Sustainable Energy, Grids and Networks 34 (2023) 101011

Contents lists available at ScienceDirect

Sustainable Energy, Grids and Networks

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

Privacy-preserving energy trading management in networked microgrids via data-driven robust optimization assisted by machine learning

Shayan Mohseni [a], Mir Saman Pishvaee [a,*], Reza Dashti [b]

[a] School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran
[b] School of Advanced Technologies, Iran University of Science and Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 18 August 2022
Received in revised form 20 December 2022
Accepted 16 January 2023
Available online 18 January 2023

Keywords:
Data-driven optimization
Peer-to-peer energy trading
Nash bargaining game
ADMM
Distributed optimization

ABSTRACT

Trading mechanism design, uncertainty treatment and privacy protection are the main issues in the energy management of networked microgrids (MGs). To address these issues in a comprehensive manner, this paper proposes a data-driven two-level transactive energy management framework, where the upper-level determines the optimal strategies of internal scheduling within MGs and external trading between MGs while the lower-level formulates a Nash bargaining game model for the fair allocation of trading benefits. The uncertainties of renewable energy sources are fully captured by an adjustable data-driven robust optimization approach with an uncertainty set constructed using the robust kernel density estimation (RKDE) as a machine learning technique. The resulting uncertainty set can provide robust scheduling and trading schemes even when the power generation data of wind turbines and photovoltaic systems are contaminated with anomalous samples, whereas conventional sets are not reliable in the case of contaminated data. To preserve the operational independence and information privacy of MGs, the proposed model is solved in a distributed manner by the alternating direction method of multipliers (ADMM) and the augmented Lagrange-based alternating direction inexact Newton (ALADIN) algorithms. ADMM is commonly used in previous studies, but it has low computational efficiency for handling the consensus process among a large number of MGs. This paper applies ADMM and ALADIN to enhance the applicability of the proposed model when the size and complexity of the networks increase. Numerical tests show the effectiveness of the proposed framework and solution methodology in terms of system cost, solution robustness, and convergence speed.

1. Introduction

1.1. Motivation

Recent years have seen the rapid adoption of renewable energy due to growing concerns about energy security and environmental pollution caused by conventional power generation. Since the dependence on individual renewable energy sources is highly unreliable, it is of great importance to develop hybrid renewable energy systems (HRESs) that combine two or more renewable resources [1]. Microgrid (MG) is a controllable entity of interconnected distributed energy resources (DERs) such as microturbines (MTs), photovoltaic systems (PVs), wind turbines (WTs) and battery energy storage systems, which can operate in stand-alone and grid-connected modes. With its flexible integration capabilities, MG is expected to become an important part of future distribution networks (DNs) [2]. As the number of MGs in a certain area increases, it is beneficial to form a multi-MG system connecting several individual MGs that are close to each other geographically. This system enables peer-to-peer (P2P) power trading among MGs, thereby relieving the burden on the utility grid and improving the reliability of energy supply [3,4]. The networking strategy, however, poses new technical challenges to the performance of power systems. DNs and MGs are generally operated and owned by private entities with different utility functions, making it difficult to design a power trading mechanism that satisfies the interests of all the entities without violating their technical constraints [5,6]. Moreover, the location-dependent and time-varying nature of renewable energy sources can lead to power system instability and generation-consumption unbalancing problems [7]. Therefore, in order to ensure system-wide stability and avoid suboptimal performance, it is necessary to effectively coordinate power scheduling and trading for networked MGs, especially when they are subject to the uncertainty of renewable generation [8]. This calls for designing a scheduling

* Corresponding author.
E-mail address: pishvaee@iust.ac.ir (M.S. Pishvaee).

https://doi.org/10.1016/j.segan.2023.101011
2352-4677/

Nomenclature

Indices and Sets

$\mathcal{M}$	Set of MGs
$\mathcal{D}$	Set of DN
$\mathcal{A}$	Set of all entities ($\mathcal{A} := \mathcal{M} \cup \mathcal{D}$)
$i/e/t$	Index of buses/entities/time periods
$\mathcal{N}_e^c$	Set of bus connecting entity e to another entity
$\mathcal{N}_e^a$	Set of all buses belonging to entity e
$\mathcal{N}^m$	Set of bus connecting DN to the main grid
$\pi(i)/\delta(i)$	Set of buses whose child/parent is bus i
$\mathcal{N}_e^{es}/\mathcal{N}_e^{r}/\mathcal{N}_e^{mt}$	Set of buses of ESSs/RGs/MTs belonging to entity e
$\tilde{\cdot}/\hat{\cdot}$	Actual value/forecasted value of uncertain parameters
$\cdot^M/\cdot^D$	Index of MG/DN

Parameters

$\bar{I}$	Maximum allowed line current
r_i	Resistance of line i- π (i)
x_i	Reactance of line i- π (i)
$\overline{V}/\underline{V}$	Upper/lower bound of bus voltage
η^c/η^{dc}	Charge/discharge efficiency of ESS
$\overline{P}^c/\overline{P}^{dc}$	Maximum charge/discharge power of ESS
$\overline{E}^{es}/\underline{E}^{es}$	Upper/lower capacity limit for ESS
η^c/η^{dc}	Charge/discharge efficiency of ESS
$\overline{Q^m}/\underline{Q^m}$	Upper/lower bound of reactive power exchanged between DN and the main grid
$\overline{P^{ex}}/\underline{P^{ex}}$	Upper/lower bound of active power exchanged between entities
$\overline{Q^{ex}}/\underline{Q^{ex}}$	Upper/lower bound of reactive power exchanged between entities
$\overline{P}^g/\underline{P}^g$	Upper/lower capacity limit for MT
$\rho^-_{e,t}/\rho^+_{e,t}$	Penalty cost for power shortage/surplus of entity e at time t
$\vartheta^s_t/\vartheta^b_t$	Prices of selling/buying power to/from the main grid at time t
$P^d_{i,t}/Q^d_{i,t}$	Active/reactive power demand of bus i at time stage t
a_i, b_i, c	Coefficients of cost function of MT at bus i
$C^{rg}_{i,t}$	Unit cost of the spinning reserve provided by MT at bus i at time t
C^{es}	Unit cost of ESS charge and discharge
$P^{vp}_{i,t}$	Active power output of RG at bus i at time stage t

Variables

$I_{i,t}$	Scheduled complex current from bus i to its parent bus at time t
$V_{i,t}$	Scheduled complex voltage at bus i at time t
$P^{rg}_{i,t}$	Scheduled reserve power of MT at bus i at time t
$P^v_{i,t}$	Scheduled active power output of RG at bus i at time t
$E^{es}_{i,t}$	Scheduled capacity of ESS at bus i at time t
$R^g_{i,t}$	Real-time active power of MT at bus i at time t
$R^g_{i,t}$	Real-time active power of MT at bus i at time t
$\omega_{e,j,t}$	Payment from entity e to entity j at time t
$\Delta P^-_{e,t}/\Delta P^+_{e,t}$	Power shortage/surplus of entity e at time t
$P^g_{i,t}/Q^g_{i,t}$	Scheduled active/reactive power output of MT at bus i at time t
$P^c_{i,t}/P^{dc}_{i,t}$	Scheduled charge/discharge power of ESS at bus i at time t
$P^f_{i,t}/Q^f_{i,t}$	Scheduled active/reactive power flow from bus i to its parent bus at time t
$P^{ex}_{i,j,t}/Q^{ex}_{i,j,t}$	Scheduled active/reactive power exchanged between bus i of entity e and entity $j \in \mathcal{A} \backslash e$ at time t
$P^m_{i,t}/Q^m_{i,t}$	Scheduled active/reactive power exchanged between bus i of DN and the main grid at time t

model with an appropriate trading and payment scheme that encourages bilateral power trading among MGs while protecting their privacy and hedging against uncertainty. For handling uncertainty in networked MGs, stochastic programming and robust optimization have been extensively used in previous studies. A major drawback of stochastic programming is that it typically models uncertainty by a large set of scenarios, which leads to computationally intractable problems [9,10]. Robust optimization describes uncertainty by an uncertainty set that is assumed to contain the possible realizations of uncertain parameters [11]. However, conventional uncertainty sets utilized by robust optimization result in over-conservative solutions with a high cost of robustness. This paper develops a privacy-preserving energy management framework under the P2P trading mechanism to advise MGs on how to externally interact with each other and how to internally schedule their power generation and consumption. The proposed framework consists of two main parts: (1) a data-driven robust (DDR) optimization model coupled with machine learning techniques for determining scheduling and trading decisions that are protected against uncertainty, and (2) a payment bargaining model for the fair allocation of trading benefits.

1.2. Literature review

Since MG systems are subject to various uncertainties arising mainly from the error in load forecasting and the dependence of renewable energy on environmental factors such as wind speed and solar irradiance, uncertainty management has become an active research topic in the scheduling of DN and MGs [12]. There are two common approaches in the literature to hedging against uncertainty: stochastic programming and robust optimization. In stochastic programming, uncertain parameters are modeled as random variables that are assumed to belong to a known probability distribution [13]. In practice, however, the true probability distribution can seldom be accurately estimated because of the insufficiency of historical data [14]. Instead of using continuous distribution functions, scenario-based stochastic programming represents uncertainty by a large set of discrete scenarios generated based on a predefined distribution function, but it generally leads to computationally intractable problems [15]. As an attractive alternative, robust optimization seeks to optimize the worst-case of the possible realizations of uncertain parameters enclosed within an uncertainty set [16]. Robust optimization has gained considerable attention in the area of networked MGs since it overcomes the computational problems caused by stochastic programming and only needs the bounds of uncertain parameters rather than the complete knowledge of their probability distribution [17]. For example, Ref. [18] ensures the robustness of the optimal scheduling scheme against all realizations of the uncertainty of MGs by developing a robust optimization approach in which the lower bounds of renewable outputs and the upper bounds of loads are considered as potential worst-case scenarios. Refs. [19,20] deal with renewable generation uncertainty that affects the scheduling problem of interconnected MGs using robust optimization under the box uncertainty set. This approach makes all uncertain parameters reach their worst-case value, which results in overly conservative scheduling strategies. To handle uncertainties encountered in an energy dispatch problem for multiple MGs, the polyhedral uncertainty set-induced robust optimization is employed by a large number of previous studies [21,22]. The polyhedral set controls the number of uncertain parameters reaching their worst-case values by an adjustable parameter named the uncertainty budget [23]. Traditional uncertainty sets (e.g., box and polyhedral sets) have fixed geometric structures and usually fail to capture the intrinsic characteristics of uncertainty, especially in the case of asymmetric and correlated uncertainty data. The size of these sets needs to be manually tuned by the user so that the possible realizations of uncertain parameters are covered. A common strategy for ensuring full coverage is to scale up the set, but this significantly increases the cost of robustness [24].

