Computers & Industrial Engineering 218 (2026) 112112

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Scalable multi-level optimization for sequentially cleared energy markets with a case study on gas and carbon aware unit commitment

Yuxin Xia [a,*], Iacopo Savelli [b], Thomas Morstyn [a]

[a] Department of Engineering Science, University of Oxford, Oxford, OX1 3PJ, UK
[b] GREEN center, Bocconi University, Milano, 20136, Italy

ARTICLE INFO

Keywords:
Multi-level problems
Multi-scale markets
Sequential decision-making
Benders subproblem separability
Unit commitment

ABSTRACT

Complex decision-making problems in industrial and infrastructure systems often involve hierarchical and sequential interactions across multiple layers, but existing optimization approaches face challenges in coordinating decision-making across these levels in a computationally efficient manner. To this end, this paper examines Mixed-Integer Multi-Level problems with Sequential Followers (MIMLSF), a specialized optimization model aimed at enhancing upper-level decision-making by incorporating anticipated outcomes from lower-level sequential market-clearing processes. We introduce a novel approach that combines lexicographic optimization with a weighted-sum method to asymptotically approximate the MIMLSF as a single-level problem, capable of managing multi-level problems exceeding three levels. To enhance computational efficiency and scalability, we propose a dedicated Benders decomposition method with multi-level subproblem separability. To demonstrate the practical application of our MIMLSF solution technique, we tackle a unit commitment (UC) problem within an integrated electricity, gas, and carbon market clearing framework in the Northeastern United States, enabling the incorporation of anticipated costs and revenues from gas and carbon markets into UC decisions. This ensures that only profitable gas-fired power plants (GFPPs) are committed, allowing system operators to make informed decisions that prevent GFPP economic losses and reduce total operational costs under stressed electricity and gas systems. The case study not only demonstrates the applicability of the MIMLSF model but also highlights the computational benefits of the dedicated Benders decomposition technique, achieving average reductions of 32.23% in computing time and 94.23% in optimality gaps compared to state-of-the-art methods.

1. Introduction

Sequential decision-making is fundamental to many real-world systems and plays a central role in the operation and clearing of multi-scale energy markets spanning temporal, spatial and sectoral dimensions. Moreover, critical operations such as unit commitment (UC) are executed ahead of time, often without full knowledge of subsequent sequential system and market outcomes. This lack of awareness and coordination can result in suboptimal or economically inefficient decisions, as the initial decisions may fail to incorporate the market conditions that subsequently emerge. Approaches based on multi-level or hierarchical structures can help decision-makers such as system operators, regulators, and strategic agents anticipate and adapt to expected market conditions, thereby enhancing risk management and overall decision performance.

1.1. Bilevel and multi-level optimization

To mathematically formulate this hierarchical decision-making framework, hierarchical optimization problems have been extensively studied. In classic bi-level programming, the structure comprises an upper-level leader and a lower-level follower, with the follower's actions acting as constraints on the leader's decisions, as represented in Fig. 1(a). This bi-level optimization programming is extensively reviewed in the literature, see Beck et al. (2023) and Kleinert et al. (2021). Multi-level programming, with more than two levels, has been gaining increasing attention due to its applicability to real-world problems (Lu et al., 2016). Multi-level programming can address nested hierarchies where multiple agents make decisions in sequence driven by individual objectives. As an example, the structure of a classic tri-level problem with a nested decision-making hierarchy is depicted in Fig. 1(b) (Avraamidou & Pistikopoulos, 2019). In this

* Corresponding author.
E-mail addresses: yuxin.xia@eng.ox.ac.uk (Y. Xia), iacopo.savelli@unibocconi.it (I. Savelli), thomas.morstyn@eng.ox.ac.uk (T. Morstyn).

https://doi.org/10.1016/j.cie.2026.112112
Received 17 November 2025; Received in revised form 9 March 2026; Accepted 9 May 2026
Available online 16 May 2026
0360-8352/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

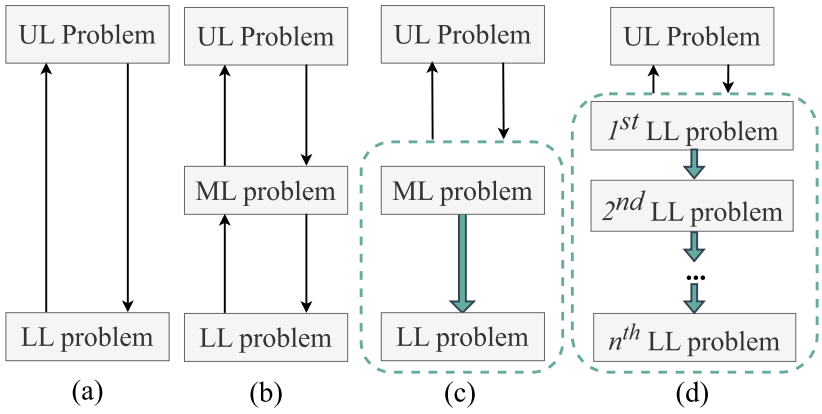

Fig. 1. Different models of hierarchical decision making process. (a) classic bi-level problem; (b) classic tri-level problem; (c) tri-level problem with sequential ML and LL problems; (d) multi-level problem with n sequential LL problems. UL: Upper-Level, ML: Middle-Level, LL: Lower-Level.

classic tri-level optimization problem, the first decision-maker at the upper-level formulates an optimization problem whose constraints include the optimization problem of the second decision-maker. This second decision-maker's problem further incorporates the optimization problem of a third decision-maker in the constraints, making for a nested decision-making hierarchy (Avraamidou & Pistikopoulos, 2019). Recently, a special tri-level structure involving two sequential followers has been studied in Byeon and Van Hentenryck (2020), as shown in Fig. 1(c). In this special framework, unlike the classic tri-level problem, the middle-level problem does not anticipate the solutions of the lower-level problem. This special tri-level optimization framework has been applied to diverse power system problems, such as heat unit commitment in sequential heat–electricity markets (Mitridati et al., 2019) and sequential national–local market operations for flexible service bidding (Paredes et al., 2023) and merchant transmission investment (Xia et al., 2025).

Multi-level problems, even for the case of bi-levels, are recognized as strongly NP-hard, presenting significant computational challenges (Fakhry et al., 2022). There are two main solution techniques for bilevel problems with convex lower-level problems: the Karush–Kuhn–Tucker (KKT)-based formulation and the strong duality-based formulation (Zare et al., 2019). The KKT-based approach incorporates the complementary slackness conditions, which are typically linearized using big-M constraints and require additional auxiliary binary variables (Nasrolahpour et al., 2018). Although effective for small to medium-sized instances, this method often encounters computational difficulties in large-scale problems due to the rapid increase in the number of auxiliary variables (Byeon, 2020). In contrast, the strong duality approach enforces the optimality of the lower-level problem through a single strong duality constraint, thereby eliminating the need for auxiliary variables associated with the linearization of complementary slackness conditions (Savelli et al., 2020; Xia et al., 2024). However, this formulation may lead to bilinear terms in the strong duality condition. These bilinear terms can be linearized using exact McCormick relaxations for some specific problems (McCormick, 1976), or alternatively reformulated through problem-specific techniques when applicable (Savelli et al., 2020; Xia et al., 2024). It has been shown that the strong duality approach can significantly reduce computational time, in some cases by orders of magnitude, particularly when the lower-level problem is more complex than the upper-level problem (Zare et al., 2019).

Unlike bi-level problems, tri-level and more complex multi-level problems often cannot be efficiently solved using traditional methods (Avraamidou & Pistikopoulos, 2019). To address these problems, researchers have developed both decomposition methods and column-and-constraint generation approaches (Dvorkin et al., 2018; Florensa et al., 2017; Gan et al., 2022; Ghorbani-Renani et al., 2021; Muñoz-Delgado et al., 2019; Sadeghi et al., 2017; Wu & Conejo, 2017). Additionally, multi-parametric and heuristic approaches offer alternative solutions (Avraamidou & Pistikopoulos, 2019; Chouhan et al., 2022; El-Meligy et al., 2021; Fakhry et al., 2022). In general, solution methods with theoretical guarantees for multi-level hierarchical optimization remain limited. Existing approaches often rely on metaheuristic or perturbation-based schemes that neither fully exploit hierarchical dependencies nor provide rigorous convergence guarantees (Tilahun et al., 2012). Recent work has proposed proximal-gradient methods for particular convex tri-level settings (Shafiei et al., 2024); however, the associated convergence results typically require strong convexity, and their practical computational performance remains largely unexplored (Shafiei et al., 2024). More recently, Sato et al. (2021) introduced a gradient method with theoretical guarantees. Nevertheless, these methods do not readily extend to general multi-level hierarchical problems (Sato et al., 2021).

In multi-level problems with multiple lower levels, one strategy is to treat upper-level decisions and preceding lower-level outputs as fixed parameters. While KKT- or strong-duality-based reformulations can in principle handle such structures, they result in highly complex single-level models. KKT-based formulations introduce numerous auxiliary variables, limiting scalability as the number of lower levels grows. Strong-duality approaches, in contrast, produce bilinear terms where dual variables are multiplied by primal variables from preceding stages, resulting in nonlinear constraints across levels. The single-level formulations resulting from both approaches are not only computationally challenging, but also significantly hinder the use of decomposition techniques for efficient solution.

1.2. Sequential multi-scale markets

In energy systems, multi-scale market operations involve interdependent decision-making across *temporal*, *spatial*, and *sectoral* layers. These interdependencies stem from physical and economic linkages introduced by specific technologies and market mechanisms. Technologies such as gas-fired power plants (GFPPs), combined heat and power (CHP) units, and heat pumps create cross-carrier physical couplings, while instruments such as carbon pricing introduce economic interactions across sectors. GFPPs connect the gas and electricity sectors through fuel-to-power conversion, while CHP units extend this coupling to the heat sector by simultaneously producing electricity and heat. As a result, decisions made in one market, such as electricity, can influence subsequent operations in other markets, such as natural gas and carbon, as well as across different temporal stages.

For example, electricity markets are organized hierarchically and cleared *sequentially*, each operating at different *temporal* resolutions ranging from months to minutes. This structure has been widely modeled in strategic bidding (Rintamäki et al., 2020; Wozabal & Rameseder,

2020) and energy management (Jiao et al., 2022; Putratama et al., 2023). Moreover, while markets for different energy sectors are *sectorally* independent, they are becoming increasingly interconnected. For example, although the electricity and natural gas markets typically operate independently, they are interconnected through GFPPs. Currently, these two markets are cleared sequentially in the United States (Byeon & Van Hentenryck, 2020), and there is significant research interest in how their interdependence affects problems including generation expansion planning (Bent et al., 2018), strategic bidding (Dimitriadis et al., 2023) and UC (Byeon & Van Hentenryck, 2020). Additionally, the sequence in which electricity and heat markets are cleared (Mitridati et al., 2020, 2019), as well as the design of carbon markets—which can be cleared either simultaneously or sequentially with the electricity market (Chen et al., 2021; Dimitriadis et al., 2023; Jiang et al., 2023)–will affect market outcomes. The *temporal and sectoral* dimensions are also crucial for current market models, particularly impacting the coordination of day-ahead and real-time markets for electricity and natural gas. A range of coordination approaches have been proposed, ranging from sequentially decoupled to fully uncoordinated and purely stochastic approaches (Chen et al., 2022; Ordoudis et al., 2019; Schwele et al., 2021). In addition, recent research has increasingly focused on the *spatial and temporal* dynamics of local electricity markets. Notably, Khan et al. (2022) have developed a bidding strategy for producers participating in sequential market clearing at wholesale and retail levels. Moreover, the study in Paredes et al. (2023) explores a sequential market clearing framework, which models the stacking of flexibility revenues of distributed energy resources in sequential national and local markets.

As previously discussed, GFPPs play a critical role in linking the electricity and gas markets. These plants are notable for providing environmental benefits, low capital costs, high efficiency, and operational flexibility (Wang et al., 2018). Nevertheless, GFPPs encounter significant challenges when these two markets operate independently (Byeon & Van Hentenryck, 2020). This is particularly problematic when GFPPs must bid in the electricity market without knowledge of actual gas prices (i.e., fuel costs), which can lead to substantial economic losses, as evidenced during the 2014 polar vortex (Byeon & Van Hentenryck, 2020; Gil et al., 2016; PJM, 2014) in the Northeastern United States. Proposals to enhance coordination between electricity and gas networks have been suggested (Chen et al., 2022; Ordoudis et al., 2019; Schwele et al., 2021), yet these markets typically clear independently, heightening financial risks during high stresses for GFPPs on both networks. In response, a new bid-validity constraint has been introduced in Byeon and Van Hentenryck (2020) to ensure that only profitable GFPPs can be committed, taking into account solely the gas costs.

The growing significance of climate change has prompted stakeholders and regulators to push organizations toward incorporating environmental externalities into their strategic and investment decisions through carbon taxes, permits, and market-based mechanisms (Dhavale & Sarkis, 2018). The carbon cap-and-trade system, which enables the trading of emission allowances within regulatory constraints (European Commission, 2024), is a cost-effective mechanism for reducing greenhouse gas emissions and promoting decarbonization across the energy sector. Recent studies have focused on integrating carbon markets with traditional energy sectors, including electricity (Cheng et al., 2023; Hua et al., 2024; Li et al., 2024), natural gas, and heat markets (Dimitriadis et al., 2023; Zhu et al., 2024), to optimize environmental and economic outcomes. The introduction of the carbon market has redefined the roles of GFPPs, which traditionally participate as sellers in the electricity market and buyers in the gas market. Now, these entities also engage as either sellers or buyers in the carbon market, depending on their emission levels and allowances. Notably, the study by Chen et al. (2021) provides an analysis of GFPPs' participation in the electricity, gas, and carbon markets, focusing on market equilibria and the potential for market power exploitation by strategic energy producers. However, their analysis does not extend to operational decision-making, such as day-ahead UC, and thus overlooks how participation in multiple interdependent markets affects short-term scheduling. This gap can result in suboptimal UC outcomes and increased financial risks for GFPPs, particularly under volatile gas and carbon prices. As carbon markets expand, their influence on operational costs and investment decisions will grow, making it essential to capture carbon-market feedback within UC models. Ignoring these interactions may reduce GFPP profitability and system flexibility and increase overall operational costs.

In theory, a fully integrated market framework that jointly optimizes across energy systems and timescales could achieve the highest level of cost efficiency (Ordoudis et al., 2019). In practice, however, such integration remains challenging due to technical limitations and regulatory fragmentation. Accordingly, there is a need for modeling approaches that enable coordinated decision-making across markets while preserving existing market arrangements, including the sequential nature of real-world market operations.

1.3. Formulations for multi-scale market coordination

The proposed multi-level formulation, illustrated in Fig. 1(d), provides a structural representation of hierarchical multi-scale market coordination. A comparison of the proposed method with different decision structures and solution paradigms is summarized in Table 1.

Unlike the classical bilevel model shown in Fig. 1(a), in which the leader anticipates the outcome of a single lower-level problem (possibly aggregating multiple parallel lower-level problems), the proposed model considers a sequence of interdependent follower stages. In practice, due to temporal differentiation in market clearing and institutional separation across markets, lower-level decisions are cleared sequentially rather than jointly optimized. Representing the stage-wise process as a single aggregated lower-level problem ignores the sequential nature of real market clearing. The leader therefore anticipates a centralized and jointly optimized response that does not reflect actual market operations. As a result, the implied upper-level strategy is derived from distorted market signals and may yield suboptimal decisions and economic inefficiencies in practice.

It also differs from classic tri-level models as shown in Fig. 1(b), where the middle and lower levels are embedded within a fully nested optimization hierarchy. In practical multi-scale market settings, clearing typically follows sequential procedures rather than iterative bidirectional information exchange across decision levels. Classical nested tri-level structures do not explicitly capture this non-iterative sequential clearing process, nor do they naturally represent temporally ordered market stages that unfold sequentially. Moreover, the proposed model extends the sequential tri-level structure illustrated in Fig. 1(c) by accommodating multiple sequentially linked market stages, thereby generalizing the classic three-level setting.

The proposed formulation further departs from fully centralized integrated models that jointly optimize all sectors and timescales within a single problem, thereby disregarding institutional separation, as well as from purely sequential but non-hierarchical market models in which stages are cleared consecutively without a cross-level coordination mechanism. It preserves the sequential structure of multi-scale markets and their institutional separation while enabling market coordination.

1.4. Research gaps and contributions

The increasing interdependence among energy markets has introduced new coordination challenges. In practice, operational and planning decisions are often made without fully accounting for the couplings that exist across multiple scales, leading to suboptimal decisions and reduced system-level efficiency. This hierarchical and sequential structure of decision-making can be naturally represented as a multi-level optimization problem, where each level models interdependent

Table 1
Comparison of decision structures and solution paradigms.

Model type	Solution approach	Scalability	Consistency w/ sequential market structure	Market coordination	≥ 3 hierarchical levels
Bilevel (convex LL)	KKT reformulation; strong duality-based reformulation; decomposition methods	Medium–High	No	Limited	No
Classic tri-level	Metaheuristic algorithms; gradient-based methods	Low	No	Yes	Yes (3 levels only)
Sequential	Stage-by-stage solving	High	Yes	No	N/A
Centralized	Direct centralized optimization; decomposition methods; heuristics	Low–Medium	No	Yes	N/A
Proposed MIMLSF	Asymptotic approximation; dedicated decomposition	High	Yes	Yes	Yes (n levels)

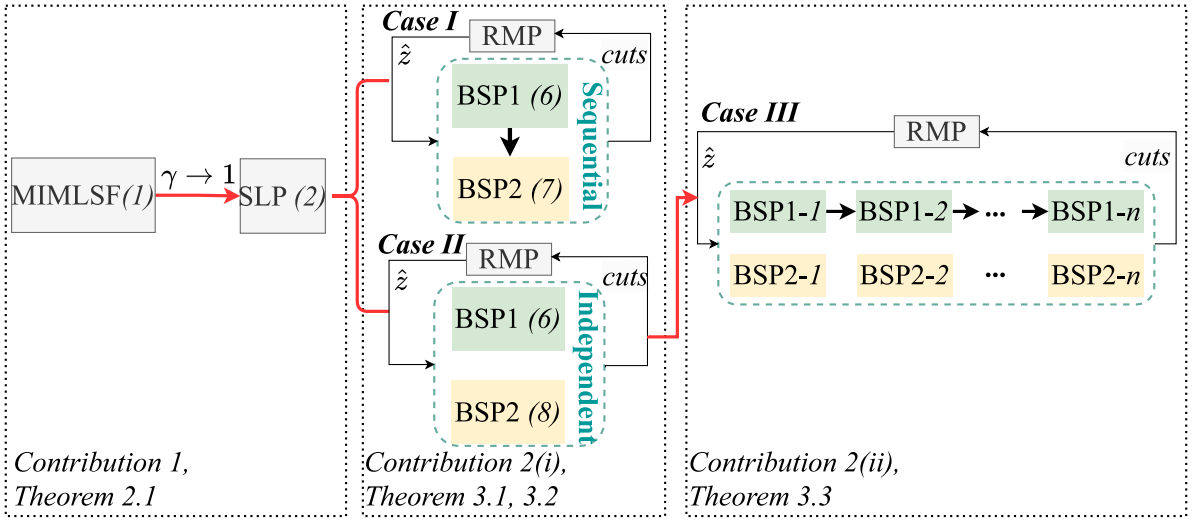

Fig. 2. Theoretical Contributions. The equation numbers in parentheses denote the complete mathematical formulation of the optimization model. Case I: general formulation; Case II: coefficients in Problem (1) satisfy $c_{x1}^T = c_1^T, c_{xi}^T = 0, \forall i \in [n]^+$; Case III: Case II without the *upper dual complicating constraints* (1e). The first block corresponds to Contribution **C1** (Section 2), while the second and third blocks correspond to Contribution **C2** (Section 3). RMP: Relaxed Master Problem; BSP: Benders Subproblem.

market decisions occurring across distinct temporal, spatial, and sectoral dimensions. Addressing cross-scale interactions demands modeling frameworks capable of representing hierarchical and sequential decision structures, coupled with efficient solution methods that deliver high-quality results. Despite recent progress, significant methodological gaps remain in developing scalable approaches capable of handling the added complexity of multi-scale market operations. This need for scalable and efficient decision frameworks is consistent with broader transformations in industrial systems under Industry 4.0, where data-driven and automated decision-making tools are increasingly adopted to manage large-scale, interconnected operations (Addula & Sekhar Sajja, 2024)

To bridge this gap, this paper proposes a multi-level solution framework and validates it through a real-world case study. The contributions of this paper are threefold: the theoretical contributions (**C1** and **C2**) are illustrated in Fig. 2, where each block represents a key theoretical component of the proposed framework, while the modeling contribution (**C3**) is demonstrated through a real-world case study on integrated electricity–gas–carbon market operations in the Northeastern United States:

C1 *Single-level approximation of the MIMLSF problem*: We investigate the Mixed-Integer Multi-Level problems with Sequential Followers (MIMLSF) (i.e., a n+1-level mixed-integer problem with n sequential followers problem), as presented in Fig. 1(d). We propose a lexicographic optimization combined with a weighted-sum approach to transform the MIMLSF problem into a single-level problem (SLP), which asymptotically approximates the original MIMLSF problem as the scaling factor γ approaches 1, defined in Section 2 (see the first block of Fig. 2).

C2 *Dedicated Benders subproblem decomposition for the SLP*: We propose a dedicated Benders decomposition approach for the single-level MIMLSF approximation problem (i.e., SLP). The key contributions are: (i) proving that the complex Bender Subproblem (BSP) can be separated into two more tractable problems (BSP1 and BSP2), which can be solved either sequentially or independently, depending on the upper-level objectives (see the second block of Fig. 2); and (ii) demonstrating that these separated BSPs can be further decomposed into n problems while addressing numerical challenges (see the third block of Fig. 2).

C3 *A practical MIMLSF case study on an integrated electricity–gas–carbon market framework*: We demonstrate our MIMLSF approach through a four-level Unit Commitment with Gas and Carbon Awareness (UCGCA) case study in the Northeastern United States networks. By incorporating economic feedback from gas and carbon markets, the UCGCA model makes improved UC decisions that mitigate gas price spikes and prevent GFPP financial losses while reducing total operational costs compared to existing approaches. The computational experiments indicate that the proposed dedicated Benders decomposition method significantly outperforms the direct centralized approach implemented with a state-of-the-art solver.

$$\textbf{MIMLSF}: \min_{\substack{z\in\{0,1\}^m,\\ x_1\geq 0, y_1\geq 0,\ldots,\\ x_n\geq 0, y_n\geq 0}} \quad c_z^T z + \sum_{i=1}^{n} c_{xi}^T x_i \tag{1a}$$

$$\text{s.t.} \quad z \in \mathcal{Z}, z \in \{0,1\}^m, \tag{1b}$$

$$H_{xi} z + G_{xi} x_i \geq b_{xi}, \quad \forall i \in [n], \tag{1c}$$

$$H_{yi} z + G_{yi} y_i \geq b_{yi}, \quad \forall i \in [n], \tag{1d}$$

$$R_y z + \sum_{i=1}^{n} S_{yi} y_i \geq q_y, \tag{1e}$$

$$(x_1, y_1, \ldots, x_n, y_n) \in \arg\min c_1^T x_1 \tag{1f}$$

$$\text{s.t.} \quad A_1 x_1 + D_1 z \geq b_1, \tag{1g}$$

$$(x_2, y_2, \ldots, x_n, y_n) \in \arg\min c_2^T x_2 \tag{1h}$$

$$\text{s.t.} \quad A_2 x_2 + B_2 x_1 + D_2 z \geq b_2, \tag{1i}$$

$$\ldots\ldots\ldots\ldots$$

$$(x_n, y_n) \in \arg\min c_n^T x_n \tag{1j}$$

$$\text{s.t.} \quad A_n x_n + B_n x_{n-1} + D_n z \geq b_n. \tag{1k}$$

Box I.

