. Distrib.*, 2024. doi: 10.1049/gtd2.13033.
- [9] X. Xu, Z. Qiu, T. Zhang, and H. Gao, "Distributed Source-Load-Storage Cooperative Low-carbon Scheduling Strategy Considering Vehicle-to-grid Aggregators," *J. Mod. Power Syst. Clean Energy*, vol. 12, no. 2, pp. 440-453, 2024. doi: 10.35833/MPCE.2023.000000.
- [10] Z. Guo, J. Bai, W. Wei, S. Mei, and W. Hu, "Partially Affine Policy for Multistage Robust Unit Commitment with Fast-Ramping Units," in *CSEE Journal of Power and Energy Systems*, vol. 11, no. 1, pp. 477-480, January 2025. doi: 10.17775/CSEEJPES.2023.08300.
- [11] H. Karimianfard, M. R. Salehizadeh, and P. Siano, "Economic Profit Enhancement of a Demand Response Aggregator Through Investment of Large-scale Energy Storage Systems," in *CSEE Journal of Power and Energy Systems*, vol. 8, no. 5, pp. 1468-1476, September 2022. doi: 10.17775/CSEEJPES.2021.02650.
- [12] X. Yu, D. Pan, and Y. Zhou, "A stochastic vehicle schedule model for demand response and grid flexibility in a renewable-building-e-transportation microgrid," *Sol. Energy*, p. 112388, 2024. doi: 10.1016/j.solener.2024.112388.
- [13] J. Kazempour et al., "Electrolysis as a flexibility resource on energy islands: The case of the North Sea," *Energy Policy*, vol. 185, p. 113921, 2024. doi: 10.1016/j.enpol.2024.113921.
- [14] Hossein Karimianfard, A robust optimization framework for smart home energy management: Integrating photovoltaic storage, electric vehicle charging, and demand response, *Journal of Energy Storage*, Volume 110, 115259, 2025. doi.org/10.1016/j.est.2024.115259.
- [15] H. Liu, F. Liang, T. Hu, J. Hong, and H. Ma, "Multi-scale Fusion Model Based on Gated Recurrent Unit for Enhancing Prediction Accuracy of State-of-charge in Battery Energy Storage Systems," *J. Mod. Power Syst. Clean Energy*, vol. 12, no. 2, pp. 405-414, 2024. doi: 10.35833/MPCE.2023.000000.
- [16] H. Haghighat, H. Karimianfard, and B. Zeng, "Integrating Energy Management of Autonomous Smart Grids in Electricity Market Operation," in *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4044-4055, Sept. 2020. doi: 10.1109/TSG.2020.2992570.
- [17] J. Bian et al., "Optimal Bidding Strategy for PV and BESSs in Joint Energy and Frequency Regulation Markets Considering Carbon Reduction Benefits," *J. Mod. Power Syst. Clean Energy*, vol. 12, no. 2, pp. 427-439, 2024. doi: 10.35833/MPCE.2023.000000.

- [18] Y. Zhuang, Z. Li, Q. Tan, Y. Li, and M. Wan, "Multi-time-scale Resource Allocation Based on Long-term Contracts and Real-time Rental Business Models for Shared Energy Storage Systems," *J. Mod. Power Syst. Clean Energy*, vol. 12, no. 2, pp. 454–465, 2024. doi: 10.35833/MPCE.2023.000000.
- [19] M. Agüero *et al.*, "Virtual Transmission Solution Based on Battery Energy Storage Systems to Boost Transmission Capacity," *J. Mod. Power Syst. Clean Energy*, vol. 12, no. 2, pp. 466–474, 2024. doi: 10.35833/MPCE.2023.000000.
- [20] Y. Zhou, J. Liu, and J. L. M. Hensen, "Advances in emerging digital technologies for energy efficiency and energy integration in smart cities," *Energy Build.*, p. 114289, 2024. doi: 10.1016/j.enbuild.2024.114289.
- [21] H. Karimianfard, H. Haghighat, and B. Zeng, "Co-Optimization of Battery Storage Investment and Grid Expansion in Integrated Energy Systems," in *IEEE Systems Journal*, vol. 16, no. 4, pp. 5928–5938, Dec. 2022, doi: 10.1109/JSYST.2021.3130057.