Motivated by the need to overcome the shortcomings of stochastic programming and classic robust optimization, DDR optimization has been proposed as a promising approach to improve the quality of robust solutions while avoiding unnecessary conservatism [25]. DDR optimization seeks to unify the ideas that underline robust optimization and stochastic programming. Instead of considering a single probability distribution, DDR optimization employs a family of possible distributions (named the ambiguity set) that is assumed to accommodate the true distribution and then searches for a robust solution with respect to the worst-case distribution within the ambiguity set [26]. The ambiguity sets used in previous studies can be classified into two main types: (1) moment-based and (2) distance-based sets. The first type includes all probability distributions with moments satisfying certain conditions [26], while the second type forms the ambiguity set based on the statistical distance of probability distributions to a nominal distribution measured using metrics such as Wasserstein distance, ϕ-divergence and Prohorov metric [27]. The ambiguity set can be built without complete distributional information, making DRO applicable to the scheduling problems of MGs where precise information on the underlying uncertainty is not available [14]. Ref. [28] models the probability of line outages caused by natural disasters using a moment-based ambiguity set that is constructed based on first-order information extracted from historical data. Applying the Hellinger distance as a ϕ-divergence function, Ref. [29] formulates a chance-constrained energy management problem with a distance-based ambiguity set to capture the uncertainty in the probability distribution of MG loads. In Refs. [4,30,31], the uncertainties of multiple MGs are handled by forming an ambiguity set based on the Wasserstein distance that measures the difference between empirical and actual distributions associated with uncertain parameters using the 1- and ∞-norms.

Robust energy management models for networked MGs follow two main architectures, namely, centralized and distributed [32]. In the centralized architecture, the generation and load information of all the system entities is collected and processed in a central computation center to solve a single optimization problem and determine unified dispatch schemes and scheduling decisions. Ref. [33] formulates a two-stage robust optimization problem for the optimal collaborative operation of multiple interconnected MGs with the aim of obtaining the optimal dispatch strategy that minimizes the operational cost of each MG under uncertainty in PV power generation. The proposed two-stage model is decomposed into master and sub-problems, which are iteratively solved using the column-and-constraint generation (CCG) algorithm. A risk-constrained robust optimization model is proposed in Ref. [34], enhancing the resiliency of scheduling decisions against extreme weather conditions and possible errors made by the central controller of MGs. Ref. [35] presents a robust model to find the optimal daily operation and system expansion planning strategy for DN with several MGs while taking into account multiple renewable energy resources, power production contingency, and demand response. Ref. [36] develops a hybrid robust-stochastic model for the scheduling of networked MGs in grid-connected and islanding modes, in which stochastic programming is used to deal with uncertainty in renewable generation, load, and time of disconnection from the upstream grid, and robust optimization is used to model the worst-case realization of market prices. Ref. [19] formulates a robust game theory optimization model to optimize the economic performance of multiple MGs while suppressing the impact of variations in renewable energy generation on power transactions between them. The proposed model is formulated as a two-stage scheduling problem and then is solved by the CCG algorithm.

In the distributed architecture, on the other hand, the information storage and computation process are performed in a distributed manner and the main problem is decomposed into several sub-problems with less computational burden, each of which is related to one of the system entities including DN and MGs [37]. Considering the benefit of each entity, this architecture coordinates the operation of the whole system without having to share sensitive information between the autonomous entities, which provides higher robustness to communication failures and greater privacy protection [38]. Ref. [39] develops an energy management framework for coordination among several MGs connected to DN, which is designed in a distributed way with the help of the alternating direction method of multipliers (ADMM) algorithm. A robust model is suggested in [40] to optimize the economic performance of grid-connected private MGs under the uncertainty of renewable generation and electricity prices. The ADMM algorithm is adopted to decompose the developed model and reduce its computational complexity. Ref. [41] presents a robust scheduling problem to manage P2P energy transactions between multiple interconnected MGs with a decentralized pricing strategy based on ADMM. Ref. [42] develops a robust economic dispatch framework for collaboration between multiple MGs and their upstream network, which is solved in a distributed way by ADMM. In [43], the authors first use the auxiliary problem principle (APP) to reformulate the centralized energy management problem of multiple AC–DC MGs for a distributed implementation and then adopt ADMM to solve the resulting problem. Aiming at coordinating the operation of DN and MGs, Ref. [18] develops two types of decentralized methodologies based on the independent CCG and combined ATC and CCG algorithms. The results show that the first methodology can obtain a faster convergence speed, while the second one provides lower communication burden and higher privacy protection.

A lot of effort has been dedicated to designing efficient P2P power trading mechanisms for networked MGs. Game theory is one of the main approaches to model the cooperative and competitive behavior of MGs that act as autonomous entities in the power trading market. Non-cooperative games are applied when different stakeholders with unequal power and conflicting interests are in charge of different parts of the system and independently make decisions to gain more profit. Ref. [44] develops a leader-follower interactive framework based on the Stackelberg game to determine the scheduling decisions of DN with multiple MGs, where the DN operator as the leader minimizes the operation cost by determining the optimal quantity and price of power exchange with each MG, and the MG operators as the followers respond to the price signals by optimizing power trading with the DN operator in a way that the constraints of their inner device operation are satisfied. Using a market-clearing mechanism, a bi-level robust optimization model is proposed in [45] for managing energy interactions between DN and MGs under uncertainties caused by renewable energy resources. An economic dispatch problem for DN is formulated at the upper level, by which locational marginal price signals are derived and communicated to MGs at the lower level to optimize their dispatch strategy based on the received signals. The corresponding optimization models in both levels are formulated as two-stage robust optimization problems and solved using the CCG algorithm. Cooperative games are used when a group of MGs agrees to coordinate their strategies because the expected profit is greater under cooperation than under competition. Ref. [46] designs a cooperative game-based robust planning model using the Nash bargaining solution to determine power trading and payments for DN and MGs. The robust model is formulated under the polyhedral uncertainty set and is solved by a bi-level iterative framework. In [47], an incentive P2P trading mechanism assisted with robust optimization is developed to encourage MGs to actively trade power and ensure fair benefit-sharing between MGs through a Nash bargaining payment problem.

Based on the literature review conducted above, the following research gaps are identified.

(1) Previous studies on power trading management have mainly focused on transaction mechanisms that are optimal only for a pre-defined set of power generation and consumption data and thus may become biased toward either buyers or sellers with a small change in input data. In a flexible and true P2P market, however, it is natural for MGs to dynamically adjust their power-sharing profiles and associated payments according to the realized value of uncertainty data.

(2) Most of the existing robust optimization models for handling uncertainty in networked MGs adopt fixed-shape uncertainty sets, such as those proposed in Refs. [46,47]. A major drawback of such sets is that they are constructed without incorporating the geometric structure of uncertainty data, yielding over-conservative solutions with a high cost of robustness. In some previous studies, DDR optimization models under the moment-based and distance-based uncertainty sets are utilized to cope with the uncertainties affecting the performance of MGs. Such models need rich historical data free of contamination to ensure the high reliability of robust solutions. Moreover, the moment-based sets cannot guarantee that the ambiguity set converges to the true distribution [31], and the distance-based sets require nontrivial assumptions on the distance metrics [48].

(3) To take into account the sequential decision-making process in the energy management of MGs, a number of studies have developed adjustable robust optimization models that incorporate recourse decisions with adaptability to the realization of uncertainty [19,49]. The adjustable robust models are usually formulated as multi-level min–max–min optimization problems that cannot be directly handled by any commercial solvers and need to be solved by decomposition algorithms such as the CCG and Benders decomposition.

(4) To coordinate the energy management problems of individual MGs, the ADMM algorithm is commonly used in previous studies [39,40]. However, the computational efficiency of this algorithm in terms of optimality gap and convergence rate degrades considerably as the complexity of the networks and the number of MGs connected to DN increase. The problem becomes more severe for DDR optimization models that are generally more computationally intensive than robust models with fixed-shape sets. In addition, ADMM-based distributed energy management models can only be applied to convex problems since there is no guarantee they converge to the optimal solution for non-convex problems [50].

1.3. Contributions

This paper proposes a data-driven power scheduling and trading framework that addresses the main issues in the planning of networked MGs: (1) the design of power trading mechanisms (transactions and payments), (2) the treatment of uncertainty, and (3) the protection of privacy. A Nash bargaining game-based energy management model is designed to optimize internal power scheduling within MGs as well as external power trading and related payments between MGs. The uncertainties associated with renewable generation are handled by developing an adaptive DDR optimization approach under a kernel-based uncertainty set. The proposed DDR model uses the robust kernel density estimation (RKDE) as a machine learning technique to form an uncertainty set by extracting statistical information from the power output data of WTs and PVs. The resulting uncertainty set is protected against contamination of the dataset with anomalous measurements. To preserve the privacy of individual MGs and alleviate the computational burden of centralized optimization, the proposed model is solved in a distributed manner using the ADMM and the augmented Lagrange-based alternating direction inexact Newton (ALADIN) algorithms. ADMM has found widespread application as a straightforward and powerful tool for decomposing complex energy management problems into simpler sub-problems [51], but its performance in finding a solution with satisfactory accuracy declines as the size of optimization problems grows [52]. Thanks to the favorable properties of ALADIN in terms of optimality gap and convergence speed [53], the distributed model developed in this paper maintains its computational tractability when the number of MGs in the network increases. A comparison of the proposed model with other existing robust optimization models for networked MGs can be found in Table 1. The main contributions of this work are summarized as follows:

(1) By combining the Nash bargaining theory with DDR optimization, a data-driven P2P trading mechanism is designed for networked MGs to ensure the robustness of power trading and scheduling decisions against uncertainty and fairy allocate the benefits of trading among MGs. The proposed mechanism leverages the power of kernel smoothing methods to accurately extract statistical information from uncertainty data and provide a safeguard against those uncertainty realizations that are realistic in practice rather than all possible realizations. The result is that MGs can earn a higher profit by sacrificing less optimality for the robustness.