The rest of the paper is organized as follows: Section 2 introduces the MIMLSF model and we prove that it can be asymptotically approximated as an SLP. Section 3 presents the dedicated Benders decomposition with further decomposition on the Benders subproblems for the SLP. Section 4 presents the UCGCA formulation and results for the UCGCA case studies are presented in Section 5. Section 6 discusses the policy and operational implications of the proposed model and its case study findings. Lastly, Section 7 concludes the paper.

2. Single-level approximation of the MIMLSF problem

In this section, we present the general MIMLSF formulation and prove its asymptotic approximation to a single-level problem. This section corresponds the first block of Fig. 2. In what follows, the notation $[n]$ represents the set $\{1, \ldots, n\}$, and $[n]^+$ denotes the set $\{2, \ldots, n\}$ for an integer $n \geq 2$, where n indicates the number of sequential problems. The symbol $\mathbf{1}$ represents the vector of ones. Square brackets around variables indicate the dual variables associated with each constraint.

The MIMLSF problem, characterized as an n+1-level structure with n sequential followers (where '1' is the upper level), is presented in Problem (1) and illustrated in Fig. 1(d). To improve readability, the problem at level i+1 is referred to as the ith lower-level problem, where $i \in [n]$ indicates its sequential position. The term '*lower-level*' is used to denote the position of n sequential problems within the hierarchy of decision-making. We use boldface letters to represent vectors of variables, such as $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$, where $\boldsymbol{x}_i$ indicates the primal variable vector of the ith lower-level problem, and $\boldsymbol{y}_i$ represents the dual variable vector. A summary of the notation is provided in Table 2 (see Box I).

In the upper-level problem, the decision variables are modeled as an m-dimensional binary vector, $\boldsymbol{z}$, where $\boldsymbol{z} \in \{0,1\}^m$. The feasible region, denoted as $\mathcal{Z}$, encompasses all binary vectors of dimension m, where each component z_i of the vector $\boldsymbol{z}$ can take a value of either 0 or 1. Furthermore, the vectors $\boldsymbol{x}_i \in \mathbb{R}^{n_{xi}}$ and $\boldsymbol{y}_i \in \mathbb{R}^{n_{yi}}$ are defined for each $i \in [n]$, representing the primal and dual variables of the ith lower-level problem, respectively. The mathematical representation of (1) with appropriate dimensions for any specific problem can be used to derive the vectors $(c_z, c_{xi}, b_{xi}, b_{yi}, q_y, c_i, b_i)$ and matrices $(H_{xi}, G_{xi}, H_{yi}, G_{yi}, R_y, S_{yi}, A_i, B_{i^+}, D_i)$, $\forall i \in [n]$, $\forall i^+ \in [n]^+$. The upper-level constraints (1c) and (1d) are defined as *upper primal coupling constraints* and *upper dual coupling constraints* for each $i \in [n]$, respectively. Moreover, constraint (1e) is defined as *upper dual complicating constraints* as it accounts for the cumulative impact of the dual variables $\boldsymbol{y}_i$ from all lower-level problems.

The binary decision variables $\boldsymbol{z}$ from the upper-level problem are passed onto n sequential lower-level constraints, as suggested in constraints (1g), (1i) (1k). In addition, the *nomination* determined in the $(i-1)^{th}$ lower-level problem will be fed into the ith (next) lower-level problem and is modeled by the constraints for ith (next) lower-level problem $A_i x_i + B_i x_{i-1} + D_i z \geq b_i$ where $i \in [n]^+$. Sequential followers framework requires that the $(i-1)^{th}$ lower-level problem does not anticipate the solutions of the ith (next) lower-level problem $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ since constraints and objective function of the $(i-1)^{th}$ lower-level problem are not dependent on the variables of the ith lower-level problem. In this formulation, *sequential followers* refer to the hierarchical decision-making stages in the problem structure, and their decision-making processes are modeled in Eqs. (1f)–(1k). The corresponding mathematical optimization formulations are referred to as *sequential lower-level problems*.

Without loss of generality, we present the lower-level problems in (1) with linear constraints. However, the models can be extended to convex lower-level problems (e.g., second-order cone programming (SOCP)) by replacing the linear constraints with appropriate conic forms. This extension is particularly relevant for applications in radial distribution networks (Savelli & Morstyn, 2021; Xia et al., 2025) and gas network problems (Byeon & Van Hentenryck, 2020).

This MIMLSF (1) is formulated as a nested multi-level optimization problem that captures hierarchical coordination in multi-scale energy markets. Although these markets are structurally interdependent, they are not jointly optimized in practice; instead, each layer clears conditional on prior outcomes. The proposed MIMLSF reflects this institutional structure while enabling coordinated decision-making through an integrated mathematical representation of cross-sector and intertemporal dependencies. Within this framework, the upper-level decision anticipates and incorporates the lower-level best-response mapping, ensuring that the leader's optimization is constrained by the followers'

Table 2
Notation summary for the MIMLSF and the SLP.

Symbol	Type	Description
$\boldsymbol{z}$	Decision variable	Binary decision vector at the upper level
$\boldsymbol{x}_i$	Decision variable	Vector of primal variables for follower i
$\boldsymbol{y}_i$	Decision variable	Vector of dual variables for follower i
$\boldsymbol{u}_{xi}, \boldsymbol{u}_{yi}, \boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}_{yi}$	Decision variable	Vectors of dual variables associated with upper-level constraints, strong-duality conditions, and linearization constraints in the SLP
$\boldsymbol{s}_{yi}$	Decision variable	Auxiliary variable vector introduced for bilinear term linearization in the SLP
A_i, B_i, D_i	Parameter	Lower-level constraint matrices for follower i
H_{xi}, G_{xi}	Parameter	Upper-level primal coupling matrices associated with follower i
H_{yi}, G_{yi}	Parameter	Upper-level dual coupling matrices associated with follower i
R_y, S_{yi}	Parameter	Upper-level dual complicating matrices
Y_{yi}, K_{yi}, E_{yi}	Parameter	Matrices associated with bilinear term linearization in the SLP
$b_i, b_{xi}, b_{yi}, q_y, e_{yi}$	Parameter	Right-hand-side vectors of the corresponding constraints
c_z, c_{xi}	Parameter	Upper-level objective coefficient vectors
c_i	Parameter	Objective coefficient vectors of follower i
γ	Parameter	Asymptotic scaling parameter
γ_i	Parameter	Scaling factor for follower i

optimal reactions. Section 4 presents a detailed four-level UCGCA example to demonstrate the practical applicability of the model.

We make the following assumptions throughout this paper:

Assumption 1. The following problem is feasible and bounded.

$$\min_{\substack{z\in\{0,1\}^m, z\in\mathcal{Z},\\ x_1\geq 0, y_1\geq 0,\ldots,\\ x_n\geq 0, y_n\geq 0}} \quad c_z^T z + \sum_{i=1}^{n} c_{xi}^T x_i$$

s.t. *Eqs.* (1c), (1d), (1e), (1g), (1i), (1k) : All primal constraints

Assumption 1 provides a lower bound on the optimal objective value of Problem (1) by relaxing the optimality of the sequential lower-level problems.

Assumption 2. The dual of each follower problem is feasible once the leader variables and the realized preceding lower-level primal variables are given.

Since the feasible regions of the followers' dual problems are independent of both $\hat{z}$ and the preceding lower-level primal solutions $x_{i-1}, \forall i \in [n]^+$, this assumption ensures that each follower problem is bounded from below for any admissible combination of $\hat{z}$ and preceding-stage outcomes. Consequently, strong duality holds between the corresponding primal and dual formulations at every stage of the sequential multi-level framework. This is a reasonable assumption because, otherwise, the follower problem would be infeasible or unbounded for any admissible leader decision and preceding-stage realization. In the bilevel optimization problem, if the lower-level problem is infeasible for a feasible upper-level decision, then no bilevel-feasible solution exists. Likewise, if there is a feasible upper-level solution such that the corresponding lower-level problem is unbounded below, then the bilevel problem is infeasible (Xu & Wang, 2014). In the MIMLSF setting, the same statement holds, with respect to a feasible upper-level decision and the realized optimal responses from preceding lower-level markets. In our case study, follower problems remain feasible and bounded under the considered system setting, as further discussed in Section 5. In general, infeasibility or unboundedness of a follower problem can be detected through feasibility checks and refinement procedures to ensure that the multi-level problem remains well-defined (Fischetti et al., 2018).

We also assume that each follower stage clears in a timely manner upon receiving upstream signals. This reflects the institutional design of sequential and time-constrained market operations, where clearing is required to deliver an outcome within regulated deadlines. Even under rapidly changing conditions, market rules and established operating protocols, together with predefined operating envelopes, provide operational safeguards that ensure the clearing stage can still return a feasible and economically efficient decision. Nevertheless, if substantial inter-market delays occur, the resulting uncertainty motivates future extensions of the proposed framework toward uncertainty-aware multi-level formulations. Such extensions could incorporate stochastic economic feedback from lower-level multi-scale market-clearing processes to represent delay-induced information lags.

Theorem 2.1. *The MIMLSF problem (1) can be asymptotically approximated by the following single-level problem (SLP) (2) for some $\gamma \in (0, 1)$. Moreover, when $\gamma \to 1$, the optimal solution of problem (2) converges to the optimal solution of problem (1). In SLP (2), the scaling factor applied for the ith lower-level/sequential problem is denoted by γ_i, with $\gamma_1 = \gamma$, $\gamma_2 = \gamma(1-\gamma)$, and continuing as $\gamma_n = (1-\gamma)^{n-1}$. The terms enclosed in squared brackets are dual variables.*

$$\textbf{SLP}: \min_{\substack{z\in\{0,1\}^m,\\ x\geq 0, y\geq 0,\\ u_x\geq 0, u_y\geq 0,\\ u\geq 0, s_y\geq 0,\\ w\geq 0, v_y\geq 0}} \quad \gamma c_z^T z + \sum_{i=1}^{n} \gamma c_{xi}^T x_i \tag{2a}$$

$$s.t. \quad z \in \mathcal{Z}, z \in \{0,1\}^m, \tag{2b}$$

$$H_{xi} z + G_{xi} x_i \geq b_{xi}, \quad \forall i \in [n], \quad [u_{xi} \geq 0] \tag{2c}$$

$$\gamma_i H_{yi} z + G_{yi} y_i \geq \gamma_i b_{yi}, \quad \forall i \in [n], \quad [u_{yi} \geq 0] \tag{2d}$$

$$R_y z + \sum_{i=1}^{n} S_{yi} \frac{y_i}{\gamma_i} \geq q_y, \quad \forall i \in [n], \quad [u \geq 0] \tag{2e}$$

$$A_1 x_1 + D_1 z \geq b_1, \quad [y_1 \geq 0] \tag{2f}$$

$$A_i x_i + B_i x_{i-1} + D_i z \geq b_i, \quad \forall i \in [n]^+, \quad [y_i \geq 0] \tag{2g}$$

$$y_i^T A_i + y_{i+1}^T B_{i+1} \leq \gamma_i c_i^T, \quad \forall i \in [n-1], \quad [x_i \geq 0] \tag{2h}$$

$$y_n^T A_n \leq \gamma_n c_n^T, \quad [x_n \geq 0] \tag{2i}$$

$$\sum_{i=1}^{n} \left(y_i^T b_i - s_{yi}^T \mathbf{1} \right) \geq \sum_{i=1}^{n} \gamma_i c_i^T x_i \quad [w \geq 0] \tag{2j}$$

$$Y_{yi} y_i + K_{yi} s_{yi} \geq \gamma_i (e_{yi} + E_{yi} z), \quad \forall i \in [n], \quad [v_{yi} \geq 0] \tag{2k}$$

Proof. Please refer to Appendix A. □

In Problem (2), the upper-level constraints (2b)–(2e) correspond to those in (1b)–(1e), where the dual variables y_i are rescaled by γ_i to recover their original scales. The lower-level primal constraints (2f)–(2g) are identical to (1g), (1i), and (1k). The corresponding dual constraints (2h)–(2i) involve y_i as the dual variables associated with the lower-level primal problem. The proposed formulation employs a *single linear* strong duality constraint (2j), which, together with the primal and dual feasibility conditions, exactly enforces the lower-level optimality conditions (see proof in Appendix A). The bilinear terms $y_i^T D_i z$ in the strong duality condition (2j) are linearized via the exact

McCormick relaxation technique in Problem (2), introducing auxiliary variables $s_{yi}^T\mathbf{1}$ satisfying $s_{yi}^T\mathbf{1} = y_i^T D_i z$ and enforcing the additional constraints specified in (2k). The number of introduced auxiliary variables and constraints depends on the sparsity pattern of D_i, as each nonzero entry induces a corresponding bilinear interaction between y_i and z. For bilinear terms involving binary variables z and bounded continuous variables y_i, the McCormick linearization (also known as the Big-M method in this setting) provides an exact linearization of the original bilinear terms.

In this section, Theorem 2.1 establishes a single-level asymptotic reformulation of MIMLSF, clarifying how the original multi-level problem can be approximated by a SLP. The resulting SLP formulation provides the structural foundation for the solution approach developed in the next section. The aforementioned linearization step introduces additional variables and constraints, but it does not constitute the primary source of the problem's computational complexity. The large-scale nature of the problem arises from the embedded binary decisions and hierarchical structural coupling inherent to the original formulation. To address this intrinsic mixed-integer complexity, a dedicated Benders decomposition scheme will be developed in the next section.

3. Dedicated Benders decomposition with multi-level subproblem separability

In this section, we propose a dedicated Benders decomposition approach with enhanced subproblem separability to effectively solving the single-level problem (2), building upon Byeon and Van Hentenryck (2022). This section relates to the second and third blocks of Fig. 2. While standard solvers may be capable of solving the single-level approximation of a tri-level model with sequential middle and lower levels of moderate complexity (i.e., $n = 2$), as indicated in Xia et al. (2025) and Paredes et al. (2023), they might struggle when addressing more intricate, larger-scale, or deeper hierarchical (i.e., multi-level) problems. In this case, Benders decomposition offers an alternative for solving the large-scale complex SLP (2).

Benders decomposition involves iteratively solving a Relaxed Master Problem (RMP) with binary variables and a Benders Subproblem (BSP) with continuous variables to find the optimal solution of the original problem by introducing cuts based on the feasibility and optimality of subproblems until the upper and lower bounds converge within a predefined tolerance. For a detailed review of Benders decomposition, see Rahmaniani et al. (2017). To facilitate the application of Benders decomposition, (2) is initially reformulated as follows with a given $\hat{z}$:

$$\min_{\hat{z}} \quad \gamma c_z^T \hat{z} + f(\hat{z}) \tag{3a}$$

$$\text{s.t.} \quad \hat{z} \in \mathcal{Z}, \hat{z} \in \{0,1\}^m, \tag{3b}$$

where $f(\hat{z})$ is defined as

$$f(\hat{z}) := \min_{\substack{x\geq 0, y\geq 0,\\ u_x\geq 0, u_y\geq 0, u\geq 0,\\ s_y\geq 0, w\geq 0, v_y\geq 0}} \sum_{i=1}^{n} \gamma c_{xi}^T x_i \tag{4a}$$

$$\text{s.t.} \quad \text{Eq. (2c)} - \text{Eq. (2k)} \textit{ with } z \textit{ replaced by } \hat{z} \tag{4b}$$

For a guess $\hat{z}$, then the corresponding BSP of Problem (2) is derived by taking the dual of Problem (4):

$$\mathbf{BSP}: \max_{\substack{x\geq 0, y\geq 0, v_y\geq 0,\\ w\geq 0, u\geq 0, u_x\geq 0,\\ u_y\geq 0, s_y\geq 0}} \sum_{i=1}^{n} \left(y_i^T(b_i - D_i\hat{z}) + u_{xi}^T(b_{xi} - H_{xi}\hat{z})\right) - \left[\sum_{i=1}^{n}\left(\gamma_i c_i^T x_i - \gamma_i u_{yi}^T(b_{yi} - H_{yi}\hat{z}) - \gamma_i v_{yi}^T(e_{yi} + E_{yi}\hat{z})\right) - u^T(q_y - R_y\hat{z})\right] \tag{5a}$$

$$\text{s.t.} \quad y_i^T A_i + y_{i+1}^T B_{i+1} + G_{xi}^T u_{xi} \leq \gamma_i c_i^T w + \gamma c_{xi}^T, \quad \forall i \in [n-1], \tag{5b}$$

$$y_n^T A_n + G_{xn}^T u_{xn} \leq \gamma_n c_n^T w + \gamma c_{xn}^T, \tag{5c}$$

$$A_1 x_1 - G_{y1}^T u_{y1} - Y_{y1}^T v_{y1} - S_{y1}^T u/\gamma_1 \geq b_1 w, \tag{5d}$$

$$A_i x_i + B_i x_{i-1} - G_{yi}^T u_{yi} - Y_{yi}^T v_{yi} - S_{yi}^T u/\gamma_i \geq b_i w, \quad \forall i \in [n]^+, \tag{5e}$$

$$K_{yi}^T v_{yi} \leq w, \quad \forall i \in [n]. \tag{5f}$$

While the single-level problem (2) can be iteratively solved through the RMP and BSP (5), the latter is computationally intensive because it involves constraints associated with both the primal variables x_i ($i \in [n]$, see (5d) and (5e)) and the dual variables y_i ($i \in [n]$, see (5b) and (5c)). Moreover, the complexity of BSP (5) scales with the number of lower-level problems n. The strong duality condition (2j) acts as a complicating constraint in the SLP (2), coupling the n sequential lower-level problems. Its corresponding dual variable w therefore serves as a linking variable in BSP (5), appearing on the right-hand side of each constraint. The following subsections develop solution techniques to efficiently solve the complex BSP (5), addressing three cases with distinct upper-level objectives (1a) and constraint requirements on (1e), as summarized in Table 3.

3.1. Benders subproblem decomposition for Case I

This subsection discusses the methodologies outlined in the upper second blocks of Fig. 2 for Case I. To further decompose the Benders subproblem (5), we show that the complex BSP (5) can be solved by sequentially solving two more tractable problems. In other words, the Benders cuts of (2) can be obtained by solving two problems that are more tractable than the original Benders subproblem.

Theorem 3.1. *Benders subproblem (5) can be solved by two more tractable problems sequentially. First, solve problem (6)*

$$\mathbf{BSP1}: \min_{\substack{x\geq 0, u_y\geq 0,\\ v_y\geq 0, u\geq 0}} \sum_{i=1}^{n}\left(\gamma_i c_i^T x_i - \gamma_i u_{yi}^T(b_{yi} - H_{yi}\hat{z}) - \gamma_i v_{yi}^T(e_{yi} + E_{yi}\hat{z})\right) - u^T(q_y - R_y\hat{z}) \tag{6a}$$

$$\text{s.t.} \quad A_1 x_1 - G_{y1}^T u_{y1} - Y_{y1}^T v_{y1} - S_{yi}^T u/\gamma_1 \geq b_1, \tag{6b}$$

$$A_i x_i + B_i x_{i-1} - G_{yi}^T u_{yi} - Y_{yi}^T v_{yi} - S_{yi}^T u/\gamma_i \geq b_i, \quad \forall i \in [n]^+, \tag{6c}$$

$$K_{yi}^T v_{yi} \leq \mathbf{1}, \quad i \in [n]. \tag{6d}$$

Then, solve problem (7) where $\mathfrak{D}_6$ is the optimal value of problem (6) if it has a finite optimum, ∞ otherwise (i.e., by fixing $w = 0$):

$$\mathbf{BSP2}: \max_{\substack{y\geq 0, u_x\geq 0,\\ w\geq 0}} \sum_{i=1}^{n}\left(y_i^T(b_i - D_i\hat{z}) + u_{xi}^T(b_{xi} - H_{xi}\hat{z})\right) - w\mathfrak{D}_6 \tag{7a}$$

$$\text{s.t.} \quad y_i^T A_i + y_{i+1}^T B_{i+1} + G_{xi}^T u_{xi} \leq \gamma_i c_i^T w + \gamma c_{xi}^T, \quad \forall i \in [n-1], \tag{7b}$$

$$y_n^T A_n + G_{xn}^T u_{xn} \leq \gamma_n c_n^T w + \gamma c_{xn}^T. \tag{7c}$$

Proof. Please refer to Appendix D of the supplementary material. □

Theorem 3.1 suggests that the BSP (5) can be solved by two tractable problems (6) and (7), and the RMP with feasibility and optimality cuts are given in Corollary 1 and the dedicated Benders decomposition algorithm is presented in Algorithm 1. For further details on Corollary 1 and Algorithm 1, please see Appendix E in the supplementary material.

3.2. Benders subproblem decomposition for Case II

In this section, we examine a specific instance of Case I, referred to as Case II, which allows for a significantly strong alternative to Theorem 2.1. This subsection details the methodologies presented in the lower second blocks of Fig. 2 for Case II. We examine a special configuration of the upper-level objectives represented by $c_z^T z + c_1^T x_1$, i.e., $c_{x1}^T = c_1^T$, $c_{xi}^T = 0$ for all $i \in [n]^+$. These conditions can be approximated as $c_{xi}^T = \gamma_i c_i^T$ as γ approaches 1. This special configuration

Table 3
Summary of cases and their descriptions.

Case	Upper-level objective requirements	Constraint requirements	Extra notes
I	$c_z^T \boldsymbol{z} + \sum_{i=1}^{n} c_{xi}^T \boldsymbol{x}_i$	None	None
II	$c_z^T \boldsymbol{z} + c_1^T \boldsymbol{x}_1$	None	Special case of Case I
III	$c_z^T \boldsymbol{z} + c_1^T \boldsymbol{x}_1$	No (1e)	Special case of Case II

is adopted by the UCGCA model introduced in the next section. We denote the BSP under this special condition is BSP' (5)' where the RHS of constraints (5b) and (5c) becomes $\gamma_i c_i^T (\boldsymbol{w}+1)$, $\forall i \in [n]$. We prove that, unlike the approach specified in Theorem 2.1 which involves solving two problems *sequentially*, the BSP' (5)' can be addressed through two *independent* problems.