(2) While previous game-theoretic energy management models use fixed-shape uncertainty sets with a high robustness cost, this paper constructs a data-driven uncertainty set based on the RKDE [54,55], whose strengths lie in the following three aspects. First, the data-driven set can adjust itself to the intrinsic complexity and structure of uncertainty data, thus avoiding the costs of handling uncertain parameters that are unlikely to be realized. Second, it provides reliable power trading strategies even when the power output data of PVs and WTs are large in scale or contaminated by noise. Third, uncertainty data is automatically incorporated into the robust optimization process through a nonparametric interface, while the type of fixed-shape sets must be endogenously selected by the modeler on the basis of historical data. To the best of our knowledge, this is the first time that the RKDE-based uncertainty set is used to deal with the uncertainty of multiple MGs.

(3) Another advantage of the proposed robust model is its adaptability, which allows recourse decisions to adjust after observing the realizations of uncertain parameters. Different from previous adaptive optimization models for networked MGs that are formulated as nonconvex min–max–min problems and need

Table 1
Comparison of the proposed model with previous robust optimization models for networked MGs.

Refs	System architecture		Robust optimization		Data-driven robust optimization			Distributed optimization method				Game-theoretic trading mechanism
	Centralized	Distributed	Box set	Polyhedral set	Moment-based set	Distance-based set	RKDE-based set	ADMM	APP	ATC	ALADIN	
[44]	*		*									*
[49]		*		*				*				
[20]	*		*									
[19]	*		*									
[57]		*	*					*				
[58]	*			*								
[59]	*			*								
[39]		*		*				*				
[21]	*			*								
[22]	*			*								
[33]	*			*								
[44]												
[34]	*			*								
[18]		*		*						*		
[40]		*		*				*				
[43]		*		*				*	*			
[41]		*		*				*				
[45]	*			*								*
[30]	*			*								
[47]		*		*				*				*
[46]		*		*				*		*		*
[36]	*			*								
[28]	*				*							
[4]		*				*		*				
[29]		*				*		*				
[42]		*				*		*				
[31]		*				*		*				
This work		*					*	*			*	*

to be tackled by decomposition algorithms, this paper adopts affine decision rules to provide a convex adaptive robust counterpart that can be directly solved using off-the-shelf optimization solvers.

(4) Unlike most prior studies that only use ADMM to recast the centralized optimization problem as a distributed one, this paper utilizes ADMM and ALADIN to improve the applicability of the proposed model to both small and large-scale networks. With more information shared between the sub-problems, ALADIN is applicable to nonconvex problems and generally converges to the optimal solution more quickly than ADMM. Although ALADIN has shown significant improvements in the computational performance of distributed optimization problems [50,56], it has not been previously used for the distributed energy management problem of multiple MGs.

The remainder of the paper is structured as follows: The general structure of the DN connected with multiple MGs is described in Section 2. In Section 3, the mathematical formulations of the deterministic and data-driven scheduling and trading problems are presented. The two kinds of solution algorithms for the distributed coordination of DN and MGs are introduced in Section 4. Case studies evaluating the proposed model are reported in Section 5. This is followed by Section 6, which concludes the paper with the main findings.

2. Model description

As displayed in Fig. 1, DN is connected to the high voltage (HV) system and at the same time connected to multiple MGs, being able to exchange information and power through communication networks and power grids, respectively. Both DN and MGs incorporate dispatchable distributed generators (DGs), renewable energy generators (RGs), energy storage systems (ESS), and internal loads. Because of different energy generation and consumption profiles, some MGs may be in need of power while others may have excess power. MGs trade power with each other and with DN to improve their own performance and enhance the reliability of the whole system. Taking into account the power transactions between the MGs, the DN operator compensates for the difference between production and consumption in the distribution system by exchanging power with the main grid.

The proposed framework for ensuring optimal power scheduling and trading in the system described above consists of four main parts. The first part formulates a deterministic scheduling problem within each MG and DN while ignoring uncertainty by fixing the outputs of RGs at their nominal values. The optimal decisions obtained by deterministic optimization models are unreliable for practical applications because they may be rendered sub-optimal or even infeasible by small fluctuations in renewable sources. In order to improve the ability to tolerate the fluctuations that may influence the normal operation of the involved entities, data-driven robust scheduling models are developed in the second part to determine robust scheduling schemes that remain feasible for all possible realizations of the renewable power outputs within their corresponding uncertainty set. While fixed-shape uncertainty sets fail to fit neatly over historical realizations, the proposed robust model uses machine learning techniques to build a convex uncertainty set that adapts itself to the structure of the realizations. Since uncertainty data is sequentially revealed

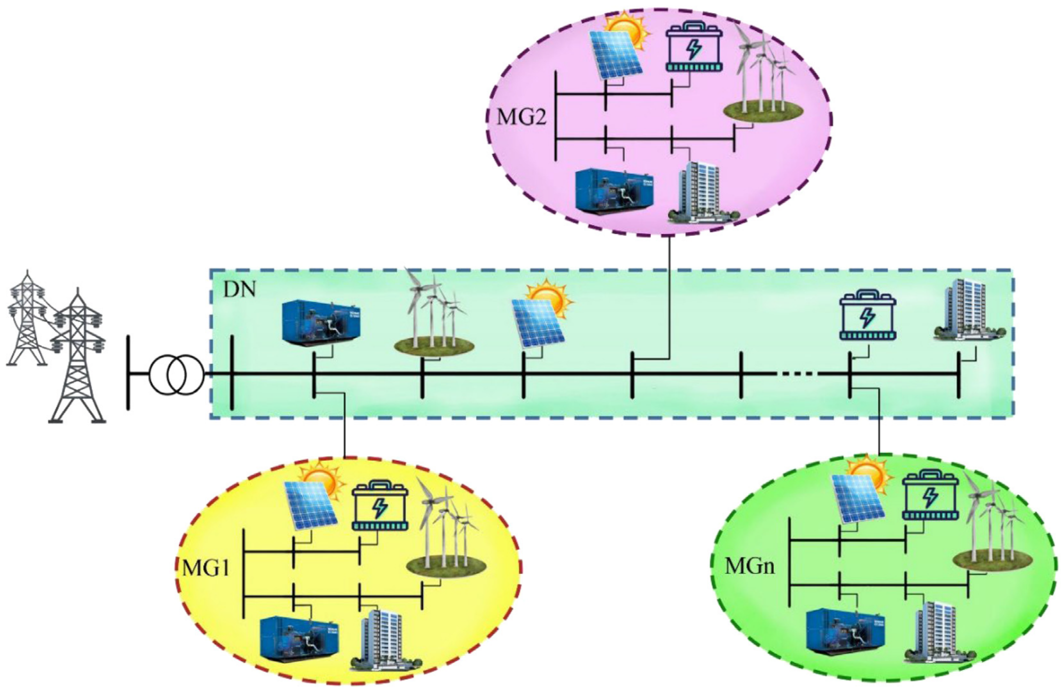

Fig. 1. The structure of DN connected with multiple MGs.

over the scheduling horizon, the model incorporates recourse decisions so that operators can adjust decisions according to the realized values of renewable power generation. Under this setting, decision variables are categorized into two classes: the first-stage decisions that have to be made here-and-now and cannot be changed over time, and the second-stage decisions that are handled in a wait-and-see manner and can flexibly be adjusted after observing actual values taken by uncertain parameters. In the proposed model, the first-stage decisions include the day-ahead scheduling decisions of power generation and spinning reserve of the dispatchable generators, charging and discharging power of the ESSs and power trading between the entities. Adjusting the output of the dispatchable generators in real-time operation is set as the second-stage decision, which aims to minimize system imbalances that might occur due to possible variations in renewable production. In the third part, the data-driven robust model is recast as a bi-level scheduling and trading payment problem. The upper level determines robust scheduling and trading decisions, and the lower level uses a game-based bargaining model to fairly allocate trading benefits to the entities based on the quantities of bilateral power transactions obtained from the upper level. The fourth part designs distributed solution methods based on the ADMM and ALADIN algorithms to decompose and solve the optimization problems of the upper and lower levels in a privacy-preserving manner. An overview of the proposed framework is shown in Fig. 2.

3. Problem formulation

The deterministic scheduling models of DN and MG are first presented in this section, followed by formulating their corresponding data-driven adjustable robust scheduling and trading models that address the uncertainty of renewable energy resources.