Theorem 3.2. *BSP (5) can be decomposed into two more tractable problems, (6) and (8), which can be solved separately.*

$$\textbf{BSP2}: \quad \max_{\boldsymbol{y}\geq 0, \boldsymbol{u}_x \geq 0} \sum_{i=1}^{n} \boldsymbol{y}_i^T (b_i - D_i \hat{\boldsymbol{z}}) + \boldsymbol{u}_{xi}^T (b_{xi} - H_{xi} \hat{\boldsymbol{z}}) \tag{8a}$$

$$s.t. \quad \boldsymbol{y}_i^T A_i + \boldsymbol{y}_{i+1}^T B_{i+1} + G_{xi}^T \boldsymbol{u}_{xi} \leq \gamma_i c_i^T, \quad \forall i \in [n-1], \tag{8b}$$

$$\boldsymbol{y}_n^T A_n + G_{xn}^T \boldsymbol{u}_{xn} \leq \gamma_n c_n^T. \tag{8c}$$

Proof. Please refer to Appendix F of the supplementary material. □

The proof of Theorem 3.2 implies that the Benders cuts for the SLP (2) with $c_{xi}^T = \gamma_i c_i^T$ conditions can be obtained by solving Problems (6) and (8) *separately* and comparing their objective values. This simplifies the Benders cut generation algorithm as described in Algorithm 2 and the RMP is presented in Corollary 2. For further details on Corollary 2 and Algorithm 2, please see Appendix G in the supplementary materials.

3.3. Benders decomposed subproblems with further multi-level separation for Case III

In this section, we investigate a particular instance of Case II, referred to as Case III, that introduces advanced features building on Theorem 3.2. This subsection examines the approach suggested in the third blocks of Fig. 2. While Theorem 3.2 demonstrates that the computation of BSPs with the special upper-level objectives can be theoretically separated into two more tractable problems (6) and (8) (i.e., decomposed into a primal-related problem (6) and a dual-related problem (8)), the complexity of solving these problems still increases with the number of lower-level problems n. Additionally, the use of the weighted-sum method introduces scaling factors γ_i to the objective functions for variables $(\boldsymbol{x}, \boldsymbol{u}_x, \boldsymbol{u}_y)$ in the subproblems defined by (6) and the dual variables $\boldsymbol{y}$ in (8) are scaled accordingly. This scaling may introduce numerical issues, particularly as γ approaches 1 (and therefore $(1-\gamma) \rightarrow 0$) and as the number of lower-level problems increases. To address this issue, we propose the following Theorem:

Theorem 3.3. *BSP1 (6) can be further decomposed into n problems that can be solved sequentially if there are no upper dual complicating constraints (1e) and BSP2 (8) can be further decomposed into n problems and solved in parallel.*

Proof. Please refer to Appendix H of the supplementary material. □

Theorem 3.3 demonstrates that when solving BSP (6) with a fixed upper-level decision $\hat{\boldsymbol{z}}$ from the master problem, only connecting variables (e.g. $\boldsymbol{x}_i$, $i \in [n-1]$) need to be exchanged between sequential lower-level problems, limiting the information flow between them. Without it, solving the single-level problem (2) would require directly incorporating variables and constraints from n lower-level agents into BSP1 (6) and BSP2 (8). Such direct incorporation would require scaling factors to capture the sequential relationships between lower-level problems, leading to significant numerical scalability challenges. Instead, Theorem 3.3 enables a more scalable solution approach. It allows BSP (6) and BSP (8) to be solved either sequentially or independently by leveraging the interpretation of scaling factors across n problems. This approach eliminates computational burden, reduces information exchange, and avoids the scalability issues associated with solving BSP (6) and BSP (8) as n coupled lower-level problems.

4. Practical implementation of MIMLSF: Four-level unit commitment with gas and carbon awareness model

In this section, we provide an application of the MIMLSF model by proposing a four-level Unit Commitment with Gas and Carbon Awareness (UCGCA) model that coordinates the electricity, gas, and carbon markets. The design of the bid-validity constraints for GFPPs and the coupling between three markets are introduced in Section 4.1. The relationship between the UCGCA model and the MIMLSF problem is then discussed in Section 4.2. The mathematical formulations for UC, economic dispatch (ED), gas market clearing, and carbon market clearing problems are detailed in Appendix B. The ED problem is also referred to as the electricity market clearing problem.

Fig. 3 illustrates the interactions among the electricity, gas, and carbon markets. Fig. 3(a) illustrates the proposed four-level hierarchical UCGCA model, which incorporates bid-validity constraints for GFPPs. In contrast, Fig. 3(b) presents a non-hierarchical benchmark model (BM) without gas and carbon coordination, which may lead to invalid GFPP bids and financial losses. The four-level UCGCA illustrates the hierarchical and cross-market propagation of decisions across electricity, gas, and carbon markets. The upper-level UC determines the commitment status of generating units and their eligibility for dispatch. The resulting electricity dispatch specifies the output of GFPPs, which determines gas demand through the heat-rate relationship. Electricity generation and associated gas consumption jointly determine carbon emissions. Deviations from allocated allowances induce carbon allowance trading. By incorporating downstream price signals into early-stage UC decisions, the proposed UCGCA model improves cross-market coordination and reduces financial risk. In contrast, the BM model follows a three-stage sequential structure in which stages are cleared consecutively without incorporating anticipated downstream price signals into UC decisions, and thus lacks cross-market coordination.

Specifically, in the UCGCA model, the upper-level UC problem is managed by Independent System Operators (ISOs). The commitment decisions, including the on/off statuses $o_{u,t}$, and indicators for generator u start-ups $v_{u,t}^{+}$ and shut-downs $v_{u,t}^{-}$ at time t, are fed into the ED problem (i.e., electricity market clearing) to determine the power output $p_{u,t}$. This power schedule $p_{u,t}$ for GFPPs generates a gas demand $l_{j,t}^{GFPP}$ at the corresponding gas network junction j, which is approximated using the heat rate curve (9), as further discussed in Section 4.1. The power output $p_{u,t}$ and the satisfied gas demand $l_{j,t}$ lead to carbon emissions, enabling participants to trade surplus allowances or purchase additional allowances in the carbon market. A net allowance greater than zero indicates a surplus, allowing participants to sell on the market for a profit; conversely, a deficit requires purchasing allowances, incurring costs. The derived dual solutions, including zonal gas prices $\psi_{k,t}$ and carbon prices π_t^e, are then incorporated into the upper-level bid-validity constraints to ensure only profitable GFPPs are committed. The design of this bid-validity constraint is discussed in Section 4.1.

The BM is structured as a three-stage sequential optimization problem: initially, the UC and ED are solved to determine the power outputs

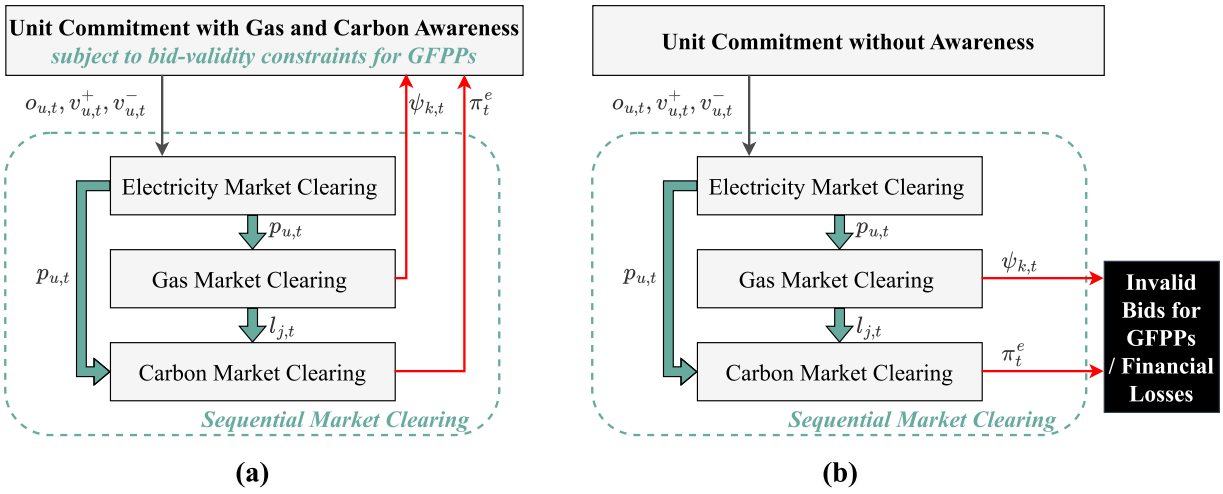

Fig. 3. Unit commitment with and without gas and carbon awareness. (a) UCGCA: A four-level hierarchical unit commitment framework in which upper-level commitment decisions anticipate sequential electricity–gas–carbon market clearing outcomes. Gas and carbon price signals are incorporated into the upper-level via bid-validity constraints, ensuring that GFPP commitments remain financially viable and mitigating financial risk exposure under volatile gas and carbon prices. (b) BM: A conventional unit commitment model without cross-market coordination. Commitment decisions are made without anticipating downstream gas and carbon market outcomes; under elevated gas and carbon prices, this may lead to economically infeasible GFPP bids and associated financial losses.

of GFPPs. These outputs then inform the gas demand induced by GFPPs $l^{GFPP}_{j,t}$ required for gas market clearing, which is calculated using the heat rate curve (9). Subsequently, the carbon market is cleared based on actual emissions and submitted bids. Fig. 3(b) shows the three-stage BM structure without any bid validity constraints specifically for GFPPs, and invalid bids can be calculated after the three-stage market clearing. Notably, the BM does not account for the economic viability of GFPPs. These plants submit their bids in the electricity market prior to the realization of gas and carbon prices. As a result, they may incur financial losses from invalid bids, especially under conditions of high stress in gas and electricity networks. The fundamental difference between the UCGCA and BM frameworks is that the UCGCA anticipates gas and carbon market outcomes within the UC process, whereas the BM clears these markets sequentially without such awareness, leaving GFPPs vulnerable to price volatility and financial losses, as observed during the 2014 polar vortex (PJM, 2014) in the Northeastern United States.

4.1. The physical and economic coupling of three markets and the bid validity constraints

The physical coupling between electric and gas network is expressed through the heat rate curve (9) for GFPPs, which is approximated as a quadratic function relating gas consumption $l^{GFPP}_{j,t}$ to electricity generation $p_{u,t}$:

$$l^{GFPP}_{j,t} = \sum_{u \in \mathcal{U}(j) \cap \mathcal{U}^g} H^G_{u,2} p^2_{u,t} + H^G_{u,1} p_{u,t} + H^G_{u,0}, \quad \forall j \in \mathcal{V}, t \in \mathcal{T}. \tag{9}$$

The physical coupling between the electric and gas markets with the carbon market is facilitated through the power generation level $p_{u,t}$ and the satisfied gas demand $l_{j,t}$. The latter is computed as the difference between the gas demand profile $d^g_{j,t}$ and the gas load shed $q_{j,t}$. Specifically, the emission levels are determined by the Carbon Intensity (CI) factors κ_u and κ_j, which are multiplied by their respective cleared quantities, as shown in Eq. (10).

$$E_{u,t} = \kappa_u p_{u,t}, \quad \forall u \in \mathcal{U}, t \in [T], \tag{10a}$$

$$E_{j,t} = \kappa_j l_{j,t}, \quad \forall j \in \mathcal{V}, t \in [T]. \tag{10b}$$

The power generation levels of GFPPs significantly influence the load on the gas system and carbon emission levels and, consequently, natural gas prices and carbon prices. These prices, in turn, determine the profitability of GFPPs, which place their bids in the electricity market prior to realizing gas and carbon prices. These dynamics are captured in the bid-validity constraints (11), designed to compare the GFPPs' marginal bids $\rho_{u,t}$ (see the definition and related constraints of $\rho_{u,t}$ in Appendix B.1) in the electricity market with the economic realities of the gas and carbon markets. It ensures that their participation remains economically viable and profitable by aligning commitment decisions $o_{u,t}$ with anticipated outcomes from gas and carbon markets. α_u represents the risk aversion level of GFPP u, modulating the relationship between the plant's expected profits and the risks it faces in volatile markets. A lower value of α_u indicates a more risk-averse GFPP, meaning it would prefer to de-commit to avoid potential financial losses.

$$\alpha_u \rho_{u,t} \geq \left((2p_{u,t} H^G_{u,2} + H^G_{u,1})\psi_{k,t} + (2y_{u,t} - 1)\kappa_u \pi^e_t\right) o_{u,t}, \quad \forall k \in \mathcal{K}, i \in \mathcal{V}(k), u \in \mathcal{U}(i) \cap \mathcal{U}^g. \tag{11}$$

The right-hand side of Eq. (11) is composed of two main components:

1. Marginal Natural Gas Cost $(2p_{u,t} H^G_{u,2} + H^G_{u,1})\psi_{k,t}$: The term $(2p_{u,t} H^G_{u,2} + H^G_{u,1})$ represents the derivative of the heat rate curve (9), representing the amount of natural gas needed to generate one additional unit of electricity. This term is multiplied by the zonal gas prices $\psi_{k,t}$ to represent the marginal cost of GFPP u.
2. Marginal Carbon Trading Cost or Revenue $(2y_{u,t} - 1)\kappa_u \pi^e_t$: When $y_{u,t} = 1$, indicating the GFPP with deficit allowances which purchases emission allowances in the carbon market, adding a marginal cost $\kappa_u \pi^e_t$ to generate one additional unit of electricity. On the other hand, when $y_{u,t} = 0$, the GFPP with surplus allowances sells its carbon allowances, leading to a negative term of $-\kappa_u \pi^e_t$, which represents the marginal revenue gained.

In the gas market modeling, dual solutions associated with the flux conservation constraints represents the marginal costs at gas junction j. However, the US natural gas spot price is zonal, the zonal natural gas prices $\psi_{k,t}$ at zone k are then computed based on averaging the prices of a subset of junctions (Byeon & Van Hentenryck, 2020). In the case study, we examine two distinct gas pricing zones following Bent et al. (2018), Borraz-Sanchez et al. (2016): the Transco Zone 6 Non NY Zone and the Transco Leidy Zone. The Transco Zone 6 Non NY Zone typically exhibits higher prices due to its location in a major gas consumption

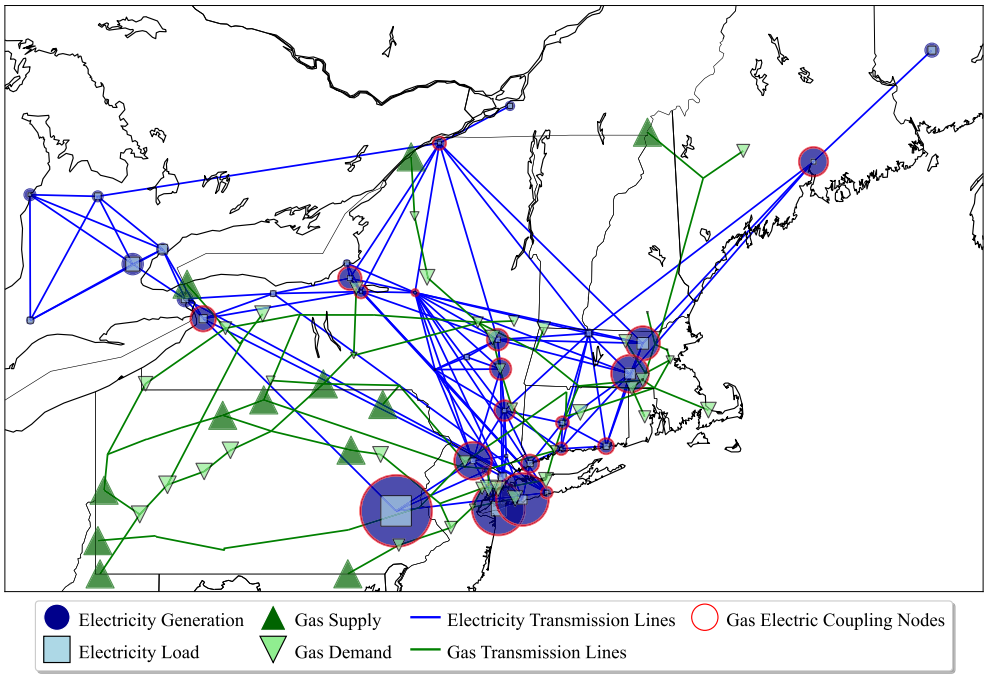

Fig. 4. Electric and gas transmission networks in the Northeastern United States, depicted with simplified point-to-point connections. Node markers are sized proportionally to maximum supply and load capacities at $(\eta_e, \eta_g) = (1, 1)$, with gas system markers shown on a logarithmic scale. GFPPs in the electric network are connected to their nearest natural gas receipt points in the gas network.

area, while the Transco Leidy Zone, situated in the Marcellus Shale production area, generally maintains lower prices due to abundant natural gas supplies. The selection of these zones enables analysis of price dynamics across regions with significant price differentials.

The bid-validity constraint (11) is added to the upper-level UC problem and it introduces nonlinearity into the model and these terms can be linearized by using the big-M technique. The bid-validity constraint (11) enables GFPPs to hedge against the risks associated with high gas prices and emission costs/revenue. This risk management is vital, ensuring that committed GFPPs consider all relevant costs and potential revenues following their bid submissions in the electricity market. By incorporating this constraint into the UC problem, system operators can improve the financial viability of GFPPs and prevent default or large financial losses. This is especially important during extreme demand events, such as the 2014 polar vortex in the Northeastern United States, to avoid defaults or financial losses (PJM, 2014).

4.2. Relationship between the four-level UCGCA model and the MIMLSF problem

The UCGCA model is a specific instance of the MIMLSF problem where $n = 3$, corresponding to the three sequential lower-level problems. The UCGCA model features a specialized structure for the dedicated Benders decomposition as outlined in Section 3.2: the objective function of the upper-level UC problem includes: (1) $c_z^T \boldsymbol{z}$: costs associated with the binary variable $\boldsymbol{z}$, encompassing no-load and start-up costs, (2) $c_{x1}^T \boldsymbol{x}_1$: costs of selected supply bids from electrical power generating units, which is exactly the objective function of the second-level ED problem (i.e., $c_{x1}^T = c_1^T, c_{xi}^T = 0, \forall i \in [n]^+$, the upper-level objective is $c_z^T \boldsymbol{z} + c_1^T \boldsymbol{x}_1$ and the first lower-level objective is $c_1^T \boldsymbol{x}_1$). Therefore, the dedicated Benders decomposition with subproblem separability introduced in Section 3.2 for Case II can be effectively applied for the UCGCA model.

5. Case study

This section presents a case study on the integration of the IEEE 36-bus Northeastern US bulk electric power system (Allen et al., 2008; Bent et al., 2018) with a 117-junction multi-company gas transmission network spanning from Pennsylvania to Northeast New England (Bent et al., 2018; Borraz-Sanchez et al., 2016; Byeon & Van Hentenryck, 2020), as depicted in Fig. 4. In Fig. 4, blue connections represent electricity transmission lines, with blue markers indicating the locations and magnitudes of electricity generation and load levels. Similarly, green connections represent natural gas transmission lines, with green markers indicating the locations and magnitudes of gas supply and demand levels. The gas electric coupling nodes are depicted in red circle which specifies the location of GFPPs and the GFPP of the electric power network were linked to the closest natural gas receipt point in the gas system. As discussed in Section 4.1, we consider two natural gas pricing zones: Transco Zone 6 Non NY Zone and Transco Leidy Line Zone. Fig. 5 shows the pricing junctions for these zones. Transco Zone 6 Non NY Zone are represented by square markers, whereas the pricing points/junctions for Transco Leidy Line Zone are represented by diamond markers. The pricing points are based on previous studies Bent et al. (2018), Byeon and Van Hentenryck (2020) and the GasPowerModels.jl repository https://github.com/lanl-ansi/GasPowerModels.jl. The gas network construction follows Bent et al. (2018), Byeon and Van Hentenryck (2020). Under this setting, fixed (firm) gas demand is relatively small, while the flexible gas production capacity beyond the firm level is large in aggregate. Consequently, total available gas supply is sufficient to meet firm and flexible gas demand, as well as the fuel requirements of GFPPs under normal operating conditions. Under stressed conditions where gas deliverability is limited, flexible gas demand may not be fully satisfied within network constraints; where necessary, high-penalty gas load shedding ensures that feasibility is maintained. Natural gas storage introduces inter-temporal flexibility that can buffer supply–demand mismatches arising from sequential clearing and network constraints; incorporating such dynamics will be addressed in future work.

To analyze system behavior under different market conditions, we varied both electrical and gas load parameters in our case study. We examined electrical load increases of 20% and 60% (i.e., $\eta_e = 1.0, 1.2, 1.6$) and gas load increases ranging from 10% to 130% (i.e., $\eta_g = 1.0, 1.1, \ldots, 2.3$). These parameter variations allowed us to evaluate the

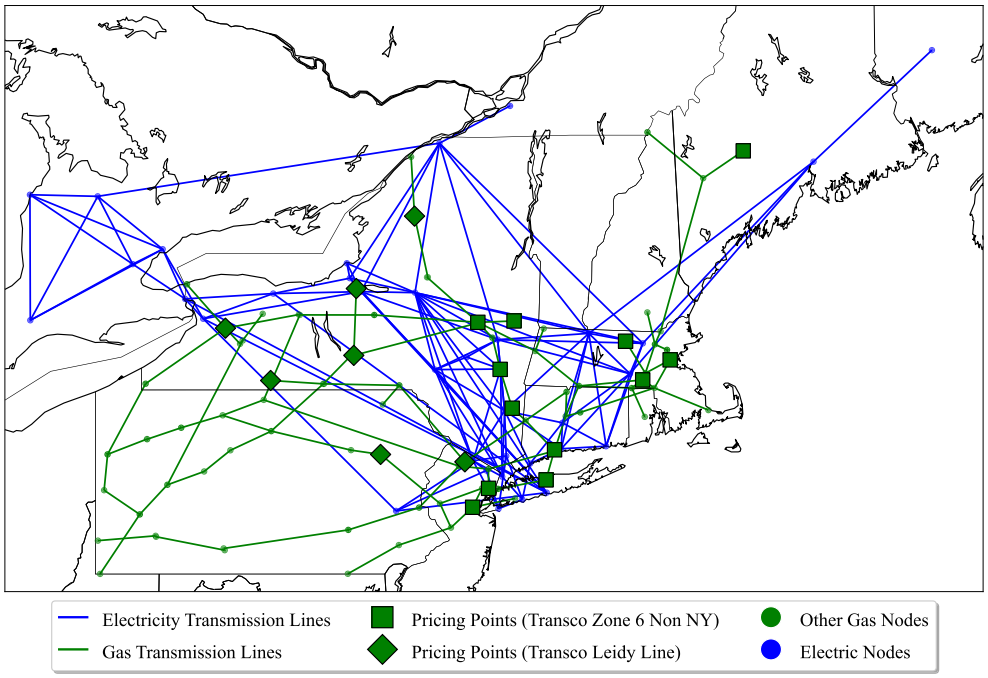

Fig. 5. Electric and Gas Transmission Networks in the Northeastern United States. Diamond markers represent pricing points/junctions in Transco Zone 6 Non NY Zone, while square markers denote those in Transco Leidy Line Zone. The pricing points are obtained based on Bent et al. (2018), Byeon and Van Hentenryck (2020) and the GasPowerModels.jl repository (https://github.com/lanl-ansi/GasPowerModels.jl).

Table 4
Carbon intensity of generators in the electric network, sourced from Savelli et al. (2022). Fuel types include Oil (O), Coal (C), Open Cycle Gas Turbine (G-O), Combined Cycle Gas Turbine (G-C), Hydro (H), Refuse (R), Nuclear (N), and Others (E).

Fuel type	O	C	G-O	G-C	H	R	N	E
CI ($t\mathrm{CO}_2$/MWh)	0.777	0.937	0.651	0.394	0	0.120	0	0.300

economic viability of GFPP under increasing gas costs, assess the effectiveness of bid-validity constraints, and demonstrate the performance advantages of the proposed UCGCA approach. The electric network examined includes 91 generators of various fuel types, each characterized by specific CI values referenced by Savelli et al. (2022), as listed in Table 4. The UC data is derived from the RTO Unit Commitment Test System, with further details provided in Federal Energy Regulatory Commission (2022). Bid prices for participants in the carbon market are sampled from normal distributions with standard deviations of $\$10/t\mathrm{CO}_2$ and means equal to $\$30.24/t\mathrm{CO}_2$. Negative prices are set to zero. The mean value is based on the California Cap-and-Trade Program's carbon allowance prices for Q3 2024 (California Air Resources Board, 2024). In the absence of specific allowance data, we allocate carbon allowances to each generator at 50% of their maximum emission potential, calculated as the maximum power output multiplied by the carbon intensity. On the other hand, carbon allowances for each gas load junction are set at 165% of the firm gas load under standard conditions (i.e., when $\eta_g = 1.65$) multiplied by the carbon intensity (CI) factor for natural gas, which is set at 55 $t\mathrm{CO}_2$/mmcf according to United States Environmental Protection Agency (2023). The selection of $\eta_g = 1.65$ represents an intermediate stress point on the gas network, positioned in the middle of the tested range ($\eta_g = 1.0, 1.1, \ldots, 2.3$). The cost associated with gas load shedding is set to \$130/mmBtu (Byeon & Van Hentenryck, 2020), while the cost for acquiring additional external carbon allowances is priced at $\$50/t\mathrm{CO}_2$. The risk aversion level, α_u, is set to 100%. The model runs were performed using C++/Gurobi 11.0.0 on an Apple M1, 3.2 GHz processor with 16 GB of RAM, with each run having a wall-time limit of 1 h. The analysis covers a single time-period (i.e., $T = 1$). The proposed UCGCA model and solution approach naturally generalizes to a multi-period unit commitment setting by introducing inter-temporal constraints, such as ramping limits, minimum up/down times, and start-up/shut-down decisions, see constraints in Problem (B.1). In that case, the first-level problem becomes a multi-period UC model with additional time-coupling constraints, while the decomposition framework remain unchanged. The computational complexity would increase due to inter-temporal coupling, but the underlying sequential multi-level structure and solution approach are directly applicable.

5.1. Effectiveness of the bid-validity constraints

The performance of the UCGCA (Fig. 3(a)) and BM (Fig. 3(b)) models is compared under varying system stress conditions to evaluate the impact of incorporating anticipated gas and carbon market outcomes into the UC problem. The electrical system stress is characterized by load increases of 20% and 60% (stress levels $\eta_e = \{1.0, 1.2, 1.6\}$, representing normal, slight, and high stress), while the gas network stress reflects load increases from 10% to 130% (stress levels $\eta_g = \{1.0, 1.1, \ldots, 2.3\}$). The financial losses in BM are quantified by calculating the sum of the difference between the violated GFPPs' marginal bids $\rho^*_{u,t}$ times the risk aversion level α_u in the electricity market and the realized costs of gas and carbon $(2p^*_{u,t}H^G_{u,2} + H^G_{u,1})\psi^*_{k,t} + (2y^*_{u,t} - 1)\kappa_u\pi^{e*}_t$, multiplied by the power produced $p^*_{u,t}$, i.e.,

Financial Losses

$$= \sum_{t\in[T]}\sum_{u\in\mathcal{U}_g} \max\left(0, (2p^*_{u,t}H^G_{u,2} + H^G_{u,1})\psi^*_{k,t} + (2y^*_{u,t} - 1)\kappa_u\pi^{e*}_t - \alpha_u\rho^*_{u,t}\right)p^*_{u,t}.$$

In contrast, UCGCA does not incur any financial loss due to the enforcement of the bid-validity constraints (11).

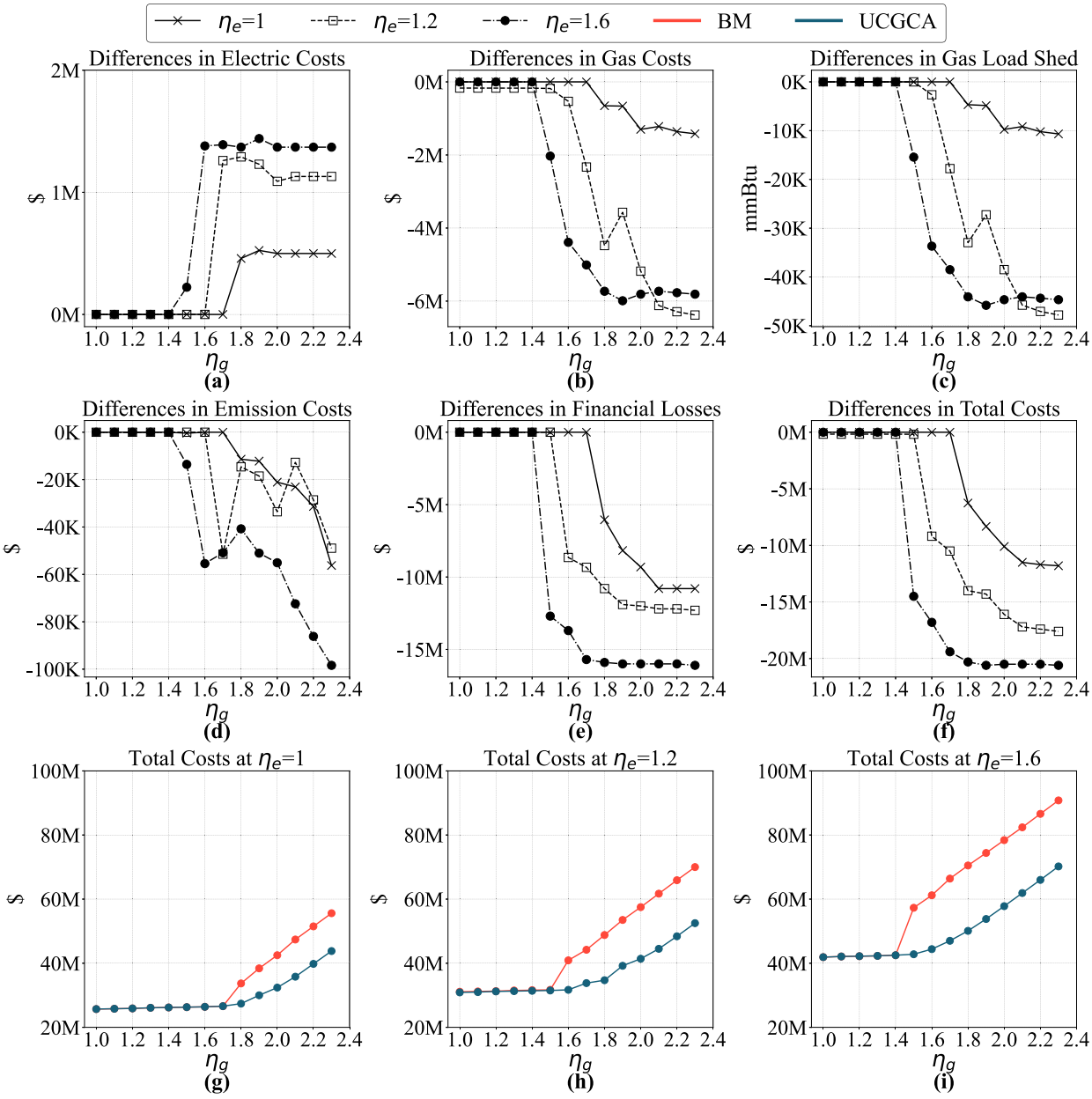

Fig. 6. Cost performance comparison between BM and UCGCA under varying electrical (η_e) and gas (η_g) stress levels, with the x-axis representing η_g. Positive/negative differences indicate higher/lower UCGCA values relative to BM. Panels show differences in: (a) electric costs, (b) gas costs, (c) gas load shed, (d) emission costs, (e) financial losses, where UCGCA avoids GFPP losses through bid-validity constraints, and (f) total costs, where BM total costs include electric, gas, carbon market costs, and financial losses, while UCGCA total costs comprise only market costs because UCGCA avoids financial losses. (g)–(i) Direct comparison of total costs between models at $\eta_e = \{1, 1.2, 1.6\}$.

Fig. 6 illustrates the differences in cost and gas load shed between the UCGCA and the BM. Fig. 7 extends this comparison to differences in gas prices between the two models across all instances. Note that the positive/negative differences indicate higher/lower values achieved in the UCGCA. Fig. 6 illustrates the differences in cost and gas load shed between the UCGCA and the BM. BM's total costs include electric, gas, carbon market costs, and financial losses from unprofitable GFPPs. In contrast, UCGCA only incurs electric, gas, and carbon market costs, as its bid-validity constraints prevent unprofitable operations and financial losses. For direct comparison, Fig. 6(g), (h) and (i) show the actual total costs of both models. Furthermore, Fig. 7 illustrates gas price differences between the models across all instances. We focus on two gas price regions, represented by black solid lines for Transco Zone 6 Non-NY and dashed lines for Transco Leidy Line Zone. These zonal prices are derived from average values of selected junctions (see Fig. 5). To capture system-wide effects, the blue line shows the mean price difference across *all* junctions, with the shaded blue area around the blue line indicates the 95% percentile interval, showing the range where most of the price differences occur. For direct comparison, Fig. 7(d), (e) and (f) show the actual zonal gas prices of both models.

Under normal electrical conditions ($\eta_e = 1$), the BM begins to suffer financial losses at a gas network stress level of $\eta_g = 1.8$, with financial losses escalating to \$10.8M as η_g reaches 2.3 (see Fig. 6(e)). These losses are primarily due to unprofitable GFPPs, which result from the lack of awareness of gas and carbon market clearing outcomes. The BM also experiences higher gas load shedding due to increased network

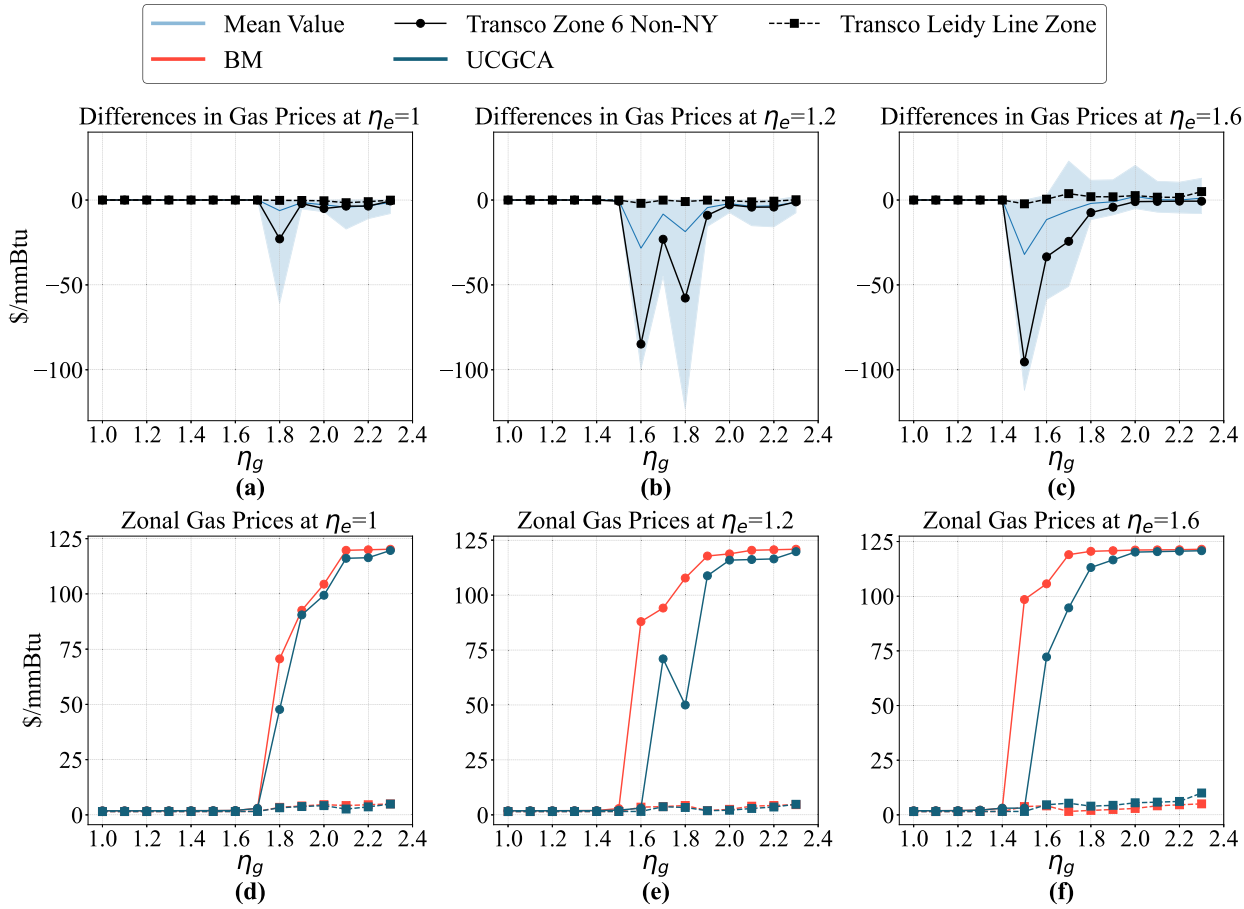

Fig. 7. Comparison of natural gas prices between BM and UCGCA under varying electrical (η_e) and gas stresses (η_g), with the x-axis representing η_g. Note that positive/negative differences indicate higher/lower values in UCGCA compared to BM. The shaded area around the blue line indicates the 95% percentile interval, showing the range where most of the junction prices differences fall. Plots (a)–(c) show the natural gas price differences between UCGCA and BM at electrical stress levels $\eta_e = \{1, 1.2, 1.6\}$, while plots (d)–(f) present the corresponding zonal gas prices for both models at the same η_e values.

congestion, leading to significant rises in gas costs and abrupt spikes in zonal natural gas prices in the Transco Zone 6 Non-NY area. In contrast, the UCGCA model shows a more resilient response to these conditions, as evidenced by the negative differences in zonal gas prices at Transco Zone 6 Non-NY and reduced gas load shed from $\eta_g = 1.8$, as shown in Fig. 7(a) and Fig. 6(c), respectively. Additionally, the reduction in carbon allowance deficits among gas and power market participants contributes to lower overall emission costs. Although the UCGCA has increased electric costs from de-committing unprofitable GFPPs with lower bids (see the positive difference in Fig. 6(a)), its total costs are significantly reduced (see the large negative difference in Fig. 6(f)). This cost efficiency is achieved through bid-validity constraints that prevent financial losses from invalid bids, mitigate the impacts of gas congestion, and significantly lower gas costs. Fig. 6(g) presents a direct cost comparison between the two models under normal electrical system conditions ($\eta_e = 1$). The results, with BM shown in red and UCGCA in blue, demonstrate that UCGCA consistently achieves cost reductions when the gas network experiences congestion.

As electrical stress increases to $\eta_e = 1.2$, the advantages of the UCGCA's operational adjustments become more apparent. For the BM, financial losses and gas price spikes occur earlier, starting at $\eta_g = 1.6$, with losses increasing to \$12.3M by $\eta_g = 2.3$. Conversely, the UCGCA maintains a lower cost profile even as gas stress intensifies, demonstrating effective management of network operations and costs, see the significant cost difference of \$17.6M at $\eta_g = 2.3$ in Fig. 6(f). This is due to the UCGCA's effective selection of better-committed generators, which reduces the severity of sudden price increases in the Transco Zone 6 Non-NY area, as evidenced by the maximum price difference of \$84.88/mmBtu at $\eta_g = 1.6$ in Fig. 7(b). Note that in Fig. 7(b) and (e), there is a drop in Transco Zone 6 Non-NY gas prices at $\eta_g = 1.8$, which results from the presence of optimality gaps for some hard instances.

The contrast between the outcomes of the BM and UCGCA becomes even more pronounced under the highest electrical stress level, $\eta_e = 1.6$. In this scenario, the UCGCA successfully avoids the significant financial losses from invalid bids that the BM begins to incur at $\eta_g = 1.5$. The UCGCA's overall costs are \$20.6M lower than those of the BM at $\eta_g = 2.3$, suggesting its superior ability to manage costs effectively. Fig. 6(i) further provides the actual values of total costs under BM and UCGCA at $\eta_e = 1.6$. Notably, at $\eta_g = 1.5$, the difference in zonal gas prices at Transco Zone 6 Non-NY between the BM and UCGCA reaches \$95.39/mmBtu (see Fig. 7(c) and (f)), indicating that the UCGCA significantly mitigates the impact of gas network congestion by committing alternative generators, thereby effectively reducing gas load shedding costs.

Table 5 presents the distribution of committed generators across different fuel types for nine representative instances. In the BM, the number of committed generators remains consistent regardless of variations in η_g. Conversely, the UCGCA model responds to a highly stressed gas network by altering its commitments based on the value of η_g. For

Table 5
Number of committed generators by fuel type when $\alpha_u = 100\%$. Columns represent committed generator counts by fuel types: Oil (O), Coal (C), Gas (G), Hydro (H), Refuse (R), Nuclear (N), and Others (E).

η_e	η_g	Model	O	C	G	H	R	N	E
1.0	1.0	UCGCA	3	4	9	11	0	10	3
	1.8	UCGCA	5	6	9	11	0	10	3
	2.3	UCGCA	5	6	8	11	1	11	3
	*	BM	3	4	9	11	0	10	3
1.2	1.0	UCGCA	5	5	11	11	0	10	3
	1.8	UCGCA	6	4	10	11	0	11	3
	2.3	UCGCA	7	6	8	11	0	12	3
	*	BM	5	5	11	11	0	10	3
1.6	1.0	UCGCA	7	6	12	11	0	12	3
	1.8	UCGCA	8	5	11	11	0	13	3
	2.3	UCGCA	8	5	10	11	0	13	3
	*	BM	7	6	12	11	0	12	3

example, at $\eta_e = 1.6$, the UCGCA model strategically de-commits two GFPPs and one coal-fired generator and replaced by one oil-type and one nuclear-type generator as η_g increases from 1.0 to 2.3.

In summary, the UCGCA model, by incorporating anticipated outcomes from sequentially cleared markets into the unit commitment process, not only prevents financial losses for GFPPs but also moderates gas prices, reduces network congestion, and significantly lowers overall system costs.

5.2. Effectiveness of the dedicated Benders decomposition method

The computational efficiency of the UCGCA model was assessed using two approaches: conventional approach using Gurobi to solve the SLP (2) without decomposition (denoted as C) and a dedicated Benders Decomposition (denoted as B) for solving SLP (2) introduced in Section 3.2. The dedicated Benders Decomposition is enhanced with two acceleration schemes: an in-out acceleration scheme with perturbation (Byeon & Van Hentenryck, 2022; Fischetti et al., 2017) and a normalization condition in the Benders subproblem (BSP) (Byeon & Van Hentenryck, 2022; Fischetti et al., 2010). We evaluate the performances of C and B across various scenarios, using combinations of gas network stress ($\eta_g = \{1, 1.1, \ldots, 2.3\}$), electrical system conditions ($\eta_e = \{1, 1.2, 1.6\}$), and risk-aversion levels ($\alpha_u = \{80\%, 100\%, 120\%\}$). This parametric combination yields 126 distinct test instances (14 gas stress levels × 3 electrical stress levels × 3 risk factors). We adopt the conservative tolerance in Gurobi to ensure numerical stability and solution reliability. For C, we let Gurobi parameters with `NumericFocus=3`, `FeasibilityTol=1e-9`, `OptimalityTol=1e-9`, `IntFeasTol=1e-9` and `TimeLimit=3600`. For B subproblems, we let Gurobi parameters with `NumericFocus=3`, `FeasibilityTol=1e-9`, `DualReductions=0`, `BarHomogeneous=1`, `BarQCPConvTol=1e-7`, `Aggregate=0` and `ScaleFlag=0`. For B master problems, we let Gurobi parameters with `FeasibilityTol=1e-9`, `OptimalityTol=1e-9` and `IntFeasTol=1e-9`. The detailed computational performance of C and B for varying parameters (η_e, η_g, and α_u) is presented in Tables C.14–C.16 of Appendix C in the supplementary material.

The 126 test instances are categorized according to their computational complexity. We classify an instance as *hard* if either C or B fails to achieve optimality within the time limit (3600 s), and as *easy* if both methods converge to optimal solutions within this limit. Fig. 8 compares computational performance between methods C and B for *hard* instances. In Fig. 8(a) and (b), the red dashed line represents equal performance between the two methods. Among 52 hard instances, C only outperforms B in only three cases $(\alpha_u, \eta_e, \eta_g) = \{(80\%, 1, 2.1), (80\%, 1, 2.2), (80\%, 1.2, 2)\}$, achieving slightly smaller optimality gaps as shown by the few points above the reference line in Fig. 8(b). However, B demonstrates superior performance in most hard cases, with the majority of points falling below the red dashed line, indicating both smaller optimality gaps and shorter computing times. C's performance deteriorates significantly for more challenging instances, either failing to find any incumbent solution (shown by 100% optimality gaps in Fig. 8(b)) or producing solutions with large optimality gaps.

Table 6 summarizes the computational performance of both methods across different instance types. In summary, B demonstrates superior computational performance compared to C across all test instances. On average, B demonstrates superior performance, reducing solution times by 32.23% (from 1549.85 s to 1050.55 s) and optimality gaps by 94.23% (from 16.80% to 0.97%). B's advantage is particularly pronounced in computationally challenging instances, reducing both solution times by 31.72% (from 3600 s to 2458.11 s) and optimality gaps by 94.20% (from 40.71% to 2.36%). Even for easy instances, B shows efficiency improvements, solving problems 43.73% faster than C (61.45 s versus 109.20 s). These results demonstrate that B not only provides high quality solutions but also achieves them more efficiently, making it particularly valuable for large-scale applications where computational performance is important.