3.1. Deterministic scheduling model

3.1.1. Cost function

MG cost function: The objective of each MG is to minimize its own operation cost while satisfying all constraints by keeping them within their safe range. The objective function of the day-ahead scheduling problem for MG $e \in \mathcal{M}$ can be formulated as follows:

$$\min f_e\left(\boldsymbol{P}_e^{g,M}, \boldsymbol{P}_e^{rg,M}, \boldsymbol{P}_e^{c,M}, \boldsymbol{P}_e^{dc,M}, \boldsymbol{\omega}_e^{M}\right) = f_e^{g}\left(\boldsymbol{P}_e^{g,M}, \boldsymbol{P}_e^{rg,M}\right) + f_e^{es}\left(\boldsymbol{P}_e^{c,M}, \boldsymbol{P}_e^{dc,M}\right) + f_e^{ex}\left(\boldsymbol{\omega}_e^{M}\right) \quad \forall e \in \mathcal{M} \tag{1}$$

where f_e^g, f_e^{es} and f_e^{ex} are the generation and reserve cost of MTs, the operation cost of ESSs and the total trading payment, respectively. The generation cost of MTs is modeled as a function of their real power output with the help of a quadratic polynomial approximation. Therefore, f_e^g, f_e^{es} and f_e^{ex} can be expressed as follows:

$$f_e^{g}\left(\boldsymbol{P}_e^{g,M}, \boldsymbol{P}_e^{rg,M}\right) = \sum_{t\in T}\left[\sum_{i\in\mathcal{N}_e^m}\left(a_i P_{i,t}^{g,M^2} + b_i P_{i,t}^{g,M} + c\right) + \sum_{i\in\mathcal{N}_e^m} C_{i,t}^{rg,M} P_{i,t}^{rg,M}\right] \tag{2}$$

$$f_e^{es}\left(\boldsymbol{P}_e^{c,M}, \boldsymbol{P}_e^{dc,M}\right) = \sum_{t\in T}\sum_{i\in\mathcal{N}_e^{es}} C^{es}\left(P_{i,t}^{c,M} + P_{i,t}^{dc,M}\right) \tag{3}$$

$$f_e^{ex}\left(\boldsymbol{\omega}_e^{M}\right) = \sum_{t\in T}\sum_{j\in\mathcal{A}\setminus e} \omega_{e,j,t}^{M} \tag{4}$$

DN cost function: In addition to power trading with MGs, DN can sell/buy power to/from the main grid. In comparison with the cost function of MG, the cost of power transaction with the main grid is an extra cost that needs to be included in the objective function of DN as follows:

$$\min f_e\left(\boldsymbol{P}_e^{g,D}, \boldsymbol{P}_e^{rg,D}, \boldsymbol{P}_e^{c,D}, \boldsymbol{P}_e^{dc,D}, \boldsymbol{\omega}_e^{D}\right) = f_e^{g}\left(\boldsymbol{P}_e^{g,D}, \boldsymbol{P}_e^{rg,D}\right) + f_e^{es}\left(\boldsymbol{P}_e^{c,D}, \boldsymbol{P}_e^{dc,D}\right) + f_e^{m}\left(\boldsymbol{P}_e^{m,D}\right) + f_e^{ex}\left(\boldsymbol{\omega}_e^{D}\right) \forall e \in \mathcal{D} \tag{5}$$

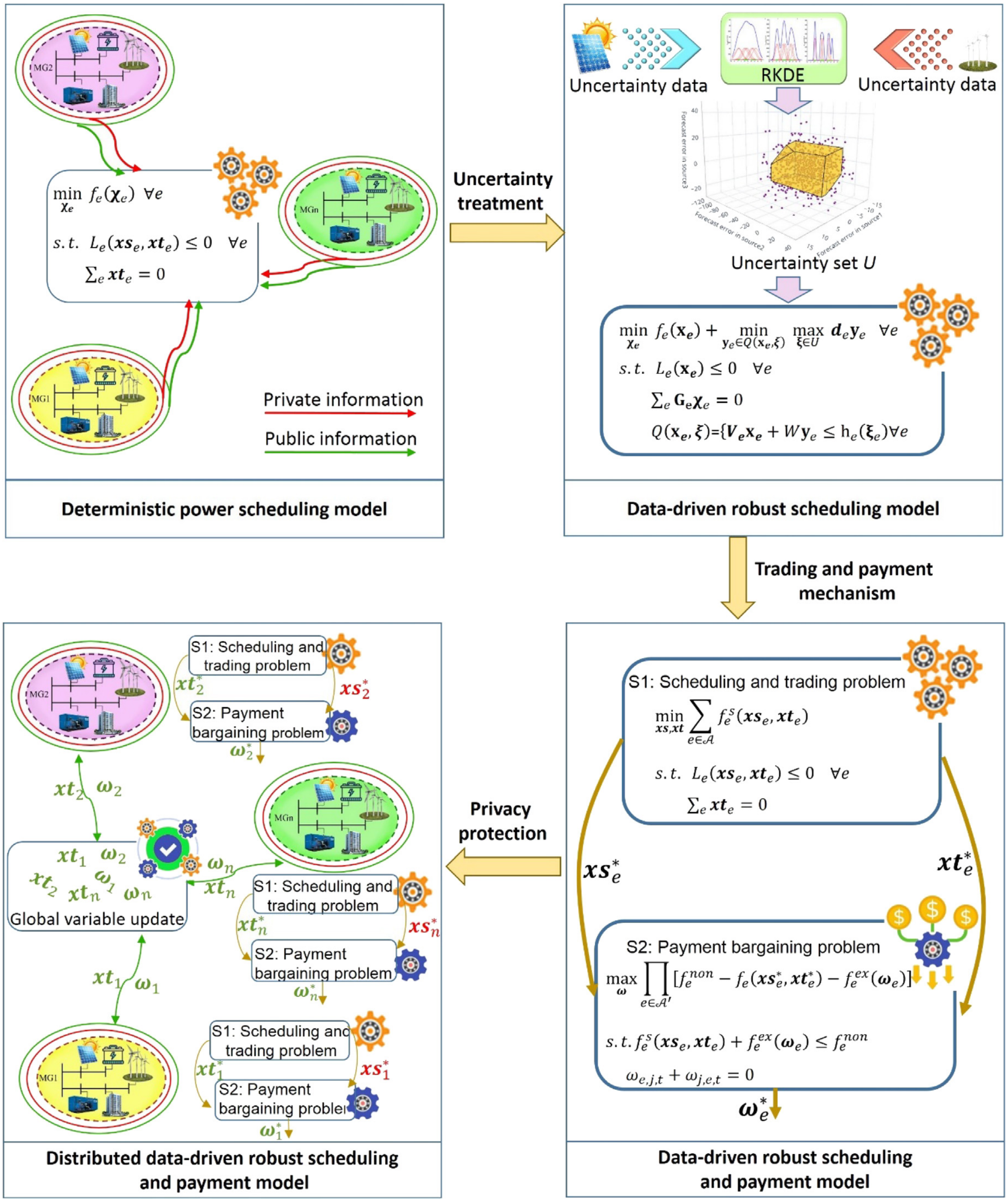

Fig. 2. An overview of the proposed data-driven robust scheduling and trading framework.

where f_e^g, f_e^{es}, and f_e^{ex} have the same definition as in the cost function of MG, and f_e^m is given as:

$$f_e^m\left(\boldsymbol{P}_e^{m,D}\right)=\sum_{t\in T}\sum_{i\in\mathcal{N}^m}\left(\vartheta_t^b\left[P_{i,t}^{m,D}\right]^+-\vartheta_t^s\left[-P_{i,t}^{m,D}\right]^+\right) \tag{6}$$

Since the selling price ϑ_t^s is generally lower than the purchase price ϑ_t^b, Eq. (6) can be linearized by replacing the ramp function $[x]^+=\max(x,0)$ with auxiliary variable $P_{i,t}^{m'}$ as follows:

$$f_e^m\left(\boldsymbol{P}_e^{m,D}\right)=\sum_{t\in T}\sum_{i\in\mathcal{N}^m}P_{i,t}^{m,D'} \tag{7}$$

$$P_{i,t}^{m,D'}\geq\vartheta_t^bP_{i,t}^{m,D},\ P_{i,t}^{m,D'}\geq\vartheta_t^sP_{i,t}^{m,D}\forall t\in T,i\in\mathcal{N}^m \tag{8}$$

3.1.2. Constraints

MG power flow constraints: The MG considered in this paper is modeled as a radial network $\mathcal{G}(\mathcal{N},\mathcal{L})$ comprised of a set of nodes $\mathcal{N}$ connected by distribution lines $\mathcal{L}$. The nodes are indexed by $i=0,1,\ldots,n$, where 0 indicates the root node and others indicate branch nodes. Each node i has one parent node $\pi(i)$ (that is closer to the root node) and a set of child nodes $\delta(i)$. Assuming that $S_i=P_i+\mathbf{i}Q_i$ is the complex power that flows from node i to its parent node $\pi(i)$, the AC power flow equations can be formulated as follows [60]:

$$P_{i,t}^{g,M} + P_{i,t}^{v,M} + P_{i,t}^{dc,M} - P_{i,t}^{c,M} - P_{i,t}^{d,M} - \sum_{i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e} P_{i,j,t}^{ex,M} = P_{i,t}^{f,M} + r_i^M l_i^M - \sum_{k \in \delta(i)} P_{k,t}^{f,M} \forall t \in T, i \in \mathcal{N}_e^a \tag{9}$$

$$Q_{i,t}^{g,M} - Q_{i,t}^{d,M} - \sum_{i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e} Q_{i,j,t}^{ex,M} = Q_{i,t}^{f,M} + x_i l_{i,t}^M - \sum_{k \in \delta(i)} Q_{k,t}^{f,M} \forall t \in T, i \in \mathcal{N}_e^a \tag{10}$$

$$\upsilon_{i,t}^M = \upsilon_{\pi(i),t}^D - \left(r_i^2 + x_i^2\right) l_{i,t}^M + 2\left(r_i P_{i,t}^{f,M} + x_i Q_{i,t}^{f,M}\right) \forall t \in T, i \in \mathcal{N}_e^a \tag{11}$$

$$l_{i,t}^M = \frac{\left(P_{i,t}^{f,M}\right)^2 + \left(Q_{i,t}^{f,M}\right)^2}{\upsilon_{i,t}^M} \forall t \in T, i \in \mathcal{N}_e^a \tag{12}$$

where $\upsilon_{i,t}$ and $l_{i,t}$ represent the squared voltage and current, respectively, which are defined as $\upsilon_{i,t} = \left|V_{i,t}\right|^2$ and $l_{i,t} = \left|I_{i,t}\right|^2$. To improve the computational tractability of the model, the second-order cone (SOC) relaxation is adopted to convexify the Eq. (12) as follows:

$$\left\| \begin{matrix} 2P_{i,t}^{f,M} \\ 2Q_{i,t}^{f,M} \\ l_{i,t}^M - \upsilon_{i,t}^M \end{matrix} \right\| \le \upsilon_{i,t}^M + l_{i,t}^M \Longleftrightarrow l_{i,t}^M \ge \frac{\left(P_{i,t}^{f,M}\right)^2 + \left(Q_{i,t}^{f,M}\right)^2}{\upsilon_{i,t}^M} \forall t, i \in \mathcal{N}_e^a \tag{13}$$