5.3. Strong duality analysis and numerical performance with different scaling values

In this section, we examine how the scaling factor $\gamma = \{0.9, 0.99, 0.999, 0.9999, 0.99999\}$ affects the strong duality condition and numerical stability for each lower-level market-clearing problem under three extreme conditions $(\eta_e, \eta_g) = \{(1, 1), (1.2, 1), (1.6, 1)\}$. Let $(\boldsymbol{z}^*, \boldsymbol{x}^*, \boldsymbol{y}^*)$ denote the optimal solution of SLP (2). Theorem 2.1 proves that as γ approaches 1, constraints (2j) ensure strong duality for each sequential lower-level problem. We quantify the strong duality gaps (SDG) using the absolute percentage difference between primal and dual objectives relative to the primal objective for three lower-level problems:

$$SDG_e = \left| \frac{c_1^T \boldsymbol{x}_1^* - \boldsymbol{y}_1^{*T}(b_1 - D_1 \boldsymbol{z}^*)}{c_1^T \boldsymbol{x}_1^*} \right| \times 100\%$$

$$SDG_g = \left| \frac{c_2^T \boldsymbol{x}_2^* - \boldsymbol{y}_2^{*T}(b_2 - B_2 \boldsymbol{x}_1^* - D_2 \boldsymbol{z}^*)}{c_2^T \boldsymbol{x}_2^*} \right| \times 100\%$$

$$SDG_c = \left| \frac{c_3^T \boldsymbol{x}_3^* - \boldsymbol{y}_3^{*T}(b_3 - B_3 \boldsymbol{x}_2^* - D_3 \boldsymbol{z}^*)}{c_3^T \boldsymbol{x}_3^*} \right| \times 100\%$$

For these three extreme cases where GFPPs incur no financial losses, the results from BM (without scaling factors) should match those from UCGCA (with scaling factors). We measure this convergence using the total cost difference between UCGCA and BM solutions, where total costs comprise electricity, gas, and carbon market costs:

$$TC\_Diff = \left| \frac{TC^{UCGCA} - TC^{BM}}{TC^{BM}} \right| \times 100$$

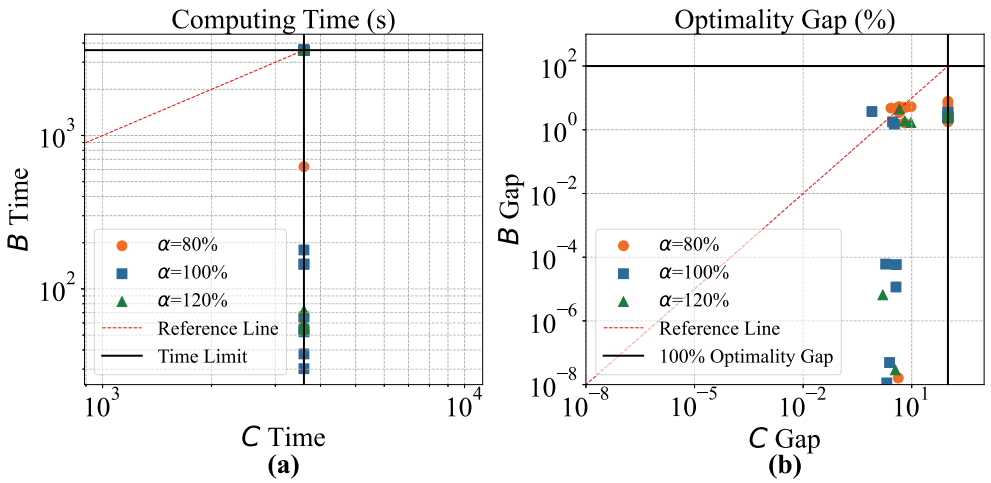

Fig. 8. Comparison of computational effectiveness between using Gurobi to solve the SLP (2) without decomposition (denoted as C) and a dedicated Benders Decomposition (denoted as B) for hard instances (i.e., when the computational time of C or B reaches the 1-hour limit). (a) Computational times of C and B. (b) Optimality gaps of C and B.

Table 6
Averaged computational performance comparison between C and B.

Instance type	C		B	
	Time (s)	Gap (%)	Time (s)	Gap (%)
All Instances	1549.85	16.80	1050.55	0.97
Easy Instances	109.20	0.00	61.45	0.00
Hard Instances	3600.00	40.71	2458.11	2.36

Fig. 9 demonstrates how the strong duality gaps and total cost differences vary with γ. The upper plots (a)–(c) show that as γ approaches 1, the total strong duality gaps decrease significantly from initial values of 130.60%, 191.03%, and 164.13% (at $\gamma = 0.9$) to less than 0.4% (at $\gamma = 0.9999$), aligning with Theorem 2.1. When γ is small ($\gamma = 0.9$), we observe both high SDG values and large TC_Diff values (reaching nearly 40% as shown in plots (d)–(f)), indicating that the single-level approximation fails to approach the original multi-level problem optimal solutions. The lower plots (d)–(f) show that TC_Diff consistently exhibits a U-shaped pattern across all three η_e cases, with values decreasing from nearly 40% at $\gamma = 0.9$ to a minimum of approximately 0.1% at $\gamma = 0.9999$. However, when γ increases to 0.99999, both performance metrics deteriorate. Notably, at $\eta_e = 1.6$, the total SDG increases to 5.89% (where SDG_c accounts for 5.86% of this total), and TC_Diff rises to approximately 0.4%, indicating numerical instability at this extreme value.

Our analysis identifies $\gamma = 0.9999$ as the optimal choice for solving UCGCA, where total SDG remains below 0.4% for all η_e conditions, and TC_Diff reaches its minimum value (approximately 0.1%) for all three cases. This value achieves the best balance between approximation accuracy and computational stability across $(\eta_e, \eta_g) = \{(1, 1), (1.2, 1), (1.6, 1)\}$.

5.4. Discussion

The single-level approximation quality depends on both the scaling factor value and the problem structure. Section 5.3 examines how different scaling factors influence the solution quality of each lower-level market-clearing problem relative to the original multi-level optimization problem using the UCGCA case study. Theoretically, as the scaling factor $\gamma \to 1$, the reformulated SLP problem satisfies the strong duality condition for all lower-level market-clearing problems and converges to the optimal solution of the original multi-level problem. We show that $\gamma = 0.9999$ as the optimal choice for solving UCGCA. In practice, the appropriate magnitude of γ depends on the number of lower-level problems n considered. For example, values around 0.99 have been effectively applied in tri-level optimization settings, see Byeon and Van Hentenryck (2020) and Mitridati et al. (2019). In practical implementation, as n increases, potential numerical scaling issues can be mitigated by appropriately normalizing the magnitudes of decision variables across levels.

Importantly, Theorem 3.3 is introduced to address the scalability issue arising when solving all lower-level problems connected through γ-scaling. By separating the BSP into primal-related and dual-related subproblems (BSP1 and BSP2), Theorem 3.3 enables each BSP to be further decomposed into n smaller problems that can be solved sequentially or in parallel according to the inherent structure of the lower-level interactions. The decomposition leverages the stage ordering encoded in the γ-based formulation to determine the direction of information flow across successive follower stages, enabling sequential or parallel computation. Importantly, γ itself does not enter the computation of the optimal solution as a multiplicative scaling factor. As a result, this approach improves numerical stability and scalability of the MIMLSF problem.

The proposed MIMLSF scales effectively due to its hierarchical decomposition. Although the approximated SLP remains a large-scale mixed-integer optimization problem, computational tractability can be achieved through the proposed Benders decomposition, which separates integer decisions in the master problem from multiple BSPs (see Theorem 3.1–3.3). As network detail increases, the number of variables and constraints naturally grows, raising computational effort. However, each subproblem incorporates only its own network constraints and can be solved independently or in parallel, significantly mitigating full cross-network coupling. Moreover, the numerical experiments are conducted on a highly complex multi-energy networks for the northeastern United States, combining an IEEE 36-bus electricity system with a multi-company gas network and carbon market interactions (see Fig. 4). The substantial computational gains observed in Section 5.2 under this high-dimensional and tightly coupled setting indicate that the

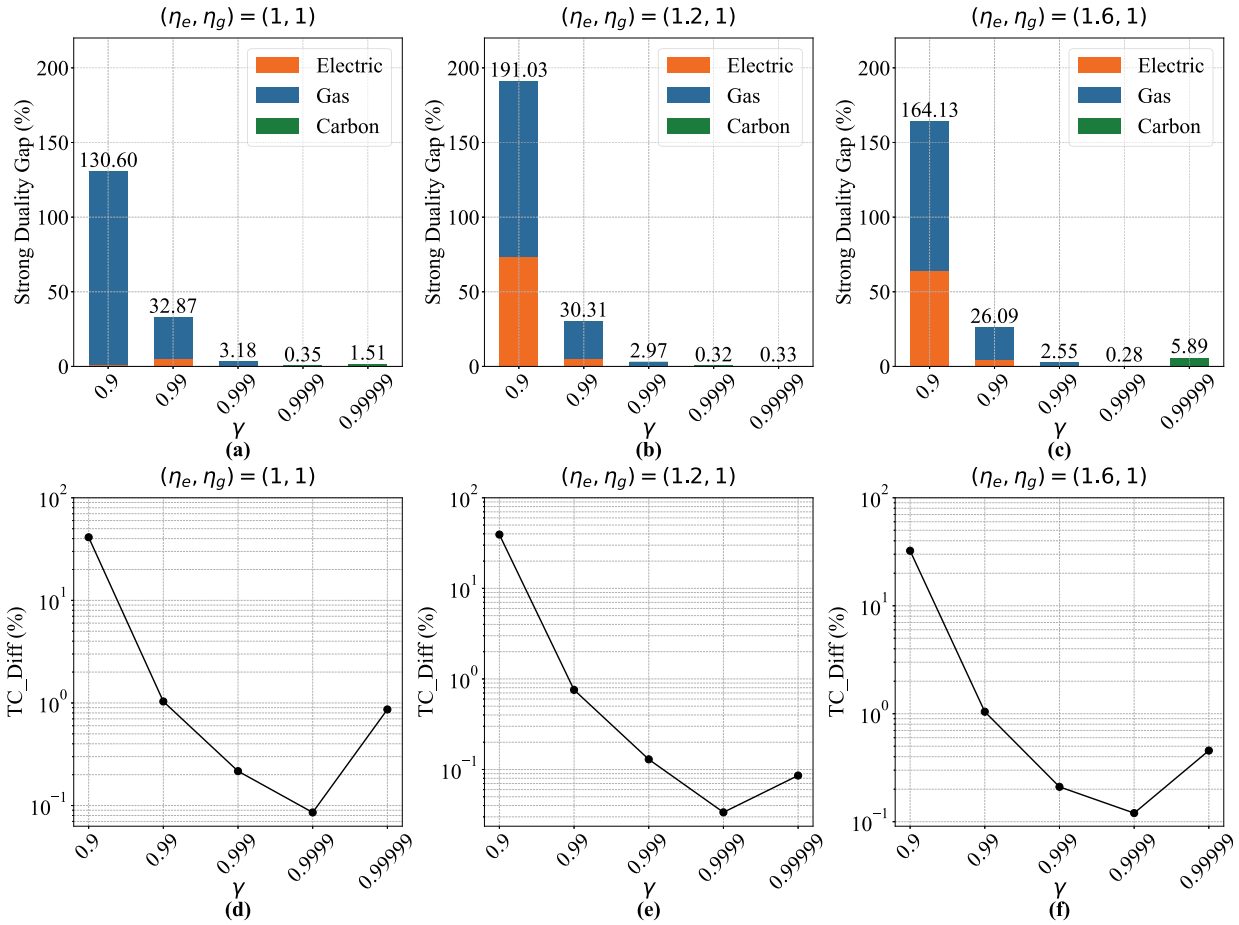

Fig. 9. Strong Duality Gap (SDG) and total cost differences under different scaling factors γ and η_e conditions: (a)–(c) show the SDG percentages for electricity, gas, and carbon market clearing problems, and (d)–(f) present the total cost differences between UCGCA and BM for three extreme instances $(\eta_e, \eta_g) = (1, 1), (1.2, 1), (1.6, 1)$.

improvements are driven by the structural properties of the proposed decomposition rather than by particular network simplifications. Consequently, for more complex multi-energy system models with stronger hierarchical coupling and higher dimensionality, similar computational advantages can reasonably be anticipated from the proposed method.

6. Management and policy implications

This section synthesizes the key insights from the UCGCA case study and the MIMLSF problem, highlighting their operational, strategic, and regulatory implications for system operators, GFPPs, and the coordinated integration of electricity, gas, and carbon markets.

6.1. Implications for system operators

The growing need for improved coordination between electricity and gas system operations has been recognized in both UK and US policy frameworks. In the UK, the Ofgem report Review of GB Energy System Operation identifies institutional and legal barriers to cross-system coordination, rooted in the separate regulatory frameworks governing the Electricity System Operator and the Gas System Operator (Ofgem, 2021). As part of recent reforms, NESO has been assigned responsibility for electricity and gas system planning (National Energy System Operator (NESO), 2025), as well as electricity system operation, while National Grid Gas retains responsibility for gas system operation (Ofgem, 2025). This fragmentation is cited as a key obstacle to whole-system decision-making and may prevent operators from delivering the most cost-effective and reliable solutions to meet net-zero targets. The report also highlights that greater coordination could be achieved through better information sharing and the development of a common language between the electricity and gas sectors. In the US, FERC Order 787 permits limited information exchange between electricity and natural gas transmission operators, acknowledging the operational risks introduced by growing interdependencies (Byeon & Van Hentenryck, 2020). These developments reflect increasing policy awareness that separate system operations can lead to inefficiencies, and that improved coordination and institutional alignment are needed to ensure reliable and efficient system outcomes.

Within this context, the UCGCA model developed in this paper offers a practical solution to support decision-making in interdependent yet decentralized energy systems. The model enables electricity system operators to account for fuel and emissions cost exposure when scheduling GFPPs. The UCGCA model incorporates bid-validity constraints that reflect anticipated gas and carbon prices, along with generator-specific risk preferences, supporting more informed, cost-effective, and lower-risk scheduling decisions. It enables conditional bidding, allowing GFPPs to avoid commitment when projected costs are too high, thereby reducing financial risk under volatile market conditions. As

system operators consider reforms to better manage market interdependencies, the UCGCA model offers a realistic and adaptable tool for improving cost-efficiency and reducing financial risk in multi-scale market environments.

6.2. Implications for gas-fired power plants

In electricity markets, GFPPs submit bids without knowing realized gas prices, as the gas market clears subsequently based on fuel demand induced by electricity dispatch (Byeon & Van Hentenryck, 2020). Repeated exposure to volatile input costs, without adequate hedging or cost recovery mechanisms, can lead to contractual defaults or market exit, with adverse implications for system adequacy, resource availability, and reliability. In extreme situations, GFPPs may be required to remain online for reliability purposes while incurring substantial financial losses. A prominent real-world example occurred during the 2014 polar vortex in the northeastern United States, when GFPPs were obligated to operate despite extreme fuel price spikes and no guarantee of cost recovery, resulting in substantial financial losses (PJM, 2014). These risks extend beyond rare events and reflect a broader structural issue in power system decision-making, where actions are taken without fully accounting for cascading effects across interdependent, multi-scale energy markets. As interdependencies among electricity, gas, and carbon markets continue to increase, there is a growing need for coordination mechanisms that explicitly capture these linkages to support more informed operational decisions and improve the risk management and financial viability of GFPPs. In response, the proposed UCGCA framework enables system operators to incorporate cross-market price signals into UC decisions and supports richer bidding profiles for GFPPs.

It is important to note that GFPPs can mitigate financial risks through various operational and contractual strategies, such as securing flexible gas supply arrangements, contracting storage access, improving gas and carbon price forecasting, and developing bilateral or hedging agreements. While these measures help manage risks, they do not address the more fundamental limitation that key operational decisions, such as UC, are made without full knowledge of the interactions among electricity, gas, and carbon markets. The UCGCA model complements these operational strategies by introducing system-level market awareness into the UC process. It acts as a form of "double insurance": even if GFPPs submit inaccurate bids due to imperfect forecasts, especially under high-demand or network-stressed conditions that amplify gas and carbon price volatility, system operators can apply the UCGCA model to make more informed commitment decisions. This helps prevent significant financial losses that would otherwise be incurred by GFPPs and enhances overall system efficiency.

6.3. Implications for cross-market coordination and decarbonization

The carbon market introduces environmental cost signals through emissions trading schemes or carbon taxes that directly influence operational dispatch decisions and long-term investment incentives (U.K. Government, 2016). The relative impact of gas and carbon markets on system outcomes depends on prevailing system conditions: gas market dynamics tend to exert stronger influence when fuel supply is constrained or demand rises, whereas carbon pricing becomes more binding as decarbonization targets tighten and emissions caps become more restrictive.

Effective integration of carbon markets with electricity and gas markets requires institutional and market mechanisms that internalize carbon price signals into operational decisions while remaining compatible with existing market structures. The proposed UCGCA framework provides such a mechanism by explicitly incorporating carbon allowance price signals into upper-level commitment. This ensures that carbon costs are reflected in generation scheduling and market-clearing outcomes while preserving institutional separation across markets. Such coordinated integration across electricity, gas, and carbon markets strengthens the role of carbon pricing as an effective decarbonization instrument, aligns operational incentives with clean energy objectives, and supports a coherent transition toward a low-carbon energy system.

7. Conclusion

Decision-making in multi-scale energy markets is becoming increasingly interdependent and sequentially coupled across temporal, spatial, and sectoral layers. A central challenge arises from the need to make forward decisions before the outcomes of subsequent, interdependent stages are realized. However, while interdependence across energy systems continues to grow, current market arrangements remain institutionally fragmented and temporally sequential, highlighting a gap in coordination mechanisms across multi-scale markets. This lack of coordination is evident in day-ahead unit commitment (UC) decisions, which are made before downstream market conditions in natural gas and carbon systems are realized. As a result, gas-fired power plants (GFPPs) committed in the day-ahead market can become exposed to volatile price risks from these markets. When downstream prices rise substantially, the committed GFPPs may become unprofitable to operate. Under normal conditions, such units might withdraw from the market; however, during extreme events such as the 2014 polar vortex, some generators were required to continue operating to maintain system reliability, resulting in substantial financial losses. These cases are not isolated events but reflect systemic shortcomings in current market designs, where decisions across coupled multi-scale markets remain poorly coordinated.

To address these challenges, we propose a Mixed-Integer Multi-Level problem with Sequential Followers (MIMLSF) for coordinating the hierarchical and sequential structure of coupled energy markets. By treating downstream markets as embedded lower-level decision problems, the MIMLSF formulation enables upper-level decisions to internalize and remain aware of the cross-market dependencies. This framework provides a general foundation for improving coordination, reliability, and economic efficiency across tightly linked systems.

To solve this computationally challenging MIMLSF problem, we asymptotically approximate the multi-level problem as a single-level problem. Specifically, we combine lexicographic optimization with a weighted-sum method through a scaling parameter γ. To enhance computational performance, we develop a dedicated Benders decomposition technique for the single-level problem. This technique first separates the complex Benders subproblem (BSP) into two tractable BSPs, and then further decomposes these two BSPs into n tractable problems where n is the number of sequential problems. This multi-level BSP separation reduces information exchange, eliminates the computational burden and avoids scalability issues of solving the complex BSP as an n-coupled problem.

We demonstrate our method's effectiveness on a four-level unit commitment with gas and carbon awareness (UCGCA) problem for the Northeastern United States, under conditions replicating the extreme market stresses of the 2014 polar vortex. This awareness is achieved by incorporating anticipated gas and carbon market outcomes into the UC problem, and those GFPPs who is expected to experience financial losses due to high gas and carbon costs will be de-committed. Compared with existing approaches in which system operators lack awareness of subsequent market outcomes, our approach achieves $20.6 million in total cost savings (a 22.7% reduction) while also avoiding financial losses for GFPPs exposed to volatile gas and carbon prices. Moreover, the dedicated Benders decomposition technique achieves a 94.23% reduction in optimality gaps and a 32.23% reduction in computing time compared to direct solution methods. We also demonstrate that $\gamma = 0.9999$ provides the best balance between approximated solution quality and numerical stability across three extreme cases in the UCGCA problem.

In summary, we propose a multi-level optimization problem with convex sequential followers to address multi-scale energy market coordination problems, and develop an asymptotic single-level reformulation with theoretical convergence guarantees to the original multi-level optimum, along with a dedicated Benders decomposition method. The proposed MIMLSF model and solution approach are validated through a real-world case study of coupled electricity, gas, and carbon markets in the Northeastern United States. Therefore, the contribution lies in a new hierarchical–sequential modeling framework, a theoretically grounded asymptotic reformulation with convergence guarantees, and a scalable decomposition architecture for coordinated decision-making in multi-scale markets. The proposed framework significantly improves computational efficiency while maintaining high solution quality, thereby enabling scalable and reliable decision-making in large-scale energy systems.

Promising directions for future work include extending the deterministic MIMLSF formulation to stochastic and chance-constrained frameworks that account for renewable uncertainty and inter-temporal decision lags in integrated energy systems. Future work will also extend the numerical study to a multi-period framework and evaluate scalability under inter-temporal constraints. In addition, the effectiveness of multi-level Benders separation techniques will be validated through real-world case studies.

CRediT authorship contribution statement

Yuxin Xia: Writing – original draft, Software, Methodology, Conceptualization. **Iacopo Savelli:** Writing – review & editing, Supervision, Conceptualization. **Thomas Morstyn:** Writing – review & editing, Supervision, Conceptualization.

Acknowledgment

Y.X. acknowledges support from an Engineering Studentship from the University of Edinburgh. I.S. acknowledges support from the European Union's Horizon Europe programme under the Marie Skłodowska-Curie grant agreement No. 101148367. The authors would like to thank Prof. G. Byeon and Prof. P. Van Hentenryck, whose work provided important inspiration for this study (Byeon & Van Hentenryck, 2020, 2022).

Appendix A. Proof of Theorem 2.1

Proof. Using strong duality conditions of the $n-1$ sequential lower-level problems (i.e., Eqs. (1f)–(1k)), the sequential problem is equivalent to the equation in Box II. In Problem (A.1), constraints (A.1b), (A.1d), (A.1h) and constraints (A.1e), (A.1i) represent the primal and dual constraints of the ith lower-level problem, respectively. Eqs. (A.1f) and (A.1j) enforce the strong duality to ensure optimality by posing a reversed weak duality constraint. Sequential followers framework requires that the $(i-1)^{th}$ lower-level problem does not anticipate the solutions of the ith (next) lower-level problem $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ since constraints and objective function of the $(i-1)^{th}$ lower-level problem are not dependent on the variables of the ith lower-level problem. Therefore, we can solve the n sequential lower-level problems in n steps:

1. solve the 1st lower-level problem and obtain the optimal primal and dual pairs $(\boldsymbol{x}_1^*, \boldsymbol{y}_1^*)$;
2. solve the 2nd lower-level problem with $\boldsymbol{x}_1$ fixed to the values $\boldsymbol{x}_1^*$ and obtain the optimal primal and dual pairs $(\boldsymbol{x}_2^*, \boldsymbol{y}_2^*)$;

...

n. solve the nth lower-level problem with $\boldsymbol{x}_{n-1}$ fixed to the values $\boldsymbol{x}_{n-1}^*$ and obtain the optimal primal and dual pairs $(\boldsymbol{x}_n^*, \boldsymbol{y}_n^*)$.