To guarantee the safe operation of MGs, constraint (14) bounds the voltage magnitude of all buses within a prescribed level of a nominal voltage. Moreover, constraint (15) takes into account the current limit during the operating cycle.

$$\underline{V}^2 \le \upsilon_{i,t}^M \le \overline{V}^2 \quad \forall t \in T, i \in \mathcal{N}_e^a \tag{14}$$

$$l_{i,t}^M \le \overline{I}^2 \quad \forall t \in T, i \in \mathcal{N}_e^a \tag{15}$$

DN power flow constraints: Since DN is directly connected to the main grid, its power balance constraint includes power transactions with the main grid. Therefore, the active and reactive power balance constraints are rewritten as constraints (16) and (17), while constraints on current and voltage remain the same. The range of power transactions between DN and the main grid is limited by constraints (18) and (19).

$$P_{i,t}^{m,D} + P_{i,t}^{g,D} + P_{i,t}^{v,D} + P_{i,t}^{dc,D} - P_{i,t}^{c,D} - P_{i,t}^{d,D} - \sum_{i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e} P_{i,j,t}^{ex,D} = P_{i,t}^{f,D} + r_i l_i^D - \sum_{k \in \delta(i)} P_{k,t}^{f,D} \quad \forall t \in T, i \in \mathcal{N}_e^a \tag{16}$$

$$Q_{i,t}^{m,D} + Q_{i,t}^{g,D} - Q_{i,t}^{d,D} - \sum_{i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e} Q_{i,j,t}^{ex,D} = Q_{i,t}^{f,D} + x_i l_{i,t}^D - \sum_{k \in \delta(i)} Q_{k,t}^{f,D} \quad \forall t \in T, i \in \mathcal{N}_e^a \tag{17}$$

$$\underline{P^m} \le P_{i,t}^{m,D} \le \overline{P^m} \quad \forall t \in T, i \in \mathcal{N}_e^a \tag{18}$$

$$\underline{Q^m} \le Q_{i,t}^{m,D} \le \overline{Q^m} \quad \forall t \in T, i \in \mathcal{N}_e^a \tag{19}$$

ESS constraints: The maximum charging and discharging power of each ESS are limited by constraints (20) and (21), respectively. The state of charge is ensured by constraint (22), which states that the power stored in ESSs at each period is equal to the power remaining at the end of the previous period plus the amount of charging or discharging power in the current period considering its efficiency rate. It is proved that the minimization of the operating cost of ESSs included in the objective function avoids simultaneous charging and discharging [45,61]. The proof of this property is included in Appendix to make this paper self-contained. The upper and lower bound of ESS capacity are expressed by constraint (23). The index of DN and MGs is omitted for notation simplicity.

$$0 \le P_{i,t}^c \le \overline{P}^c \quad \forall t \in T, i \in \mathcal{N}_e^{es} \tag{20}$$

$$0 \le P_{i,t}^{dc} \le \overline{P}^{dc} \quad \forall t \in T, i \in \mathcal{N}_e^{es} \tag{21}$$

$$E_{i,t}^{es} = E_{i,t-1}^{es} + P_{i,t}^c \eta^c - P_{i,t}^{dc} / \eta^{dc} \quad \forall t \in T, i \in \mathcal{N}_e^{es} \tag{22}$$

$$\underline{E}^{es} \le E_{i,t}^{es} \le \overline{E}^{es} \quad \forall t \in T, i \in \mathcal{N}_e^{es} \tag{23}$$

Power generation constraints: Constraint (24) states that the sum of the scheduled power of each MT with its respective reserve cannot exceed the maximum generation capacity. The scheduled power of each RG should not be higher than the forecasted output, just as enforced in constraint (25).

$$0 \le P_{i,t}^g + P_{i,t}^{rg} \le \overline{P}^g \quad \forall t \in T, i \in \mathcal{N}_e^m \tag{24}$$

$$0 \le P_{i,t}^v \le \hat{P}_{i,t}^{vp} \quad \forall t \in T, i \in \mathcal{N}_e^r \tag{25}$$

Power trading/payment constraints: P2P power transactions between different entities (MGs and DN) are represented by constraints (26) and (27), where $P_{i,j,t}^{ex}$ and $Q_{i,j,t}^{ex}$ are active/reactive power exchanged between bus i of entity e to other entity j. If entity e buys power from entity j at time t, then $P_{i,j,t}^{ex} < 0$; if entity e sells power to entity j at time t, then $P_{i,j,t}^{ex} > 0$. The payment among the entities is subject to the market-clearing constraint (28). If entity e makes payment to entity j at time t, then $\omega_{e,j,t} > 0$; otherwise entity e receives payment from entity j and $\omega_{e,j,t} < 0$. The allowable range for power transactions between the entities is defined by constraints (29) and (30).

$$P_{i,j,t}^{ex} + P_{j,i,t}^{ex} = 0 \quad \forall t \in T, i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e \tag{26}$$

$$Q_{i,j,t}^{ex} + Q_{j,i,t}^{ex} = 0 \quad \forall t \in T, i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e \tag{27}$$

$$\omega_{e,j,t} + \omega_{j,e,t} = 0 \quad \forall t \in T, e \in \mathcal{A}, j \in \mathcal{A} \setminus e \tag{28}$$

$$\underline{P^{ex}} \le P_{i,j,t}^{ex} \le \overline{P^{ex}} \quad \forall t \in T, i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e \tag{29}$$

$$\underline{Q^{ex}} \le Q_{i,j,t}^{ex} \le \overline{Q^{ex}} \quad \forall t \in T, i \in \mathcal{N}_e^c, j \in \mathcal{A} \setminus e \tag{30}$$

3.2. Data-driven robust scheduling model

In this section, the data-driven adjustable robust optimization model for energy scheduling of DN and MGs is developed. As stated before, the optimal solution of deterministic models is not immunized against infeasibility because they consider only one possible realization of renewable generation. To keep the optimal power schedule of all entities feasible in the face of possible perturbations of uncertain parameters, a two-stage robust optimization model that ensures protection against all realizations within a given uncertainty set can be formulated as follows:

$$\min_{\boldsymbol{x}} f_e(\boldsymbol{x}_e) + \max_{P^{vp} \in U} \min_{\boldsymbol{y} \in \Omega(x,U)} \sum_{t \in T} \left[\rho_{e,t}^- \Delta P_{e,t}^- + \rho_{e,t}^+ \Delta P_{e,t}^+\right] \quad \forall e \in \mathcal{A} \tag{31}$$

$s.t.$(9)–(11), (13)–(30)

$0 \le \Delta P_{e,t}^+, \Delta P_{e,t}^-$

where x represents the first-stage decisions that are the same as the decision variables defined for the deterministic model, including the day-ahead scheduling of the MTs, ESSs and power exchange with DN. $\boldsymbol{y_e} = \left[\Delta P_e^-, \Delta P_e^+, R_e^g\right]$ represents the second-stage (recourse) decisions that include the power shortage/surplus of each entity and the output of its MTs during real-time operation. While the first-stage decisions need to be made at the beginning before the uncertainty is uncovered, the second-stage decisions are adjustable based on the realized power output of RGs with respect to the data-driven uncertainty set U. Under the worst-case realization, optimization problem (31) finds a robust solution that minimizes the operating cost as well as the imbalance cost calculated by penalizing power shortages and surpluses resulting from variations in renewable generation. It is worth noting that the penalty cost of the power shortage is higher than that of the power surplus. The feasible region $\Omega(x, U)$ is expressed by the following constraints:

$$\sum_{i\in\mathcal{N}_e^{mt}} R_{i,t}^g + \sum_{i\in\mathcal{N}_e^r} \tilde{P}_{i,t}^{vp} + \sum_{i\in\mathcal{N}_e^{es}} \left(P_{i,t}^{dc} - P_{i,t}^c\right) - \sum_{i\in\mathcal{N}_e^a} P_{i,t}^d - \sum_{i\in\mathcal{N}_e^c, j\in\mathcal{A}\backslash e} P_{i,j,t}^{ex} \le \Delta P_{e,t}^+ - \Delta P_{e,t}^- \quad \forall t\in T, e\in\mathcal{A}\backslash\mathcal{D} \tag{32}$$

$$\sum_{i\in\mathcal{N}^m} P_{i,t}^m + \sum_{i\in\mathcal{N}_e^{mt}} R_{i,t}^g + \sum_{i\in\mathcal{N}_e^r} \tilde{P}_{i,t}^{vp} + \sum_{i\in\mathcal{N}_e^{es}} \left(P_{i,t}^{dc} - P_{i,t}^c\right) - \sum_{i\in\mathcal{N}_e^a} P_{i,t}^d - \sum_{i\in\mathcal{N}_e^c, j\in\mathcal{A}\backslash e} P_{i,j,t}^{ex} \le \Delta P_{e,t}^+ - \Delta P_{e,t}^- \quad \forall t\in T, e\in\mathcal{A}\backslash\mathcal{M} \tag{33}$$

$$0 \le R_{i,t}^g \le P_{i,t}^g + P_{i,t}^{rg} \quad \forall t\in T, i\in\mathcal{N}_e^{mt}, e\in\mathcal{A} \tag{34}$$

Constraints (32) and (33) are the system imbalance constraints that determine power shortages or surpluses caused by the difference between the first stage day-ahead scheduling and the second stage real-time operation. Constraint (34) states that the real-time power output of the MTs should not exceed their day-ahead scheduled generation and spinning reserve.

3.2.1. Uncertainty set construction

This section first presents the RKDE as an efficient approach for extracting probability distribution information embodied in the uncertainty data of renewable generation. This density estimation approach exhibits great robustness against contamination of input training samples and mitigates estimation errors [55]. The extracted distributional information is then utilized to form the uncertainty set U that accurately accommodates the region where data samples reside.