Accordingly, Problem (A.1) can be expressed as a Lexicographic optimization problem as follows (Ehrgott, 2005):

$$(\boldsymbol{x}, \boldsymbol{y}) \in \underset{\substack{\boldsymbol{x}_1\geq 0, \boldsymbol{y}_1\geq 0,\\ \cdots\\ \boldsymbol{x}_n\geq 0, \boldsymbol{y}_n\geq 0}}{\arg\min} < c_1^T\boldsymbol{x}_1, c_2^T\boldsymbol{x}_2, \ldots, c_n^T\boldsymbol{x}_n > \tag{A.2a}$$

$$\text{s.t.} \quad A_1\boldsymbol{x}_1 + D_1\boldsymbol{z} \geq b_1, \tag{A.2b}$$

$$A_i\boldsymbol{x}_i + B_i\boldsymbol{x}_{i-1} + D_i\boldsymbol{z} \geq b_i, \quad \forall i \in [n]^+, \tag{A.2c}$$

$$\boldsymbol{y}_i^T A_i \leq c_i^T, \quad \forall i \in [n]^+, \tag{A.2d}$$

$$\boldsymbol{y}_i^T(b_i - D_i\boldsymbol{z} - B_i\boldsymbol{x}_{i-1}) \geq c_i^T\boldsymbol{x}_i, \quad \forall i \in [n]^+. \tag{A.2e}$$

For given upper-level binary decisions $\bar{\boldsymbol{z}}$, any optimal solution $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ of Problem (A.2) satisfies the following conditions:

$$(\boldsymbol{x}_1^*, \boldsymbol{y}_1^*) \in \underset{\boldsymbol{x}_1\geq 0, \boldsymbol{y}_1\geq 0}{\arg\min} \quad c_1^T\boldsymbol{x}_1 \tag{A.3a}$$

$$\text{s.t.} \quad A_1\boldsymbol{x}_1 \geq b_1 - D_1\bar{\boldsymbol{z}},$$

$$A_2\boldsymbol{x}_2 + B_2\boldsymbol{x}_1 \geq b_2 - D_2\bar{\boldsymbol{z}}.$$

$$\left[(\boldsymbol{x}_i^*, \boldsymbol{y}_i^*) \in \underset{\boldsymbol{x}_i\geq 0, \boldsymbol{y}_i\geq 0}{\arg\min} \quad c_i^T\boldsymbol{x}_i \right. \tag{A.3b}$$

$$\text{s.t.} \quad A_i\boldsymbol{x}_i \geq b_i - D_i\bar{\boldsymbol{z}} - B_i\boldsymbol{x}_{i-1}^*,$$

$$\boldsymbol{y}_i^T A_i \leq c_i^T,$$

$$\left. \boldsymbol{y}_i^T(b_i - D_i\bar{\boldsymbol{z}} - B_i\boldsymbol{x}_{i-1}^*) \geq c_i^T\boldsymbol{x}_i\right] \quad \forall i \in [n]^+,$$

Observe from Problem (A.3b), any feasible solution $(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{y}}_i)$ of the ith lower-level problem is optimal, $\forall i \in [n]^+$. That is because, by strong duality forced in the third constraints of (A.3b), the pair $(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{y}}_i)$ satisfies for $\forall i \in [n]^+$:

$$\hat{\boldsymbol{x}}_i \in \underset{\boldsymbol{x}_i\geq 0}{\arg\min} \quad c_i^T\boldsymbol{x}_i \tag{A.4a}$$

$$\text{s.t.} \quad A_i\boldsymbol{x}_i \geq b_i - D_i\bar{\boldsymbol{z}} - B_i\boldsymbol{x}_{i-1}^*.$$

$$\hat{\boldsymbol{y}}_i \in \underset{\boldsymbol{y}_i\geq 0}{\arg\max} \quad \boldsymbol{y}_i^T(b_i - D_i\bar{\boldsymbol{z}} - B_i\boldsymbol{x}_{i-1}^*) \tag{A.4b}$$

$$\text{s.t.} \quad \boldsymbol{y}_i^T A_i \leq c_i.$$

Using the lexicographic optimization combined with the weight-sum method and the optimality conditions of Problem (A.2), given in Problems (A.3a) and (A.4), we can approximate Problem (A.2) with $\gamma \in (0,1)$ as Problem (A.5). The scaling factor applied for the ith lower-level/sequential problem is denoted by γ_i, with $\gamma_1 = \gamma$, $\gamma_2 = \gamma(1-\gamma)$, and continuing as $\gamma_n = (1-\gamma)^{n-1}$:

$$\min_{\substack{\boldsymbol{x}_1\geq 0, \boldsymbol{y}_1\geq 0, \ldots,\\ \boldsymbol{x}_n\geq 0, \boldsymbol{y}_n\geq 0}} \quad \sum_{i=1}^{n} \gamma_i c_i^T\boldsymbol{x}_i \tag{A.5a}$$

$$\text{s.t.} \quad A_1\boldsymbol{x}_1 + D_1\boldsymbol{z} \geq b_1, \quad [\boldsymbol{y}_1 \geq 0] \tag{A.5b}$$

$$A_i\boldsymbol{x}_i + B_i\boldsymbol{x}_{i-1} + D_i\boldsymbol{z} \geq b_i, \quad \forall i \in [n]^+. \quad [\boldsymbol{y}_i \geq 0] \tag{A.5c}$$

where $\boldsymbol{y}_i$ is obtained by the dual solution associated with the ith primal constraint of Problem (A.5). Note that $\boldsymbol{y}_i$ is scaled by γ_i as the objective function of ith lower-level problem is scaled by γ_i.

As a result, the multi-level problem can be approximated by Problem (1) for some $\gamma \in (0,1)$ by taking the dual constraints and reversed weak duality conditions of Problem (A.5). It remains to show that the single-level problem (2) is indeed an asymptotic approximation of the multi-level problem (1).

To this end, we introduce n auxiliary variables for dual variables $\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_n$ where $\tilde{\boldsymbol{y}}_1 = \frac{\boldsymbol{y}_1}{\gamma}, \tilde{\boldsymbol{y}}_2 = \frac{\boldsymbol{y}_2}{\gamma(1-\gamma)}, \ldots, \tilde{\boldsymbol{y}}_n = \frac{\boldsymbol{y}_n}{(1-\gamma)^{n-1}}$ (i.e., $\tilde{\boldsymbol{y}}_i = \frac{\boldsymbol{y}_i}{\gamma_i}$), then Problem (2) becomes

$$\min_{\substack{\boldsymbol{z}\in\{0,1\}^m,\\ \boldsymbol{x}_1\geq 0, \boldsymbol{y}_1\geq 0,\\ \cdots\\ \boldsymbol{x}_n\geq 0, \boldsymbol{y}_n\geq 0}} \quad \gamma h^T\boldsymbol{z} + \gamma c_{x1}^T\boldsymbol{x}_1 + \cdots + \gamma c_{xn}^T\boldsymbol{x}_n \tag{A.6a}$$

$$\text{s.t.} \quad \boldsymbol{z} \in \mathcal{Z}, \boldsymbol{z} \in \{0,1\}^m, \tag{A.6b}$$

$$H_{xi}\boldsymbol{z} + G_{xi}\boldsymbol{x}_i \geq b_{xi}, \quad \forall i \in [n], \tag{A.6c}$$

$$(\boldsymbol{x}_1, \boldsymbol{y}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_n) \in \arg\min c_1^T \boldsymbol{x}_1 \tag{A.1a}$$
$$\text{s.t.} \quad A_1\boldsymbol{x}_1 + D_1\boldsymbol{z} \geq b_1, \tag{A.1b}$$
$$(\boldsymbol{x}_2, \boldsymbol{y}_2, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_n) \in \arg\min c_2^T \boldsymbol{x}_2 \tag{A.1c}$$
$$\text{s.t.} \quad A_2\boldsymbol{x}_2 + B_2\boldsymbol{x}_1 + D_2\boldsymbol{z} \geq b_2, \tag{A.1d}$$
$$\boldsymbol{y}_2^T A_2 \leq c_2^T, \tag{A.1e}$$
$$\boldsymbol{y}_2^T(b_2 - D_2\boldsymbol{z} - B_2\boldsymbol{x}_1) \geq c_2^T \boldsymbol{x}_n, \tag{A.1f}$$
$$\cdots\cdots$$
$$(\boldsymbol{x}_n, \boldsymbol{y}_n) \in \arg\min c_n^T \boldsymbol{x}_n \tag{A.1g}$$
$$\text{s.t.} \quad A_n\boldsymbol{x}_n + B_n\boldsymbol{x}_{n-1} + D_n\boldsymbol{z} \geq b_n, \tag{A.1h}$$
$$\boldsymbol{y}_n^T A_n \leq c_n^T, \tag{A.1i}$$
$$\boldsymbol{y}_n^T(b_n - D_n\boldsymbol{z} - B_n\boldsymbol{x}_{n-1}) \geq c_n^T \boldsymbol{x}_n. \tag{A.1j}$$

Box II.

$$H_{yi}\boldsymbol{z} + G_{yi}\tilde{\boldsymbol{y}}_i \geq b_{yi}, \quad \forall i \in [n], \tag{A.6d}$$
$$R_y\boldsymbol{z} + \sum_{i=1}^{n} S_{yi}\tilde{\boldsymbol{y}}_i \geq q_y, \tag{A.6e}$$
$$A_1\boldsymbol{x}_1 + D_1\boldsymbol{z} \geq b_1, \tag{A.6f}$$
$$A_2\boldsymbol{x}_2 + B_2\boldsymbol{x}_1 + D_2\boldsymbol{z} \geq b_2, \tag{A.6g}$$
$$A_3\boldsymbol{x}_3 + B_3\boldsymbol{x}_2 + D_3\boldsymbol{z} \geq b_3, \tag{A.6h}$$
$$\cdots$$
$$A_n\boldsymbol{x}_n + B_n\boldsymbol{x}_{n-1} + D_n\boldsymbol{z} \geq b_n, \tag{A.6i}$$
$$\gamma\tilde{\boldsymbol{y}}_1^T A_1 + \gamma(1-\gamma)\tilde{\boldsymbol{y}}_2^T B_2 \leq \gamma c_1^T, \tag{A.6j}$$
$$\gamma(1-\gamma)\tilde{\boldsymbol{y}}_2^T A_2 + \gamma(1-\gamma)^2\tilde{\boldsymbol{y}}_3^T B_3 \leq \gamma(1-\gamma)c_2^T, \tag{A.6k}$$
$$\gamma(1-\gamma)^2\tilde{\boldsymbol{y}}_3^T A_3 + \gamma(1-\gamma)^3\tilde{\boldsymbol{y}}_4^T B_4 \leq \gamma(1-\gamma)^2 c_3^T, \tag{A.6l}$$
$$\cdots$$
$$(1-\gamma)^{n-1}\tilde{\boldsymbol{y}}_n^T A_n \leq (1-\gamma)^{n-1} c_n^T, \tag{A.6m}$$
$$\gamma\tilde{\boldsymbol{y}}_1^T(b_1 - D_1\boldsymbol{z}) + \cdots + (1-\gamma)^{n-1}\tilde{\boldsymbol{y}}_n^T(b_n - D_n\boldsymbol{z}) \geq \gamma c_1^T \boldsymbol{x}_1 + \gamma(1-\gamma)c_2^T\boldsymbol{x}_2 + \cdots + (1-\gamma)^{n-1}c_n^T\boldsymbol{x}_n. \tag{A.6n}$$

Given that the upper level variables $\boldsymbol{z}$ is fixed to $\hat{\boldsymbol{z}}$, we denote $P(\hat{\boldsymbol{z}})$ and $\tilde{P}(\hat{\boldsymbol{z}})$ as the objective value of the Lexicographic optimization problem (A.2) and the modified single-level problem (A.6), respectively. Let the set $(\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{y}}_2, \ldots, \hat{\boldsymbol{x}}_n, \hat{\boldsymbol{y}}_n)$ be the optimal solution of $\tilde{P}(\hat{\boldsymbol{z}})$.

For the 1st lower-level problem, as γ approaches to 1, Eqs. (A.6j) and (A.6n) becomes:

$$\tilde{\boldsymbol{y}}_1^T A_1 \leq c_1^T, \tag{A.7a}$$
$$\tilde{\boldsymbol{y}}_1^T(b_1 - D_1\hat{\boldsymbol{z}}) \geq c_1^T\boldsymbol{x}_1. \tag{A.7b}$$

which implies that given fixed upper-level decision $\hat{\boldsymbol{z}}$, $(\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{y}}_1)$ approximate the optimal and dual solution of Problem (A.3a). This is because Eq. (A.6f) poses the primal feasibility for problem (A.3a), and when $\gamma \to 1$, $\hat{\boldsymbol{y}}_1$, becomes feasible to the dual of Problem (A.3a) as ensured by constraints (A.7a). Eq. (A.7b) ensures the strong duality condition. Therefore, as $\gamma \to 1$, the pair $(\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{y}}_1)$ is a feasible solution to $P(\hat{\boldsymbol{z}})$ and has the same optimal objective value.

For the 2nd lower-level problem, as γ approaches to 1, (A.6k) becomes

$$\tilde{\boldsymbol{y}}_2^T A_2 \leq c_2^T. \tag{A.8a}$$

Similarly, the primal feasibility is guaranteed by constraint (A.6g). To prove the strong duality condition, we combine Eq. (A.6j)×$\boldsymbol{x}_1$ and Eq. (A.6n) and we have:

$$c_1^T\boldsymbol{x}_1 \geq \tilde{\boldsymbol{y}}_1^T A_1\boldsymbol{x}_1 + (1-\gamma)\tilde{\boldsymbol{y}}_2^T B_2\boldsymbol{x}_1, \tag{A.9a}$$
$$c_1^T\hat{\boldsymbol{x}}_1 \leq \tilde{\boldsymbol{y}}_1^T(b_1 - D_1\hat{\boldsymbol{z}}) + (1-\gamma)\tilde{\boldsymbol{y}}_2^T(b_2 - D_2\hat{\boldsymbol{z}}) + \cdots + \frac{(1-\gamma)^{n-1}}{\gamma}\tilde{\boldsymbol{y}}_n^T(b_n - D_n\hat{\boldsymbol{z}}) - \left((1-\gamma)c_2^T\boldsymbol{x}_2 + (1-\gamma)^2 c_3^T\boldsymbol{x}_3 + \cdots + \frac{(1-\gamma)^{n-1}}{\gamma}c_n^T\boldsymbol{x}_n\right). \tag{A.9b}$$

Therefore, we have

$$\tilde{\boldsymbol{y}}_1^T(A_1\boldsymbol{x}_1 - b_1 + D_1\hat{\boldsymbol{z}}) \leq (1-\gamma)\Big(\tilde{\boldsymbol{y}}_2^T(b_2 - D_2\hat{\boldsymbol{z}} - B_2\boldsymbol{x}_1) - c_2^T\boldsymbol{x}_2 + (1-\gamma)(\tilde{\boldsymbol{y}}_3^T(b_3 - D_3\boldsymbol{z}) - c_3^T\boldsymbol{x}_3) + \cdots + \frac{(1-\gamma)^{n-2}}{\gamma}(\tilde{\boldsymbol{y}}_n^T(b_n - D_n\boldsymbol{z}) - c_n^T\boldsymbol{x}_n)\Big). \tag{A.10a}$$

The left-hand-side term is non-negative according to the primal Eq. (A.6f), which means that the right-hand side term of (A.10a) is also non-negative. As $\gamma \to 1$, the scaling terms $(1-\gamma), \ldots, \frac{(1-\gamma)^{n-2}}{\gamma}$ approach to 0, therefore we have:

$$\tilde{\boldsymbol{y}}_2^T(b_2 - D_2\hat{\boldsymbol{z}} - B_2\boldsymbol{x}_1) \geq c_2^T\boldsymbol{x}_2, \tag{A.11a}$$

which implies that $\hat{\boldsymbol{x}}_2$ and $\hat{\boldsymbol{y}}_2$ approximate the optimal primal and dual solutions of Problem (A.3b) with fixed $\hat{\boldsymbol{z}}$.

Similarly, for the 3rd lower-level problem, as γ approaches to 1, Eq. (A.6l) ensures the dual feasibility condition as follows:

$$\tilde{\boldsymbol{y}}_3^T A_3 \leq c_3^T. \tag{A.12a}$$

To prove the strong duality condition, we combine Eq. (A.6l)×$\boldsymbol{x}_2$ and the non-negativity right-hand side term of Eq. (A.10a) as follows,

$$c_2^T\boldsymbol{x}_2 \geq \tilde{\boldsymbol{y}}_2^T A_2\boldsymbol{x}_2 + (1-\gamma)\tilde{\boldsymbol{y}}_3^T B_3\boldsymbol{x}_2, \tag{A.13a}$$
$$c_2^T\boldsymbol{x}_2 \leq \tilde{\boldsymbol{y}}_2^T(b_2 - D_2\hat{\boldsymbol{z}} - B_2\boldsymbol{x}_1) + (1-\gamma)\Big((\tilde{\boldsymbol{y}}_3^T(b_3 - D_3\boldsymbol{z}) - c_3^T\boldsymbol{x}_3) + \cdots + \frac{(1-\gamma)^{n-3}}{\gamma}(\tilde{\boldsymbol{y}}_n^T(b_n - D_n\boldsymbol{z}) - c_n^T\boldsymbol{x}_n)\Big). \tag{A.13b}$$

Therefore, we have

$$\hat{\boldsymbol{y}}_2^T(A_2\hat{\boldsymbol{x}}_2 - b_2 + D_2\hat{\boldsymbol{z}} + B_2\hat{\boldsymbol{x}}_1) \leq (1-\gamma)\Big(\hat{\boldsymbol{y}}_3^T(b_3 - D_3\boldsymbol{z} - B_3\hat{\boldsymbol{x}}_2) - c_3^T\hat{\boldsymbol{x}}_3$$

$$+\cdots+\frac{(1-\gamma)^{n-3}}{\gamma}(\hat{\boldsymbol{y}}_n^T(b_n-D_n\boldsymbol{z})-c_n^T\hat{\boldsymbol{x}}_n)\Big). \tag{A.14a}$$

The left-hand side term of Eq. (A.14a) is non-negative according to the primal constraint Eq. (A.6g), which means that the right-hand side term is also non-negative. As γ approaches to 1, we have:

$$\hat{\boldsymbol{y}}_3^T(b_3-D_3\hat{\boldsymbol{z}}-B_3\hat{\boldsymbol{x}}_2)\geq c_3^T\hat{\boldsymbol{x}}_3, \tag{A.15a}$$

which implies that $\hat{\boldsymbol{x}}_3$ and $\hat{\boldsymbol{y}}_3$ approximate the optimal primal and dual solutions of Problem (A.3b) with fixed $\hat{\boldsymbol{z}}$.

The proof can be similarly applied to the remaining lower-level problems. In addition, the single-level problem (A.6) contains bilinear terms in constraints (A.6n) $\boldsymbol{y}_i^T D_i\boldsymbol{z}$ where $i\in[n]$. The exact McCormick relaxation technique can be applied to linearize the bilinear terms $\boldsymbol{y}_i^T D_i\boldsymbol{z}$ with auxiliary variables $\boldsymbol{s}_{yi}^T\mathbf{1}=\boldsymbol{y}_i^T D_i\boldsymbol{z}$. As such, additional linearization constraints $Y_{yi}\boldsymbol{y}_i+K_{yi}\boldsymbol{s}_{yi}\geq\gamma_i(e_{yi}+E_{yi}\boldsymbol{z})$, $\forall i\in[n]$ with proper dimensions of matrices Y_y, K_y, E_y and vectors e_y are introduced in constraint (2k). The reason to incorporate γ_i in the right-hand side is that the upper-bound of variable $\boldsymbol{y}_i$ will also be scaled by γ_i since the ith lower-level objective function is scaled by γ_i and consequently the range of the dual variables $\boldsymbol{y}_i$ will also be scaled by γ_i.

In summary, $\hat{\boldsymbol{x}}_1$ is an approximate solution of $P(\hat{\boldsymbol{z}})$ that becomes increasingly close to the optimal solution of Problem $P(\hat{\boldsymbol{z}})$ as $\gamma\to1$, and the pair $(\hat{\boldsymbol{x}}_i,\hat{\boldsymbol{y}}_i)$ with $i\in[n]^+$ is the exact response of the follower i with respect to $\hat{\boldsymbol{x}}_{i-1}$ for any $\gamma\in(0,1)$. Therefore, the approximation may sacrifice the leader's optimality when γ is not large enough, but it always gives a feasible solution. The proposed single-level reformulation (2) can represent all lower-level optimality conditions using a single linear constraint, which is not achievable with conventional KKT- or strong-duality-based approaches. The resulting single-level formulation further enables efficient decomposition, as detailed in Section 3. □

Appendix B. Detailed formulation of the UCGCA model

In what follows, the electricity transmission grid is represented by an undirected graph $\mathcal{G}^e=(\mathcal{N},\mathcal{E})$ and the natural gas transmission system is by a directed graph $\mathcal{G}^g=(\mathcal{V},\mathcal{A})$. The letter $\mathcal{T}$ denotes the set of time periods $\{0,1,\ldots,T\}$. $[a,b]_{\mathbb{Z}}$ denotes the set of integers in interval $[a,b]$.

B.1. The upper-level unit commitment problem

The problem (B.1) is a classic UC problem with additional calculation of marginal bids for GFPPs and the bid validity constraints. Tables B.7 and B.8 summarize the parameters and variables of the UC problem.

$$\min\quad\sum_{t\in[T]}\sum_{u\in\mathcal{U}}\left(c_u o_{u,t}+r_{u,t}+\sum_{b\in B_u}\beta_{u,b}s_{u,b,t}\right) \tag{B.1a}$$

s.t.

$$r_{u,t}\geq C_{u,h}(o_{u,t}-\sum_{n\in[h]}o_{u,t-n}),\quad\forall h\in\Psi_u,u\in\mathcal{U},t\in[T], \tag{B.1b}$$

$$r_{u,t}\geq0,\quad\forall u\in\mathcal{U},t\in[T], \tag{B.1c}$$

$$o_{u,t}=\bar{o}_{u,0},\quad\forall u\in\mathcal{U},t\in[0,\bar{\tau}_{u,0}+\underline{\tau}_{u,0}]_{\mathbb{Z}}, \tag{B.1d}$$

$$\sum_{t'\in[t-\bar{\tau}_u+1,t]_{\mathbb{Z}}}v^+_{u,t'}\leq o_{u,t},\quad\forall u\in\mathcal{U},t\in[\max\{\bar{\tau}_u,\bar{\tau}_{u,0}+1\},T]_{\mathbb{Z}}, \tag{B.1e}$$

$$\sum_{t'\in[t-\underline{\tau}_u+1,t]_{\mathbb{Z}}}v^+_{u,t'}\leq1-o_{u,t-\underline{\tau}_u},\quad\forall u\in\mathcal{U},t\in[\max\{\underline{\tau}_u,\underline{\tau}_{u,0}+1\},T]_{\mathbb{Z}}, \tag{B.1f}$$

$$v^+_{u,t}-v^-_{u,t}=o_{u,t}-o_{u,t-1},\quad\forall u\in\mathcal{U},t\in[T], \tag{B.1g}$$

$$w_{u,b,t}\leq o_{u,t},\quad\forall b\in\mathcal{B}_u,u\in\mathcal{U}^g,t\in[T], \tag{B.1h}$$

$$0\leq s^e_{u,b,t}\leq\bar{s}_{u,b}w_{u,b,t},\quad\forall b\in\mathcal{B}_u,u\in\mathcal{U}^g,t\in[T], \tag{B.1i}$$

$$\bar{s}_{u,b}w_{u,b+1,t}\leq s^e_{u,b,t},\quad\forall b\in[1,\mathcal{B}_u-1]_{\mathbb{Z}},u\in\mathcal{U}^g,t\in[T], \tag{B.1j}$$

$$\rho_{u,t}=\sum_{b\in[B_u-1]}\beta_{u,b}(w_{u,b,t}-w_{u,b+1,t})+\beta_{u,B_u}w_{u,B_u,t},\quad\forall u\in\mathcal{U}^g,t\in[T], \tag{B.1k}$$

$$\sum_{b\in\mathcal{B}_u}\kappa_u s^e_{u,b,t}-E^0_{u,t}\geq-\overline{M}^e_u(1-y_{u,t}),\quad\forall u\in\mathcal{U}^g,t\in[T], \tag{B.1l}$$

$$\sum_{b\in\mathcal{B}_u}\kappa_u s^e_{u,b,t}-E^0_{u,t}\leq\overline{M}^e_u y_{u,t},\quad\forall u\in\mathcal{U}^g,t\in[T], \tag{B.1m}$$

$$The\ bid-validity\ constraint\ (11), \tag{B.1n}$$

$$\psi_{k,t}=\frac{1}{|\mathcal{J}_k|}\sum_{j\in\mathcal{J}_k}\pi^g_{j,t},\quad\forall k\in\mathcal{K},t\in[T], \tag{B.1o}$$

$$v^+_{u,t},v^-_{u,t},o_{u,t}\in\{0,1\},\quad\forall u\in\mathcal{U},t\in[T], \tag{B.1p}$$

$$w_{u,b,t}\in\{0,1\},\quad\forall b\in\mathcal{B}_u,u\in\mathcal{U}^g,t\in[T]. \tag{B.1q}$$

$$y_{u,t}\in\{0,1\},\quad\forall u\in\mathcal{U}^g,t\in[T]. \tag{B.1r}$$

The objective of the UC is to minimize the total cost of operating the power system over T. This total cost includes the start-up cost $r_{u,t}$, no-load cost $c_u o_{u,t}$ and the costs of the selected supply bids of each power generating unit $\sum_{b\in B_u}\beta_{u,b}s_{u,b,t}$. The non-negative start-up cost $r_{u,t}$ is calculated based on the length of time generator u has been offline. The expression $o_{u,t}-\sum_{n=1}^{h}o_{u,t-n}$ equals one when generator u reactivates after being offline for h time periods, as shown in (B.1b). Eq. (B.1d) suggests the initial operational status for each unit u. Constraints (B.1e) and (B.1f) provide formulation for the minimum up- and down-time where $\underline{\tau}_u$ and $\bar{\tau}_u$ are the minimum uptime and downtime of unit u. In addition, the logical constraint between the binary variables for the on–off $o_{u,t}$, start-up $v^+_{u,t}$, and shut-down $v^-_{u,t}$ statuses of generator u is stated in Eq. (B.1g).

To include the bid-validity constraints for GFPPs, the formulation for calculating the marginal bid $\rho_{u,t}$ is necessary. To this end, we introduce the binary variables $w_{u,b,t}$ to indicate whether a bid $b\in\mathcal{B}_u$ for a generating unit u is selected during time period t. Eq. (B.1h) states that the bid of a generator u can be selected only when it is committed. Constraint (B.1i) ensure that when the power generation $s^e_{u,b,t}$ is non-negative, the bid b is used (i.e., $w_{u,b,t}=1$) and constraint (B.1j) further states that the $(b+1)^{th}$ bid is selected *iif* the bth bid is fully used. As such, the marginal/maximum bid $\rho_{u,t}$ of GFPP u at time period t is calculated by the constraints involving $w_{u,b,t}$ as shown in Eq. (B.1k). Constraints (B.1l) and (B.1m) suggests that the GFPP u has deficit carbon allowances when $y_{u,t}=1$, and 0 otherwise. Furthermore, the bid-validity constraints (B.1n) is introduced to ensure that only profitable GFPPs will be committed. Eq. (B.1o) calculates the zonal gas prices $\psi_{k,t}$ at zone k and time t. Here, $|\mathcal{J}_k|$ denotes the number of elements in the pricing point set $\mathcal{J}_k$, and $\pi^g_{j,t}$ represents the gas price at node j within zone k at time t.

B.2. The second-level economic dispatch (electricity market clearing) problem

$$\min\quad\sum_{t\in[T]}\sum_{u\in\mathcal{U}}\sum_{b\in B_u}\beta_{u,b}s_{b,t} \tag{B.2a}$$

s.t.

$$\sum_{u\in\mathcal{U}(i)}p_{u,t}-d^e_{i,t}=\sum_{l\in\mathcal{E}:l_t=i}f_{l,t}-\sum_{l\in\mathcal{E}:l_h=i}f_{l,t},\quad\forall i\in\mathcal{N},t\in[T], \tag{B.2b}$$

$$p_{u,t}=\sum_{b\in\mathcal{B}_u}s^e_{u,b,t},\quad\forall u\in\mathcal{U},t\in[T], \tag{B.2c}$$

$$0\leq s^e_{u,b,t}\leq\bar{s}_b,\quad\forall b\in\mathcal{B}_u,u\in\mathcal{U},t\in[T], \tag{B.2d}$$

$$\underline{p}_u o_{u,t}\leq p_{u,t}\leq\bar{p}_u o_{u,t},\quad\forall u\in\mathcal{U},t\in[T], \tag{B.2e}$$

Table B.7
Parameters of the electricity system.

Parameter	Description
$\mathcal{G}^e = (N, \mathcal{E})$	Undirected graph where N is a set of buses indexed by $i = \{1, \dots, N\}$ and $\mathcal{E}$ is a set of lines indexed with $l = \{1, \dots, \mathcal{E}\}$
$\mathcal{U}$	Set of generators, indexed by $u = \{1, \dots, U\}$
$\mathcal{U}^g \subseteq U$	Set of GFPPs
$\mathcal{U}(i) \subseteq U$	Set of generators located at $i \in \mathcal{N}$
$\mathcal{B}_u$	Set of supply bids submitted by $u \in \mathcal{U}$, indexed by $b = 1, \dots, B_u$
$\beta_{u,b}$	Bid price of $b \in B_u$ submitted by $u \in \mathcal{U}$
$\overline{s}_{u,b}$	Maximum amount of active power generation of bid $b \in B_u$
$\overline{M}^e_u$	Maximum gap between actual emissions and initial allowances of GFPP $u \in \mathcal{U}^g$
$\underline{p}_u, \overline{p}_u$	Minimum/maximum real power generation of generator $u \in \mathcal{U}$
$\underline{R}_u, \overline{R}_u$	Ramp-down/-up rate of generator $u \in \mathcal{U}$
c_u	No-load cost of $u \in \mathcal{U}$
Ψ_u	Set of counts of time periods with distinct start-up costs of generator u indexed by h
$C_{u,h}$	Start up cost of generator $u \in \mathcal{U}$ when u is turned on after it has been offline for some time $\in [\Psi_{u,h}, \Psi_{u,h+1}]$
$\overline{o}_{u,0}, \overline{p}_{u,0}$	Initial on–off status/real power generation of generator $u \in \mathcal{U}$
$\overline{\tau}_u, \underline{\tau}_u$	Minimum up/down time of generator $u \in \mathcal{U}$
$\overline{\tau}_{u,0}, \underline{\tau}_{u,0}$	The time that generator $u \in \mathcal{U}$ has to be active/inactive from $t = 0$
b_l	Line susceptance of line $l \in \mathcal{E}$
f_l	Real power limit of line $l \in \mathcal{E}$
$E^0_{u,t}$	Allocated carbon allowances for GFPP $u \in \mathcal{U}^g$ during $t \in [T]$
κ_u	Carbon intensity (CI) parameters of GFPP $u \in \mathcal{U}^g$
$(d^e_{i,t})_{i \in \mathcal{N}}$	Electricity load profile during $t \in [T]$
Δ_l	Maximum voltage angle difference between two end-points of line $l \in \mathcal{E}$
$\underline{\theta}_i, \overline{\theta}_i$	Minimum/maximum voltage angle at node $i \in \mathcal{N}$

Table B.8
Variables of the UC problem.

Variable	Description
Binary variables	
$o_{u,t}$	1 if generator $u \in \mathcal{U}$ is on during $t \in [T]$, 0 otherwise
$v^+_{u,t}$	1 if generator $u \in \mathcal{U}$ becomes online during $t \in [T]$, 0 otherwise
$v^-_{u,t}$	1 if generator $u \in \mathcal{U}$ becomes offline during $t \in [T]$, 0 otherwise
$w_{u,b,t}$	1 if generator $b \in \mathcal{B}_u$ is selected during $t \in [T]$, 0 otherwise
$y_{u,t}$	1 if GFPP $u \in \mathcal{U}^g$ has deficit carbon allowances during $t \in [T]$, 0 otherwise
Continuous variables	
$r_{u,t}$	Start-up cost of generator $u \in \mathcal{U}$ during $t \in [T]$
$\rho_{u,t}$	Maximum allowable bidding price for generator $u \in \mathcal{U}^g$ to generate power at its scheduled level during $t \in [T]$

$$p_{u,0} = \overline{p}_{u,0}, \quad \forall u \in \mathcal{U}, \tag{B.2f}$$

$$p_{u,t} - p_{u,t-1} \le \overline{R}_u o_{u,t-1} + \overline{p}_u v^+_{u,t}, \quad \forall u \in \mathcal{U}, t \in [T], \tag{B.2g}$$

$$p_{u,t-1} - p_{u,t} \le \underline{R}_u o_{u,t} + \underline{p}_u v^-_{u,t}, \quad \forall u \in \mathcal{U}, t \in [T], \tag{B.2h}$$

$$f_{l,t} = -b_l(\theta_{l_h,t} - \theta_{l_t,t}), \quad \forall l \in \mathcal{E}, t \in [T], \tag{B.2i}$$

$$-\overline{f}_l \le f_{l,t} \le \overline{f}_l, \quad \forall l \in \mathcal{E}, t \in [T], \tag{B.2j}$$

$$\underline{\theta}_i \le \theta_{i,t} \le \overline{\theta}_i, \quad \forall i \in \mathcal{N}, t \in [T], \tag{B.2k}$$

$$-\Delta_l \le \theta_{l_h,t} - \theta_{l_t,t} \le \Delta_l, \quad \forall l \in \mathcal{E}, t \in [T]. \tag{B.2l}$$

The problem (B.2) is a classic ED problem and the objective of the ED is to minimize the costs of the selected supply bids of each power generating unit u. Table B.9 summarizes the variables of the ED problem. Eq. (B.2b) is a real power balance constraint with l_h and l_t representing the head and tail of line $l \in \mathcal{E}$, respectively. Eq. (B.2c) shows that the real power generation of unit u is equal to the sum of quantities of all selected bids and Eq. (B.2d) constrains the power generation $s^e_{u,b,t}$ from bid $b \in \mathcal{B}_u$ to be no more than the submitted amount $\overline{s}_b$. The bounds for real power generation $p_{u,t}$ is suggested in Eq. (B.2e) and the initial power at $t = 0$ is specified in (B.2f). The ramp-up and ramp-down constraints for generators are given in constraints (B.2g) and (B.2h). This paper assumes DC-OPF and the line flow $f_{l,t}$ is formulated in Eq. (B.2i) and it is bounded by flow limits as suggested in Eq. (B.2j). Eqs. (B.2k)–(B.2l) state the voltage angle bounds on each bus and the bounds on the angle difference of two adjacent buses respectively.

B.3. The third-level gas market clearing problem

This natural gas market clearing problem (B.3) is a classic steady-state model of natural gas network with Weymouth equation to capture the relationship between pressures and flux. Tables B.10 and B.11 summarize the parameters and variables of the gas market clearing problem.

$$\min \quad \sum_{t \in [T]} \sum_{j \in \mathcal{V}} \left(\sum_{s \in S_j} c_{j,s} s^g_{s,t} + k_j q_{j,t} \right) \tag{B.3a}$$

s.t.

$$s^g_{j,t} - l_{j,t} - l^{GFPP}_{j,t} = \sum_{a \in \mathcal{A}: a_t = j} \phi_{a,t} - \sum_{a \in \mathcal{A}: a_h = j} \phi_{a,t}, \quad \forall j \in \mathcal{V}, t \in \mathcal{T}, \tag{B.3b}$$

$$s^g_{j,t} = \sum_{s \in S_j} s^g_{s,t}, \quad \forall j \in \mathcal{V}, t \in [T], \tag{B.3c}$$

$$l_{j,t} = d^g_{j,t} - q_{j,t}, \quad \forall j \in \mathcal{V}, t \in [T], \tag{B.3d}$$

$$0 \le q_{j,t} \le d^g_{j,t}, \quad \forall j \in \mathcal{V}, t \in [T], \tag{B.3e}$$

$$\phi_{a,t} \ge 0, \quad \forall a \in \mathcal{A}, t \in \mathcal{T}, \tag{B.3f}$$

$$\underline{s}^g_j \le s^g_{j,t} \le \overline{s}^g_j, \quad \forall j \in \mathcal{V}, t \in [T], \tag{B.3g}$$

$$\underline{\alpha}^c_a \pi_{a_h,t} \le \pi_{a_t,t} \le \overline{\alpha}^c_a \pi_{a_h,t}, \quad \forall a \in \mathcal{A}_c, t \in [T], \tag{B.3h}$$

$$\underline{\alpha}^v_a \pi_{a_h,t} \le \pi_{a_t,t} \le \overline{\alpha}^v_a \pi_{a_h,t}, \quad \forall a \in \mathcal{A}_v, t \in [T], \tag{B.3i}$$

$$\pi_{a_h,t} - \pi_{a_t,t} \ge W_a \phi^2_{a,t}, \quad \forall a \in \mathcal{A} \setminus (\mathcal{A}_v \cup \mathcal{A}_c), t \in [T], \tag{B.3j}$$

Table B.9
Variables of the ED problem.

Variable	Description
$s^e_{u,b,t}$	Real power generation from $b \in \mathcal{B}_u$ of generator $u \in \mathcal{U}$ during $t \in \mathcal{T}$
$p_{u,t}$	Real power generation of generator $u \in \mathcal{U}$ during $t \in \mathcal{T}$
$f_{l,t}$	Real power flow on line $l \in \mathcal{E}$ during $t \in \mathcal{T}$
$\theta_{i,t}$	Voltage angle on node $i \in \mathcal{N}$ during $t \in \mathcal{T}$

Table B.10
Parameters of the gas system.

Parameter	Description
$\mathcal{G} = (\mathcal{V}, \mathcal{A})$	Directed graph representing a natural gas transmission network, where $\mathcal{V}$ is a set of junctions, indexed with $j = \{1, \dots, V\}$, and $\mathcal{A} \subseteq \mathcal{V} \times \mathcal{V}$ is a set of connections, indexed with $a = \{1, \dots, A\}$
$\mathcal{A}_c \subseteq \mathcal{A}$	Set of compressors
$\mathcal{A}_v \subseteq \mathcal{A}$	Set of control valves
κ_j	Cost of demand shedding at junction $j \in \mathcal{V}$
$(d^g_{j,t})_{j \in \mathcal{V}}$	Gas demand profile during $t \in \mathcal{T}$
$\overline{s}^g_j, \underline{s}^g_j$	Lower/Upper limit on natural gas supply at junction $j \in \mathcal{V}$
$c_j(\cdot)$	Cost function for gas supply at junction $j \in \mathcal{V}$
$\{H^G_{u,j}\}_{j=0,1,2}$	Coefficients of the heat rate curve of GFPP $u \in \mathcal{U}^g$
W_a	Pipeline resistance (Weymouth) factor of connection $a \in \mathcal{A}$
$\underline{\pi}_j, \overline{\pi}_j$	Minimum/maximum squared pressure at junction $j \in \mathcal{V}$
$\underline{\alpha}^c_a, \overline{\alpha}^c_a$	Lower/upper compression ratio of connection $a \in \mathcal{A}_c$
$\underline{\alpha}^v_a, \overline{\alpha}^v_a$	Lower/upper control ratio of connection $a \in \mathcal{A}_v$
$\mathcal{K}$	Set of pricing zones, indexed by $k = \{1, \dots, K\}$
$\mathcal{V}(k)$	Set of junctions that belong to $k \in \mathcal{K}$

Table B.11
Variables on the gas system.

Variable	Description
$s^g_{j,t}$	Amount of gas supplied at junction $j \in \mathcal{V}$ during $t \in [T]$
$s_{s,t}$	Amount of gas supply from $s \in S_j$ during $t \in [T]$
$\pi_{j,t}$	Pressure squared at $j \in \mathcal{V}$ during $t \in [T]$
$\phi_{a,t}$	Gas flow on $a \in \mathcal{A}$ during $t \in [T]$
$l_{j,t}$	Satisfied gas demand at $j \in \mathcal{V}$ during $t \in [T]$
$q_{j,t}$	Gas demand shedding at $j \in \mathcal{V}$ during $t \in [T]$
$l^{GFPP}_{j,t}$	Total amount of gas consumed by the GFPP located at $j \in \mathcal{N} \cap \mathcal{V}$ during $t \in [T]$

$$\underline{\pi}_j \le \pi_{j,t} \le \overline{\pi}_j, \quad \forall j \in \mathcal{V}, t \in [T], \tag{B.3k}$$

The heat rate curve constraint (9). (B.3l)

Eq. (B.3b) shows the flux conservation constraint with a_h and a_t representing the head and tail of connections $a \in \mathcal{A}$, respectively. Eq. (B.3c) calculates the total gas production at junction $j \in \mathcal{V}$. The satisfied gas demand is calculated based on the gas demand profile $d^g_{j,t}$ minus the gas load shedding $q_{j,t}$ (Eq. (B.3d)), where there exists a high penalty cost for gas demand shedding. In addition, the gas demand shedding should not exceed the demand profile, as shown in Eq. (B.3e). This model assumes a predetermined gas flow direction as enforced by the non-negativity of $\phi_{a,t}$ in Eq. (B.3f). The bounds for natural gas supplies $s^g_{j,t}$ are $[\underline{s}^g_j, \overline{s}^g_j]$. In addition, a single compressor machine approximation is assumed in this work and the change in pressure through compressors and control valves are stated in Eqs. (B.3h)–(B.3i). The relaxed Weymouth equation, as stated in Eq. (B.3j), is used to model steady state gas flows. The bounds on the nodal pressure is modeled in (B.3k). Lastly, Eq. (B.3l) shows the heat rate curve for calculating the demand induced by GFPPs $l^{GFPP}_{j,t}$. The gas balancing constraint (B.3b) yields dual variable $\pi^g_{j,t}$, representing the nodal gas price at junction j for time period t.

The cost structure at each gas supply junction j consists of two components: a production cost represented by a piecewise linear function, and a gas shedding penalty κ_j. The production cost curve is divided into multiple segments S_j, spanning the range $[0, \overline{s}^g_j]$. Each segment $s \in S_j$ has an associated marginal cost $c_{j,s}$, which is non-decreasing across consecutive segments ($c_{j,s} \le c_{j,s+1}$). To model this segmented structure, we define supply variables $s^g_{j,t}$ representing the gas quantity produced from each segment s at time t. These segment-specific supply variables are connected to the junction's total gas flow through constraint (B.3c), with the total cost captured in objective function (B.3a).

B.4. The fourth-level carbon market clearing problem

The fourth level of the UCGCA model is the carbon emission trading scheme and is detailed in Problem (B.4). Tables B.12 and B.13 provide a summary of the parameters and variables.

$$\min \quad \sum_{t \in [T]} \left(\sum_{i \in \mathcal{U} \cup \mathcal{V}} c_{i,t} e_{i,t} + \kappa^e q^e_t \right) \tag{B.4a}$$

$$\text{s.t.} \quad 0 \le e_{i,t} \le E^{surplus}_{i,t}, \quad \forall i \in \mathcal{U} \cup \mathcal{V}, t \in [T], \tag{B.4b}$$

$$\sum_{i \in \mathcal{U} \cup \mathcal{V}} e_{i,t} + q^e_t = \sum_{i \in \mathcal{U} \cup \mathcal{V}} E^{deficit}_{i,t}, \quad [\pi^e_t] \quad \forall t \in [T]. \tag{B.4c}$$

We adopt an allowance-based carbon market architecture where each participant receives an initial allocation of carbon allowances $E^0_{i,t}$. Eq. (B.4a) shows the objective of the carbon trading mechanism, which aims to minimize the costs associated with carbon emissions trading. This objective function includes costs from internal transactions within the market, represented as $c_{i,t} e_{i,t}$, and external acquisitions of emission quotas as a last resort, denoted by $\kappa^e q^e_t$, where q^e_t incurs a higher cost due to external purchases. Market participants can trade their allowances based on their positions. Those experiencing deficits, where actual emissions $E_{i,t}$ exceed their initial allowances $E^0_{i,t}$, must acquire additional allowances to meet this inelastic emission demand $E^{deficit}_{i,t}$. Conversely, participants with surpluses, where initial allowances $E^0_{i,t}$ exceed their actual emissions $E_{i,t}$, can sell their excess allowances $e_{i,t}$ at bid prices $c_{i,t}$, as constrained by (B.4b). The calculation of actual carbon

Table B.12
Parameters of the carbon trading mechanism.

Parameters	Description
$c_{i,t}$	Offers to sell surplus carbon allowances by unit i during time period $t \in [T]$.
k^e	Cost incurred for acquiring additional external carbon allowances.
$E^0_{i,t}$	Allocated carbon allowances for unit i during time period $t \in [T]$.
κ_i	Carbon intensity (CI) parameters for unit i.

Table B.13
Variables of the carbon trading mechanism.

Variables	Description
$e_{i,t}$	Amount of emission allowances unit i sold during time period $t \in [T]$.
q^e_t	Quantity of external emission allowances bought at price k^e during time period $t \in [T]$.
$E^*_{i,t}$	Actual carbon emissions produced by unit i during period t.
$E^{deficit}_{i,t}$	Carbon allowance deficit for participant i in period t
$E^{surplus}_{i,t}$	Carbon allowance surplus for participant i in period t

Table C.14
Computational performance for C and B across different values of η_e and η_g levels when $\alpha = 0.8$.