3.2.1.1. Robust kernel density estimation (RKDE). Let us consider a set of N data samples $\left\{\xi^{(i)}\right\}_{i=1}^N \in \mathbb{R}^d$ from an unknown distribution with a density function f, where $\boldsymbol{\xi}$ represents the vector of uncertain parameters. The kernel density estimate (KDE) of f is given by

$$\hat{f}_{KDE}(\xi) = \frac{1}{N}\sum_{i=1}^N K_\sigma\left(\xi, \xi^{(i)}\right) \tag{35}$$

where K_σ is a non-negative kernel function with smoothing parameter σ named the bandwidth. The Gaussian kernel is a popular choice for the kernel function with well-known theoretical and practical features, expressed as:

$$K_\sigma\left(\xi, \xi^{(i)}\right) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^d exp\left(-\frac{\left\|\xi - \xi^{(i)}\right\|^2}{2\sigma^2}\right) \tag{36}$$

For any positive semi-definite kernel K_σ, there is a mapping $\Phi(\xi) : \mathbb{R} \mapsto \mathbb{H}$ such that $K_\sigma\left(\xi, \xi^{(i)}\right) = \left\langle\Phi(\xi), \Phi\left(\xi^{(i)}\right)\right\rangle$, where $\mathbb{H}$ is a high-dimensional Hilbert space of functions. It can be shown that $\hat{f}_{KDE}$ is equivalent to the solution of the following problem [62]:

$$\min_{h\in\mathbb{H}} \sum_{i=1}^N \left\|\Phi\left(\xi^{(i)}\right) - h\right\|^2 \tag{37}$$

where h represents an arbitrary function in $\mathbb{H}$. The use of the quadratic loss function in (37) makes the KDE suffer from high sensitivity to the presence of outliers. The RKDE is developed as a weighted version of the KDE to enhance robustness in dealing with outlier-contaminated data by assigning smaller weights to outlying data points. The RKDE can be obtained by solving the following problem:

$$\min_{h\in\mathbb{H}} L(h) = \sum_{i=1}^N \gamma\left(\left\|\Phi\left(\xi^{(i)}\right) - h\right\|\right) \tag{38}$$

where $\gamma(\cdot)$ represent a robust loss function that down-weights the outliers, such as the Hampel and Huber loss functions [63,64]. Since problem (38) has no closed-form solution, it cannot directly be solved by available optimization software. Fortunately, $\hat{f}_{KDE}$ can efficiently be computed by the kernelized iteratively reweighted least squares (KIRWLS), which is implemented as an iterative algorithm [62]. The algorithm starts with a set of initial weights $\mathrm{w}_i^{(0)} \ge 0$ such that $\sum_{i=1}^n \mathrm{w}_i^{(0)} = 1$. The value of $\mathrm{w}_i^{(r)}$ is updated in each iteration r by the following equations to produce a sequence of $h^{(r)}$. This procedure continues until the improvement in $L\left(h^{(r)}\right)$ becomes smaller than a predefined tolerance.

$$h^{(r)} = \sum_{i=1}^N \mathrm{w}_i^{(r-1)}\Phi\left(\xi^{(i)}\right) \tag{39}$$

$$\mathrm{w}_i^{(r)} = \frac{\eta\left(\left\|\Phi\left(\xi^{(i)}\right) - h^{(r)}\right\|\right)}{\sum_{j=1}^N \eta\left(\left\|\Phi\left(\xi^{(j)}\right) - h^{(r)}\right\|\right)} \tag{40}$$

In the above equation, η is the first-order derivative of the loss function γ, and $\left\|\Phi\left(\xi^{(i)}\right) - h^{(r)}\right\|$ is calculated as follows:

$$\left\|\Phi\left(\xi^{(j)}\right) - h^{(r)}\right\| = \sqrt{\left\langle\Phi\left(\xi^{(j)}\right), \Phi\left(\xi^{(j)}\right)\right\rangle - 2\left\langle\Phi\left(\xi^{(j)}\right), h^{(r)}\right\rangle + \left\langle h^{(r)}, h^{(r)}\right\rangle} \tag{41}$$

where the computation of the mapping Φ that is not explicitly known can be done by the kernel trick:

$$\left\langle\Phi\left(\xi^{(j)}\right), \Phi\left(\xi^{(j)}\right)\right\rangle = K_\sigma\left(\xi^{(j)}, \xi^{(j)}\right) \tag{42}$$

$$\left\langle\Phi\left(\xi^{(j)}\right), \xi^{(r)}\right\rangle = \sum_{i=1}^N \mathrm{w}_i^{(r-1)} K_\sigma\left(\xi^{(j)}, \xi^{(i)}\right) \tag{43}$$

$$\left\langle h^{(r)}, h^{(r)}\right\rangle = \sum_{i=1}^N\sum_{s=1}^N \mathrm{w}_i^{(r-1)}\mathrm{w}_s^{(r-1)} K_\sigma\left(\xi^{(i)}, \xi^{(s)}\right) \tag{44}$$

3.2.1.2. RKDE-based uncertainty set. In this subsection, an adjustable uncertainty set is constructed on the basis of the information obtained from uncertainty data with the help of the RKDE. Let $\hat{f}_{RKDE}^{(j)}\left(\xi_j\right)$ be the probability density function estimated for the jth uncertain parameter ξ_j from the uncertainty vector $\boldsymbol{\xi}$, $\hat{F}_{RKDE}^{(j)}\left(\xi_j\right)$ be the cumulative density function associated with the estimated function, and $\hat{F}_{RKDE}^{(i)-1}\left(\xi_j\right)$ be the quantile function defined as follows:

$$\hat{F}_{RKDE}^{(j)-1}(\beta) = \min\left\{\xi_j \in \mathbb{R} \middle| \hat{F}_{RKDE}^{(j)}\left(\xi_j\right) \ge \beta\right\} \tag{45}$$

where β is a predetermined parameter that confines the confidence region for each uncertain parameter given a confidence degree of $(1 - 2\beta)$. With the extracted quantile function, the uncertainty set is formed as follows [54]:

$$U = \left\{\boldsymbol{\xi} \middle| \begin{array}{l} \hat{F}_{RKDE}^{(j)-1}(\beta) \le \xi_j \le \hat{F}_{RKDE}^{(j)-1}(1-\beta) \quad \forall j \\ \sum_j \left(1 - \theta_j\varphi\right)\xi_j^0 \le \sum_j \xi_j \le \sum_j \left(1 + \theta_j\varphi\right)\xi_j^0 \end{array}\right\} \tag{46}$$

where φ is the budget of uncertainty introduced to control the maximum deviation allowed for the overall uncertainty, thereby setting the degree of conservatism. ξ_j^0 and θ_j are the set center and the scaled deviation from the center, respectively, which are computed as follows:

$$\xi_j^0 = \left[\hat{F}_{RKDE}^{(j)-1}(\beta) + \hat{F}_{RKDE}^{(j)-1}(1-\beta)\right]/2 \quad (47)$$

$$\theta_j = \left(\hat{F}_{RKDE}^{(j)-1}(1-\beta) - \xi_j^0\right)/\xi_j^0 \quad (48)$$

3.2.1.3. Robust counterpart formulation. This section presents the procedure for deriving the tractable robust counterpart formulation of the adjustable robust scheduling model (31) under the constructed uncertainty set. To keep the notation manageable, the second-stage problem of the robust model (31) is first expressed in the following compact form:

$$\min_{\mathbf{y}\in\Omega(x,U)} \max_{\xi\in U} \sum_t \mathbf{d}'_{et}\mathbf{y}_{et} \quad \forall e \in \mathcal{A} \quad (49a)$$

$$s.t.\mathbf{v}'_{\boldsymbol{e}}\mathbf{x}_{\boldsymbol{e}} + \sum_t \mathbf{w}'_{et}\mathbf{y}_{et} \le \mathrm{h}_e(\xi_e) \quad \forall \boldsymbol{\xi} \in U, e \in \mathcal{A} \quad (49b)$$

$$\text{over}\left\{\begin{array}{c} \mathbf{x}_{\boldsymbol{e}} = \left(\mathbf{P}^{\mathrm{f}}_{\mathrm{e}}, \mathbf{Q}^{\mathrm{f}}_{\mathrm{e}}, \mathbf{V}^{\mathrm{f}}_{\mathrm{e}}, \mathbf{I}^{\mathrm{f}}_{\mathrm{e}}, \mathbf{P}^{\mathrm{c}}_{\mathrm{e}}, \mathbf{P}^{\mathrm{dc}}_{\mathrm{e}}, \mathbf{E}^{\mathrm{es}}_{\mathrm{e}}, \mathbf{P}^{\mathrm{M}}_{\mathrm{e}}, \mathbf{Q}^{\mathrm{M}}_{\mathrm{e}}, \mathbf{P}^{ex}_{\mathrm{e}}, \mathbf{Q}^{ex}_{\mathrm{e}}\right) \\ \boldsymbol{y}_{\boldsymbol{et}} = \left(\Delta P^-_{et}, \Delta P^+_{et}, R^g_{et}\right) \end{array}\right\}$$

where $\mathbf{x}_{\boldsymbol{e}}$ is the vector of here-and-now (first-stage) decisions for entity $\boldsymbol{e}$ that should be made before uncertainty realizations are revealed, $\mathbf{y}_{\boldsymbol{e}}$ is the vector of wait-and-see (second-stage) decisions for entity $\boldsymbol{e}$ that can be adjusted based on observed realizations, $\mathbf{d}_e$ the vector of cost coefficients that correspond to $\mathbf{y}_{\boldsymbol{e}}$. Constraint (49b) is the matrix form of constraints (32)–(34). $\boldsymbol{\xi}$ stands for the uncertain parameter $\tilde{P}^{vp}$ and $\mathrm{h}(\boldsymbol{\xi})$ is the right-hand-side coefficient representing $\tilde{P}^{vp}$ moved from the left-hand side to the right-hand side of constraints (32) and (33). The above optimization problem is computationally intractable because of the curse of dimensionality. A pragmatic strategy to overcome this difficulty is to employ the affine decision rule approximation [65], which restricts the adjustable decisions to be an affine function of realizations:

$$\mathbf{y}_{et}\left(\boldsymbol{\xi}^t_{\boldsymbol{e}}\right) = \mathbf{M}_{et}\boldsymbol{\xi}^t_{\boldsymbol{e}} + \mathbf{n}_{et} \quad (50)$$