η_g	$\eta_e = 1$				$\eta_e = 1.2$				$\eta_e = 1.6$			
	C		B		C		B		C		B	
	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)
1.0	0.00	6.27	0.00	40.53	0.00	8.21	0.00	25.43	0.00	5.82	0.00	57.29
1.1	0.00	5.46	0.00	34.61	0.00	6.74	0.00	58.20	0.00	2.99	0.00	43.04
1.2	0.00	5.46	0.00	23.69	0.00	3.44	0.00	89.05	0.00	1.77	0.00	42.56
1.3	0.00	4.52	0.00	24.20	0.00	5.51	0.00	57.81	0.00	7.53	0.00	34.65
1.4	0.00	9.62	0.00	39.54	0.00	5.26	0.00	61.44	0.00	2.47	0.00	44.29
1.5	0.00	7.56	0.00	25.84	0.00	4.97	0.00	82.48	0.02	3600 (†)	0.00	626.37
1.6	0.00	6.49	0.00	37.43	0.15	3600 (†)	0.00	62.29	100 (‡)	3600 (†)	2.69	3600 (†)
1.7	0.00	6.48	0.00	23.87	4.50	3600 (†)	3.47	3600 (†)	100 (‡)	3600 (†)	2.57	3600 (†)
1.8	6.34	3600 (†)	1.64	3600 (†)	100 (‡)	3600 (†)	3.54	3600 (†)	100 (‡)	3600 (†)	6.86	3600 (†)
1.9	100 (‡)	3600 (†)	1.78	3600 (†)	100 (‡)	3600 (†)	3.31	3600 (†)	100 (‡)	3600 (†)	7.85	3600 (†)
2.0	4.24	3600 (†)	0.00	66.10	2.67	3600 (†)	4.79	3600 (†)	9.62	3600 (†)	5.30	3600 (†)
2.1	3.92	3600 (†)	4.67	3600 (†)	5.22	3600 (†)	5.11	3600 (†)	100 (‡)	3600 (†)	5.27	3600 (†)
2.2	4.39	3600 (†)	5.30	3600 (†)	5.70	3600 (†)	4.83	3600 (†)	6.70	3600 (†)	5.05	3600 (†)
2.3	0.00	38.42	0.00	30.37	0.00	78.00	0.00	53.30	0.00	42.23	0.00	18.23

Table C.15
Computational performance for C and B across different values of η_e and η_g levels when $\alpha = 1.0$.

η_g	$\eta_e = 1$				$\eta_e = 1.2$				$\eta_e = 1.6$			
	C		B		C		B		C		B	
	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)
1.0	0.00	5.59	0.00	38.44	0.00	5.09	0.00	25.97	0.00	3.01	0.00	55.63
1.1	0.00	9.10	0.00	35.12	0.00	4.67	0.00	59.02	0.00	2.77	0.00	41.06
1.2	0.00	8.15	0.00	24.08	0.00	5.42	0.00	88.72	0.00	3.40	0.00	42.62
1.3	0.00	7.69	0.00	24.73	0.00	7.46	0.00	57.58	0.00	5.59	0.00	43.79
1.4	0.00	6.12	0.00	38.10	0.00	5.91	0.00	59.80	0.00	3.64	0.00	37.81
1.5	0.00	7.23	0.00	23.63	0.00	5.79	0.00	82.46	0.00	893.21	0.00	88.97
1.6	0.00	5.91	0.00	37.08	0.00	8.95	0.00	79.54	100 (‡)	3600 (†)	2.60	3600 (†)
1.7	0.00	5.72	0.00	23.68	0.80	3600 (†)	3.74	3600 (†)	100 (‡)	3600 (†)	2.55	3600 (†)
1.8	3.33	3600 (†)	1.48	3600 (†)	100 (‡)	3600 (†)	3.50	3600 (†)	100 (‡)	3600 (†)	2.40	3600 (†)
1.9	2.96	3600 (†)	1.77	3600 (†)	100 (‡)	3600 (†)	3.52	3600 (†)	100 (‡)	3600 (†)	2.52	3600 (†)
2.0	4.15	3600 (†)	0.00	65.69	2.05	3600 (†)	0.00	144.64	0.00	22.35	0.00	34.15
2.1	3.67	3600 (†)	0.00	54.43	1.91	3600 (†)	0.00	145.36	0.72	3600 (†)	0.00	30.20
2.2	3.75	3600 (†)	0.00	52.53	0.21	3600 (†)	0.00	179.46	0.00	2.97	0.00	22.71
2.3	2.44	3600 (†)	0.00	37.70	0.00	2185.13	0.00	152.32	0.00	13.29	0.00	45.98

emission levels is discussed in Eq. (B.4). The carbon emission balancing constraint (B.4c) yields dual variable π^e_t, representing the carbon price for time period t.

Appendix C. Computational performance

We evaluate the performances of C and B across various scenarios, using combinations of gas network stress ($\eta_g = 1, 1.1, \ldots, 2.3$), electrical system conditions ($\eta_e = 1, 1.2, 1.6$), and risk-aversion levels ($\alpha_u = 80\%, 100\%, 120\%$). The comparison was made across all instances, as provided in Tables C.14–C.16. The symbol † indicates that a method reaches the time limit and the symbol ‡ that the method did not find any incumbent solution.

Appendix D. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cie.2026.112112.

Data availability

Data will be made available on request.

Table C.16
Computational performance for C and B across different values of η_e and η_g levels when $\alpha = 1.2$.

η_g	$\eta_e = 1$				$\eta_e = 1.2$				$\eta_e = 1.6$			
	C		B		C		B		C		B	
	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)	Gap (%)	Time (s.)
1.0	0.00	6.54	0.00	38.26	0.00	6.59	0.00	25.69	0.00	2.74	0.00	56.20
1.1	0.00	7.78	0.00	34.78	0.00	4.98	0.00	59.48	0.00	4.05	0.00	41.57
1.2	0.00	9.27	0.00	23.99	0.00	6.04	0.00	90.13	0.00	3.15	0.00	42.48
1.3	0.00	6.41	0.00	24.60	0.00	4.60	0.00	58.08	0.00	3.57	0.00	29.14
1.4	0.00	6.91	0.00	38.02	0.00	4.20	0.00	60.56	0.00	2.22	0.00	46.28
1.5	0.00	7.25	0.00	23.69	0.00	7.21	0.00	82.93	0.00	1581.03	0.00	614.12
1.6	0.00	7.36	0.00	36.63	0.78	3600 (†)	0.00	72.40	100 (‡)	3600 (†)	2.52	3600 (†)
1.7	0.00	6.94	0.00	23.53	4.56	3600 (†)	4.49	3600 (†)	100 (‡)	3600 (†)	2.47	3600 (†)
1.8	6.42	3600 (†)	1.90	3600 (†)	100 (‡)	3600 (†)	3.65	3600 (†)	100 (‡)	3600 (†)	2.39	3600 (†)
1.9	9.50	3600 (†)	1.68	3600 (†)	100 (‡)	3600 (†)	3.07	3600 (†)	100 (‡)	3600 (†)	2.38	3600 (†)
2.0	3.53	3600 (†)	0.00	57.87	0.00	1865.32	0.00	104.05	1.59	3600 (†)	0.00	55.19
2.1	3.53	3600 (†)	0.00	57.05	0.00	54.49	0.00	352.49	0.00	7.67	0.00	51.96
2.2	3.58	3600 (†)	0.00	54.30	0.00	923.93	0.00	169.06	0.00	7.77	0.00	36.24
2.3	4.19	3600 (†)	0.00	60.33	0.00	19.19	0.00	141.28	0.00	14.10	0.00	31.32

References

Addula, S. R., & Sekhar Sajja, G. (2024). Automated machine learning to streamline data-driven industrial application development. In *2024 second international conference computational and characterization techniques in engineering & sciences* (pp. 1–4).

Allen, E. H., Lang, J. H., & Ilic, M. D. (2008). A combined equivalenced-electric, economic, and market representation of the northeastern power coordinating council U.S. electric power system. *IEEE Transactions on Power Systems*, *23*(3), 896–907.

Avraamidou, S., & Pistikopoulos, E. (2019). Multi-parametric global optimization approach for tri-level mixed-integer linear optimization problems. *Journal of Global Optimization*, *74*, 443–465.

Beck, Y., Ljubić, I., & Schmidt, M. (2023). A survey on bilevel optimization under uncertainty. *European Journal of Operational Research*, *311*(2), 401–426, URL https://www.sciencedirect.com/science/article/pii/S0377221723000073.

Bent, R., Blumsack, S., Van Hentenryck, P., Borraz-Sánchez, C., & Shahriari, M. (2018). Joint electricity and natural gas transmission planning with endogenous market feedbacks. *IEEE Transactions on Power Systems*, *33*(6), 6397–6409.

Borraz-Sanchez, C., Bent, R., van Hentenryck, P., Blumsack, S., & Hijazi, H. (2016). Elasticity model for joint gas-grid expansion planning optimization. (pp. PSIG–1610). arXiv:https://onepetro.org/PSIGAM/proceedings-pdf/PSIG16/All-PSIG16/PSIG-1610/1332155/psig-1610.pdf.

Byeon, G. (2020). *Large-scale optimization for interdependent infrastructure systems* (Ph.D. thesis), University of Michigan, URL https://hdl.handle.net/2027.42/155075.

Byeon, G., & Van Hentenryck, P. (2020). Unit commitment with gas network awareness. *IEEE Transactions on Power Systems*, *35*(2), 1327–1339.

Byeon, G., & Van Hentenryck, P. (2022). Benders subproblem decomposition for bilevel problems with convex follower. *INFORMS Journal on Computing*, *34*(3), 1749–1767.

California Air Resources Board (2024). Cap-and-trade program. https://ww2.arb.ca.gov/our-work/programs/cap-and-trade-program.

Chen, S., Conejo, A. J., & Wei, Z. (2021). Conjectural-variations equilibria in electricity, natural-gas, and carbon-emission markets. *IEEE Transactions on Power Systems*, *36*(5), 4161–4171.

Chen, S., Conejo, A. J., & Wei, Z. (2022). Gas-power coordination: From day-ahead scheduling to actual operation. *IEEE Transactions on Power Systems*, *37*(2), 1532–1542.

Cheng, S., Scholes, S. C., Kong, W., Gu, C., & Li, F. (2023). Carbon-oriented electricity balancing market for dispatchable generators and flexible loads. *IEEE Transactions on Power Systems*, *38*(6), 5648–5659.

Chouhan, V. K., Khan, S. H., & Hajiaghaei-Keshteli, M. (2022). Hierarchical tri-level optimization model for effective use of by-products in a sugarcane supply chain network. *Applied Soft Computing*, *128*, Article 109468, URL https://www.sciencedirect.com/science/article/pii/S1568494622005750.

Dhavale, D. G., & Sarkis, J. (2018). Stochastic internal rate of return on investments in sustainable assets generating carbon credits. *Computers & Operations Research*, *89*, 324–336, URL https://www.sciencedirect.com/science/article/pii/S0305054817300515.

Dimitriadis, C. N., Tsimopoulos, E. G., & Georgiadis, M. C. (2023). Optimal bidding strategy of a gas-fired power plant in interdependent low-carbon electricity and natural gas markets. *Energy*, *277*, Article 127710, URL https://www.sciencedirect.com/science/article/pii/S0360544223011040.

Dvorkin, Y., Fernández-Blanco, R., Wang, Y., Xu, B., Kirschen, D. S., Pandžić, H., Watson, J.-P., & Silva-Monroy, C. A. (2018). Co-planning of investments in transmission and merchant energy storage. *IEEE Transactions on Power Systems*, *33*(1), 245–256.

Ehrgott, M. (2005). *Multicriteria optimization*: *Vol. 491*, Berlin, Germany: Springer.

El-Meligy, M. A., Sharaf, M., & Soliman, A. T. (2021). A coordinated scheme for transmission and distribution expansion planning: A tri-level approach. *Electric Power Systems Research*, *196*, Article 107274, URL https://www.sciencedirect.com/science/article/pii/S0378779621002558.

European Commission (2024). EU emissions trading system (EU ETS). Available at: https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets_en.

Fakhry, R., Hassini, E., Ezzeldin, M., & El-Dakhakhni, W. (2022). Tri-level mixed-binary linear programming: Solution approaches and application in defending critical infrastructure. *European Journal of Operational Research*, *298*(3), 1114–1131, URL https://www.sciencedirect.com/science/article/pii/S0377221721006299.

Federal Energy Regulatory Commission (2022). RTO unit commitment test system. https://www.ferc.gov/power-sales-and-markets/increasing-efficiency-through-improved-software/rto-unit-commitment-test. (Accessed 27 June 2024).

Fischetti, M., Ljubić, I., Monaci, M., & Sinnl, M. (2018). On the use of intersection cuts for bilevel optimization. *Mathematical Programming*, *172*(1), 77–103.

Fischetti, M., Ljubić, I., & Sinnl, M. (2017). Redesigning benders decomposition for large-scale facility location. *Management Science*, *63*(7), 2146–2162.

Fischetti, M., Salvagnin, D., & Zanette, A. (2010). A note on the selection of Benders' cuts. *Mathematical Programming*, *124*(1), 175–182.

Florensa, C., Garcia-Herreros, P., Misra, P., Arslan, E., Mehta, S., & Grossmann, I. E. (2017). Capacity planning with competitive decision-makers: Trilevel MILP formulation, degeneracy, and solution approaches. *European Journal of Operational Research*, *262*(2), 449–463, URL https://www.sciencedirect.com/science/article/pii/S0377221717303466.

Gan, W., Shahidehpour, M., Guo, J., Yao, W., Pandey, S., Paaso, E. A., Vukojevic, A., & Wen, J. (2022). A tri-level planning approach to resilient expansion and hardening of coupled power distribution and transportation systems. *IEEE Transactions on Power Systems*, *37*(2), 1495–1507.

Ghorbani-Renani, N., González, A. D., & Barker, K. (2021). A decomposition approach for solving tri-level defender-attacker-defender problems. *Computers & Industrial Engineering*, *153*, Article 107085, URL https://www.sciencedirect.com/science/article/pii/S0360835220307555.

Gil, M., Dueñas, P., & Reneses, J. (2016). Electricity and natural gas interdependency: Comparison of two methodologies for coupling large market models within the European regulatory framework. *IEEE Transactions on Power Systems*, *31*(1), 361–369.

Hua, H., Chen, X., Gan, L., Sun, J., Dong, N., Liu, D., Qin, Z., Li, K., & Hu, S. (2024). Demand-side joint electricity and carbon trading mechanism. *IEEE Transactions on Industrial Cyber-Physical Systems*, *2*, 14–25.

Jiang, K., Liu, N., Yan, X., Xue, Y., & Huang, J. (2023). Modeling strategic behaviors for GenCo with joint consideration on electricity and carbon markets. *IEEE Transactions on Power Systems*, *38*(5), 4724–4738.

Jiao, F., Zou, Y., Zhang, X., & Zhang, B. (2022). A three-stage multitimescale framework for online dispatch in a microgrid with EVs and renewable energy. *IEEE Transactions on Transportation Electrification*, *8*(1), 442–454.

Khan, H. A. U., Kim, J., & Dvorkin, Y. (2022). Risk-informed participation in T&D markets. *Electric Power Systems Research*, *202*, Article 107624, URL https://www.sciencedirect.com/science/article/pii/S0378779621006052.

Kleinert, T., Labbé, M., Ljubić, I., & Schmidt, M. (2021). A survey on mixed-integer programming techniques in bilevel optimization. *EURO Journal on Computational Optimization*, *9*, Article 100007, URL https://www.sciencedirect.com/science/article/pii/S2192440621001349.

Li, J., Ge, S., Liu, H., Yu, Q., Zhang, S., Wang, C., & Gu, C. (2024). An electricity and carbon trading mechanism integrated with TSO-DSO-prosumer coordination. *Applied Energy*, *356*, Article 122328, URL https://www.sciencedirect.com/science/article/pii/S0306261923016926.

Lu, J., Han, J., Hu, Y., & Zhang, G. (2016). Multilevel decision-making: A survey. *Information Sciences*, *346–347*, 463–487, URL https://www.sciencedirect.com/science/article/pii/S0020025516300202.

McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part I — Convex underestimating problems. *Mathematical Programming*, 146–175.

Mitridati, L., Kazempour, J., & Pinson, P. (2020). Heat and electricity market coordination: A scalable complementarity approach. *European Journal of Operational Research*, *283*(3), 1107–1123, URL https://www.sciencedirect.com/science/article/pii/S0377221719309841.

Mitridati, L., Van Hentenryck, P., & Kazempour, J. (2019). A bid-validity mechanism for sequential heat and electricity market clearing. arXiv:1910.08617.

Muñoz-Delgado, G., Contreras, J., & Arroyo, J. M. (2019). Distribution system expansion planning considering non-utility-owned DG and an independent distribution system operator. *IEEE Transactions on Power Systems*, *34*(4), 2588–2597.

Nasrolahpour, E., Kazempour, J., Zareipour, H., & Rosehart, W. D. (2018). A bilevel model for participation of a storage system in energy and reserve markets. *IEEE Transactions on Sustainable Energy*, *9*(2), 582–598.

National Energy System Operator (NESO) (2025). Strategic planning – what we do. Available at: https://www.neso.energy/what-we-do/strategic-planning.

Ofgem (2021). *Review of GB energy system operation: Tech. rep.*, London, UK: Office of Gas and Electricity Markets.

Ofgem (2025). Gas system operator (GSO) regulation. Available at: https://www.ofgem.gov.uk/energy-policy-and-regulation/policy-and-regulatory-programmes/gas-system-operator-gso-regulation.

Ordoudis, C., Pinson, P., & Morales, J. M. (2019). An integrated market for electricity and natural gas systems with stochastic power producers. *European Journal of Operational Research*, *272*(2), 642–654, URL https://www.sciencedirect.com/science/article/pii/S037722171830571X.

Paredes, Á., Aguado, J., Essayeh, C., Xia, Y., Savelli, I., & Morstyn, T. (2023). Stacking revenues from flexible DERs in multi-scale markets using tri-level optimization. *IEEE Transactions on Power Systems*, 1–13.

PJM (2014). *Analysis of Operational Events and Market Impacts During the January 2014 Cold Weather Events: Technical report*, Norristown, PA, USA: PJM.

Putratama, M. A., Rigo-Mariani, R., Mustika, A. D., Debusschere, V., Pachurka, A., & Bésanger, Y. (2023). A three-stage strategy with settlement for an energy community management under grid constraints. *IEEE Transactions on Smart Grid*, *14*(2), 1505–1514.

Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, *259*(3), 801–817, URL https://www.sciencedirect.com/science/article/pii/S0377221716310244.

Rintamäki, T., Siddiqui, A. S., & Salo, A. (2020). Strategic offering of a flexible producer in day-ahead and intraday power markets. *European Journal of Operational Research*, *284*(3), 1136–1153, URL https://www.sciencedirect.com/science/article/pii/S0377221720300813.

Sadeghi, S., Seifi, A., & Azizi, E. (2017). Trilevel shortest path network interdiction with partial fortification. *Computers & Industrial Engineering*, *106*, 400–411, URL https://www.sciencedirect.com/science/article/pii/S0360835217300670.

Sato, R., Tanaka, M., & Takeda, A. (2021). A gradient method for multilevel optimization. *Advances in Neural Information Processing Systems*, *34*, 7522–7533.

Savelli, I., De Paola, A., & Li, F. (2020). Ex-ante dynamic network tariffs for transmission cost recovery. *Applied Energy*, *258*, Article 113979, URL https://www.sciencedirect.com/science/article/pii/S0306261919316666.

Savelli, I., Hardy, J., Hepburn, C., & Morstyn, T. (2022). Putting wind and solar in their place: Internalising congestion and other system-wide costs with enhanced contracts for difference in Great Britain. *Energy Economics*, *113*, Article 106218, URL https://www.sciencedirect.com/science/article/pii/S0140988322003656.

Savelli, I., & Morstyn, T. (2021). Electricity prices and tariffs to keep everyone happy: A framework for fixed and nodal prices coexistence in distribution grids with optimal tariffs for investment cost recovery. *Omega*, *103*, Article 102450, URL https://www.sciencedirect.com/science/article/pii/S0305048321000591.

Schwele, A., Ordoudis, C., Pinson, P., & Kazempour, J. (2021). Coordination of power and natural gas markets via financial instruments. *Computational Management Science*, *18*, 505–538.

Shafiei, A., Kungurtsev, V., & Marecek, J. (2024). Trilevel and multilevel optimization using monotone operator theory. *Mathematical Methods of Operations Research*, *99*, 77–114.

Tilahun, S. L., Kassa, S. M., & Ong, H. C. (2012). A new algorithm for multilevel optimization problems using evolutionary strategy, inspired by natural adaptation. In *PRICAI 2012: trends in artificial intelligence* (pp. 577–588). Berlin, Heidelberg: Springer Berlin Heidelberg.

U. K. Government (2016). EU ETS: Carbon markets. Available at: https://www.gov.uk/guidance/eu-ets-carbon-markets.

United States Environmental Protection Agency (2023). Greenhouse gases equivalencies calculator - calculations and references. https://www.epa.gov/energy/greenhouse-gases-equivalencies-calculator-calculations-and-references. (Accessed 08 October 2024).

Wang, C., Wei, W., Wang, J., Liu, F., & Mei, S. (2018). Strategic offering and equilibrium in coupled gas and electricity markets. *IEEE Transactions on Power Systems*, *33*(1), 290–306.

Wozabal, D., & Rameseder, G. (2020). Optimal bidding of a virtual power plant on the Spanish day-ahead and intraday market for electricity. *European Journal of Operational Research*, *280*(2), 639–655, URL https://www.sciencedirect.com/science/article/pii/S0377221719305867.

Wu, X., & Conejo, A. J. (2017). An efficient tri-level optimization model for electric grid defense planning. *IEEE Transactions on Power Systems*, *32*(4), 2984–2994.

Xia, Y., Savelli, I., & Morstyn, T. (2024). An incentive regulation approach for balancing stakeholder interests in transmission merchant investment. *Electric Power Systems Research*, *234*, Article 110617, URL https://www.sciencedirect.com/science/article/pii/S0378779624005030.

Xia, Y., Savelli, I., & Morstyn, T. (2025). Integrating local market operations into transmission investment: A tri-level optimization approach. *Applied Energy*, *378*, Article 124721, URL https://www.sciencedirect.com/science/article/pii/S0306261924021044.

Xu, P., & Wang, L. (2014). An exact algorithm for the bilevel mixed integer linear programming problem under three simplifying assumptions. *Computers & Operations Research*, *41*, 309–318, URL https://www.sciencedirect.com/science/article/pii/S0305054813001950.

Zare, M., Borrero, J., Zeng, B., & Rodríguez-Martín, I. (2019). A note on linearized reformulations for a class of bilevel linear integer problems. *Annals of Operations Research*, *272*(1), 99–117.

Zhu, D., Yang, B., Wu, Y., Deng, H., Dong, Z., Ma, K., & Guan, X. (2024). Joint trading and scheduling among coupled carbon-electricity-heat-gas industrial clusters. *IEEE Transactions on Smart Grid*, *15*(3), 3152–3164.