where $\mathbf{M}_{et}$ and $\mathbf{n}_{et}$ are new decision variables whose optimal values need to be determined, and $\boldsymbol{\xi}^t_{\boldsymbol{e}} = \left[\boldsymbol{\xi}'_{e1}, \ldots, \boldsymbol{\xi}'_{et}\right]$ is the vector of past realizations up to stage t. This strategy extends the two-stage model (49) into a multi-stage setting where recourse decisions are allowed to adapt to the uncertainties as they unfold over time. By substituting the decision rule (50) for the adjustable decisions, a conservative approximation to the robust scheduling model (49) can be derived:

$$\min_{\mathbf{M},\mathbf{n}} \max_{\xi\in U} \sum_t \mathbf{d}'_{et}\left(\dot{\mathbf{M}}_{et}\boldsymbol{\xi}_{\boldsymbol{e}} + \mathbf{n}_{e,t}\right) \quad \forall e \in \mathcal{A} \quad (51)$$

$$s.t.\mathbf{v}'_{\boldsymbol{e}}\mathbf{x} + \sum_t \mathbf{w}'_{et}\left(\dot{\mathbf{M}}_{et}\boldsymbol{\xi} + \mathbf{n}_{e,t}\right) \le \mathbf{h}'_{\boldsymbol{e}}\boldsymbol{\xi}_{\boldsymbol{e}} + \mathrm{h}^0_e \quad \forall \boldsymbol{\xi} \in U, \quad \forall e \in \mathcal{A}$$

$$U = \left\{\boldsymbol{\xi}_{\boldsymbol{e}} \left| \begin{array}{l} \hat{F}_{RKDE}^{(ejt)-1}(\beta_e) \le \xi_{ejt} \le \hat{F}_{RKDE}^{(ejt)-1}(1-\beta_e) \quad \forall e,j,t \\ \sum_j\sum_t\left(1-\theta_{ej}\varphi_e\right)\xi^0_{ejt} \le \sum_j\sum_t \xi_{ejt} \\ \quad \le \sum_j\sum_t\left(1+\theta_{ej}\varphi_e\right)\xi^0_{ejt} \quad \forall e \end{array}\right.\right\}$$

where $\boldsymbol{\xi}_{\boldsymbol{e}} = \left[\boldsymbol{\xi}'_{e1}, \ldots, \boldsymbol{\xi}'_{eT}\right]$, ξ_{ejt} is the jth element of vector $\boldsymbol{\xi}_{et}$, and $\dot{\mathbf{M}}_{et}$ is the modified form of $\mathbf{M}_t$ to make sure $\dot{\mathbf{M}}_{et}\boldsymbol{\xi} = \mathbf{M}_{et}\boldsymbol{\xi}^{\mathrm{t}}$. The right-hand-side coefficient is affinely dependent on the uncertainties, as expressed by $\mathbf{h}_{\boldsymbol{e}}\boldsymbol{\xi}_{\boldsymbol{e}} + \mathrm{h}^0_e$. The above min–max optimization problem with infinite-dimensional constraints is still intractable and should be converted into a tractable single-level problem. To this aim, it is first converted into the following equivalent formulation:

$$\min_{\mathbf{M},\mathbf{n},k} k_e \quad \forall e \in \mathcal{A}$$

$$s.t.\max_{\xi\in U}\left\{\left(\sum_t \mathbf{d}'_{et}\dot{\mathbf{M}}_{et}\right).\boldsymbol{\xi}_{\boldsymbol{e}}\right\} \le k_e - \sum_t \mathbf{d}'_{et}\mathbf{n}_{et} \quad \forall e \in \mathcal{A} \quad (52)$$

$$\max_{\xi\in U}\left\{\left(\sum_t \mathbf{w}'_{et}\dot{\mathbf{M}}_{et} - \mathbf{h}'_e\right).\boldsymbol{\xi}_{\boldsymbol{e}}\right\} \le \mathrm{h}^0_e - \mathbf{v}'_{\boldsymbol{e}}\mathbf{x}_{\boldsymbol{e}} - \sum_t \mathbf{w}'_{et}\mathbf{n}_{et} \quad \forall e \in \mathcal{A}$$

$$U = \left\{\boldsymbol{\xi}_{\boldsymbol{e}} \left| \begin{array}{l} \hat{F}_{RKDE}^{(ejt)-1}(\beta_e) \le \xi_{ejt} \le \hat{F}_{RKDE}^{(ejt)-1}(1-\beta_e) \quad \forall e,j,t \\ \sum_j\sum_t\left(1-\theta_{ej}\varphi_e\right)\xi^0_{ejt} \le \sum_j\sum_t \xi_{ejt} \\ \le \sum_j\sum_t\left(1+\theta_{ej}\varphi_e\right)\xi^0_{ejt} \quad \forall e \end{array}\right.\right\}$$

Based on the strong duality theory, the above inner maximization problems are replaced by its dual form, which yields the following tractable robust counterpart:

$$\min_{\substack{\mathbf{M},\mathbf{n},k \\ \boldsymbol{\pi},\boldsymbol{\tau}}} k_e \quad \forall e \in \mathcal{A} \quad (53a)$$

$$\begin{aligned} s.t.\,&\overline{\pi}_{ejt}\hat{F}_{RKDE}^{(ejt)-1}(1-\beta_e) - \underline{\pi}_{ejt}\hat{F}_{RKDE}^{(ejt)-1}(\beta_e) \\ &+ \overline{\pi}^0_e\sum_j\sum_t\left(1+\theta_{ej}\varphi_e\right)\xi^0_{ejt} \\ &- \underline{\pi}^0_e\sum_j\sum_t\left(1-\theta_{ej}\varphi_e\right)\xi^0_{ejt} + \sum_t \mathbf{d}'_{et}\mathbf{n}_{et} \le k_e \quad \forall e \in \mathcal{A} \end{aligned} \quad (53b)$$

$$\overline{\pi}_{ejt} - \underline{\pi}_{ejt} + \overline{\pi}^0_e - \underline{\pi}^0_e = \left[\sum_v \mathbf{d}'_{ev}\dot{\mathbf{M}}_{ev}\right]_{j+(t-1).J} \quad \forall e,j,t \quad (53c)$$

$$\begin{aligned} &\overline{\tau}_{ejt}\hat{F}_{RKDE}^{(ejt)-1}(1-\beta_e) - \underline{\tau}_{ejt}\hat{F}_{RKDE}^{(ejt)-1}(\beta_e) + \overline{\tau}^0_e\sum_j\sum_t\left(1+\theta_{ej}\varphi_e\right)\xi^0_{ejt} \\ &- \underline{\tau}^0_e\sum_j\sum_t\left(1-\theta_{ej}\varphi_e\right)\xi^0_{ejt} \le \mathrm{h}^0_e - \mathbf{v}'_e\mathbf{x}_{\boldsymbol{e}} - \sum_t \mathbf{w}'_{et}\mathbf{n}_{et} \quad \forall e \in \mathcal{A} \end{aligned} \quad (53d)$$

$$\overline{\tau}_{ejt} - \underline{\tau}_{ejt} + \overline{\tau}^0_e - \underline{\tau}^0_e = \left[\sum_v \dot{\mathbf{M}}'_{\boldsymbol{ev}}\mathbf{w}_{ev} - \mathbf{h}_{\boldsymbol{e}}\right]_{j+(t-1).J} \quad \forall e,j,t \quad (53e)$$

$$\boldsymbol{\pi}_e, \boldsymbol{\tau}_e \ge 0 \quad (53f)$$

where $\boldsymbol{\pi}$ and $\boldsymbol{\tau}$ are the vectors of dual variables introduced to dualize the first and second inner maximization problems in (50), J is the uncertainty dimension at each stage, and $[*]_j$ denotes the jth element of vector $*$.

After replacing the second-stage problem of the robust model (31) with its tractable robust counterpart (53) and reorganizing the decision variables, we arrive at the following robust scheduling problem:

$$\begin{aligned} &\min f^s_e(\boldsymbol{xs}_e, \boldsymbol{xt}_e) + f^{ex}_e(\boldsymbol{\omega}_e) \quad \forall e \in \mathcal{A} \\ &s.t.(9)\text{–}(11), (13)\text{–}(15), (20)\text{–}(30), (53b)\text{–}(53f) \quad \forall e \in \mathcal{M} \\ &\quad (8), (11), (13)\text{–}(30), (53b)\text{–}(53f) \quad \forall e \in \mathcal{D} \end{aligned} \quad (54)$$

where $\boldsymbol{xs}_e$ includes the variables of power flow $\left\{\mathbf{P}^{\mathrm{f}}_{\mathrm{e}}, \mathbf{Q}^{\mathrm{f}}_{\mathrm{e}}, \mathbf{V}^{\mathrm{f}}_{\mathrm{e}}, \mathbf{I}^{\mathrm{f}}_{\mathrm{e}}\right\}$, power generation $\left\{\mathbf{P}^{\mathrm{g}}_{\mathrm{e}}, \mathbf{Q}^{\mathrm{g}}_{\mathrm{e}}, \mathbf{P}^{\mathrm{rg}}_{\mathrm{e}}, \mathbf{P}^{\mathrm{v}}_{\mathrm{e}}\right\}$, ESSs $\left\{\mathbf{P}^{\mathrm{c}}_{\mathrm{e}}, \mathbf{P}^{\mathrm{dc}}_{\mathrm{e}}, \mathbf{E}^{\mathrm{es}}_{\mathrm{e}}\right\}$, and power exchange with the main grid $\left\{\mathbf{P}^{\mathrm{M}}_{\mathrm{e}}, \mathbf{Q}^{\mathrm{M}}_{\mathrm{e}}\right\}_{\forall e \in \mathcal{D}}$, and the robust optimization problem $\{\mathbf{M}, \mathbf{n}, \boldsymbol{\pi}, \boldsymbol{\tau}, k\}$, $\boldsymbol{xt}_e$ is the power trading variable $\left\{\mathbf{P}^{ex}_{e,j}, \mathbf{Q}^{ex}_{e,j}\right\}_{j\in\mathcal{A}\setminus e}$, and $\boldsymbol{\omega}_e$ is the payment variable. f^s_e corresponds to $f^g_e + f^{es}_e + f^m_e$ for DN ($e \in \mathcal{D}$) and $f^g_e + f^{es}_e$ for MG ($e \in \mathcal{M}$), respectively.

3.3. Data-driven robust scheduling and transaction payment model

DN and its connected MGs are self-interested entities that are inclined to cooperate and trade power with each other to reduce costs and improve reliability. Each entity bargains with other entities to determine the amount of power to be traded and associated payment. Nash bargaining theory is a cooperative game theory that can be used to mathematically model the bargaining process and facilitate reaching a mutually beneficial power transaction agreement through coordination and negotiation [66]. With two players participating in a bargaining game, a Nash bargaining solution is a pair of payoffs that solve the following optimization problem [67]:

$$\max_{s_1,s_2} (s_1 - d_1)(s_2 - d_2)$$
$$s.t.\ (s_1, s_2) \in S;\ (d_1, d_2) < (s_1, s_2) \quad (55)$$

where (s_1, s_2) and (d_1, d_2) are the payoffs of the players from an agreement and the disagreement point, respectively, and S is a convex set that contains the payoffs of all possible agreements. Based on the multi-player Nash bargaining solution, the data-driven robust energy management problem (54) can be formulated as the following scheduling and payment problem:

$$\max_{\boldsymbol{xs},\boldsymbol{xt},\boldsymbol{\omega}} \prod_{e\in\mathcal{A}'} \left[f_e^{non} - \left(f_e^s\left(\boldsymbol{xs}_e, \boldsymbol{xt}_e\right) + f_e^{ex}\left(\boldsymbol{\omega}_e\right)\right)\right] \quad (56a)$$

$s.t.$ (9)–(11), (13)–(15), (20)–(30), (53b)–(53f) $\forall e \in \mathcal{M}$

(8), (11), (13)–(30), (53b)–(53f) $\forall e \in \mathcal{D}$

$$f_e^s\left(\boldsymbol{xs}_e, \boldsymbol{xt}_e\right) + f_e^{ex}\left(\boldsymbol{\omega}_e\right) \leq f_e^{non} e \in \mathcal{A}' \quad (56b)$$

where $\mathcal{A}' \subset \mathcal{A}$ is the set of entities that are willing to participate in power trading, and f_e^{non} is the minimum cost that entity e can obtain without power trading with others. Constraint (56b) states that each entity will only trade power if it can reduce its total costs. Solving problem (56) ensures that trading benefits are fairly allocated to different entities. Moreover, it yields the optimal power scheduling along with the optimal scheme of power trading and payment for each entity $e \in \mathcal{A}'$ [47]. However, the set of trading entities (i.e., $\mathcal{A}'$) is unknown and problem (56) is computationally difficult to solve. To overcome the above barriers, problem (56) is equivalently decomposed into two sub-problems, which are solved sequentially [68]:

S1: Scheduling and trading problem

$$\min_{\boldsymbol{xs},\boldsymbol{xt}} \sum_{e\in\mathcal{A}} f_e^s\left(\boldsymbol{xs}_e, \boldsymbol{xt}_e\right) \quad (57a)$$

$s.t.$(9)–(11), (13)–(15), (20)–(27), (29), (30), (53b)–(53f) $\forall e \in \mathcal{M}$ (57b)

(8), (11), (13)–(27), (29), (30), (53b)–(53f) $\forall e \in \mathcal{D}$

S2: Payment allocation problem

$$\max_{\boldsymbol{\omega}} \prod_{e\in\mathcal{A}'} \left[f_e^{non} - f_e\left(\boldsymbol{xs}_e^*, \boldsymbol{xt}_e^*\right) - f_e^{ex}\left(\boldsymbol{\omega}_e\right)\right] \quad (58a)$$

$$s.t.(28),\ (56b) \quad \forall e \in \mathcal{A}' \quad (58b)$$

Sub-problem S1 minimizes the operating costs of the whole system and determines optimal scheduling and trading decisions $(\boldsymbol{xs}_e^*, \boldsymbol{xt}_e^*)$ for all entities. Entities with a non-zero trading vector $\boldsymbol{xt}_e^*$ are eager to trade power and thus belong to set $\mathcal{A}'$, while other entities $(e \in \mathcal{A}\backslash\mathcal{A}')$ do not participate in the trading market. The non-trading entities are not involved in the payment allocation problem and have the same operating costs as f_e^{non}. Based on the optimal solution of S1, the trading entities bargain over the allocation of benefits through sub-problem S2, where bilateral payments are fairly determined. S2 is a geometric programming problem that can be converted to its convex form by the logarithmic transformation:

$$\min_{\boldsymbol{\omega}} \sum_{e\in\mathcal{A}'} -\ln\left[f_e^{non} - f_e\left(\boldsymbol{xs}_e^*, \boldsymbol{xt}_e^*\right) - f_e^{ex}\left(\boldsymbol{\omega}_e\right)\right] \quad (59a)$$

$$s.t.\ (28),\ (56b) \quad \forall e \in \mathcal{A}' \quad (59b)$$

4. Solution methodology

Solving the sub-problems S1 and S2 in a centralized fashion requires the internal information of DN and MGs, which may violate the autonomy and privacy of individual entities. In this regard, each sub-problem is solved in a privacy-preserving distributed way using the ADMM and ALADIN algorithms. To formulate the distributed optimization algorithms, the original decision variables are reorganized into two groups. The first group includes private variables that contain sensitive information and cannot be accessed by other entities. The second group includes coupling variables whose values are exchanged between the entities. Based on the defined variables, S1 and S2 can take the following form:

$$\min_{\boldsymbol{\chi_e}} \sum_e f_e\left(\boldsymbol{\chi}_e\right) \quad (60a)$$

$$s.t.\ L_e\left(\boldsymbol{\chi}_e\right) \leq 0 \quad \forall e \quad (60b)$$

$$\mathbf{G}_e(\boldsymbol{\chi}_e - \mathbf{z}_e) = 0 \quad \forall e \quad (60c)$$

$$\sum_e \mathbf{G}_e\mathbf{z}_e = 0 \quad (60d)$$

For S1, $\boldsymbol{\chi_e}$ combines the coupling variable $\mathbf{xt}_e$ and the private variable $\mathbf{xs}_e$, $f_e\left(\boldsymbol{\chi_e}\right)$ is the cost function (57a), and $L_e(\boldsymbol{\chi_e})$ collects all constraints in (57b) except (26) and (27). For S2, $f_e\left(\boldsymbol{\chi_e}\right)$ is the cost function (59a), $L_e(\boldsymbol{\chi_e})$ corresponds to constraint (56b), and $\boldsymbol{\chi}_e$ includes the coupling variable $\boldsymbol{\omega_e}$. Each entity has its own coupling variables and must make them equal to those of its connected entities, with the result that different entities are connected together and a consensus between them is achieved [69]. This is ensured by constraints (60c) and (60d), where $\mathbf{z_e}$ is an auxiliary (global) variable introduced for each entity as a copy of its variable and $\mathbf{G}_e$ is a selection matrix that separates coupling variables from all variables. The last constraint is referred to as the consensus constraint that enforces the equality of the coupling variables of connected entities.

By defining $\boldsymbol{\lambda}_e$ as the dual variable associated with (60c), the augmented Lagrangian function of (60) is constructed as follows:

$$L_\rho\left(\boldsymbol{\chi}, \mathbf{z}, \boldsymbol{\lambda}\right) = \sum_e \left\{f_e\left(\boldsymbol{\chi}_e\right) + \boldsymbol{\lambda}_e'\mathbf{G}_e\left(\boldsymbol{\chi}_e - \mathbf{z}_e\right) + \frac{\rho}{2}\left\|\mathbf{G}_e\left(\boldsymbol{\chi}_e - \mathbf{z}_e\right)\right\|^2\right\} \quad (61)$$

where ρ is a penalty parameter. The general idea of ADMM and ALADIN is to minimize (61) over $\boldsymbol{\chi}_e$ and $\mathbf{z}_e$ in a sequential or alternating way. Specifically, the entities perform the $\boldsymbol{\chi}$-minimization step separately in parallel on the basis of fixed global variable $\mathbf{z}_e$ and dual variable $\boldsymbol{\lambda}_e$. The new values of coupling variables are sent to the central aggregated where the global variables are updated in specific manner depending on the used algorithm. ADMM computes the average of all received coupling variables while ALADIN solves a coupled quadratic problem based on the Jacobian, Hessian and gradient estimates of the entities' local problems. The updated global variables are redistributed to the entities for updating the dual variables and carrying out a new $\boldsymbol{\chi}$-minimization. The sequential updating of $\boldsymbol{\chi}_e$, $\mathbf{z}_e$ and $\boldsymbol{\lambda}_e$ continues until the coupling variables of connected entities converge to a common value. The iterative process implemented through ADMM and ALADIN is described in more details in the following.

4.1. ADMM algorithm

The ADMM algorithm consists of three main steps [50]. In step 1, the robust model (60) is reformulated based on the augmented Lagrangian function. The local minimization problems are then solved separately in parallel, yielding the values of the private and coupling variables for each entity. In step 2, the values of the coupling variables are sent from all entities to a central aggregator, which computes the new value of the global variable through an optimization problem. It is proved that the solution of this optimization problem can be obtained in a simpler way by averaging the coupling variables of connected entities [52]. The updated value of the global variable is then scattered to the entities. In step 3, the Lagrange multipliers are updated to start a new round of local variable optimization. The updating process in steps 1–3 is repeated until the consensus constraints are fulfilled and the deviation of the coupling variables from their respective global variables becomes sufficiently small. To summarize the procedure above, the pseudocode of ADMM is provided in Algorithm 1. For better readability, the dual variables are immediately written after the constraints.

Algorithm 1

Input: Initial guesses $\boldsymbol{\lambda}_e, \mathbf{z}_e$; penalty parameter ρ, and tolerance ϵ

while $\|\mathbf{G}\boldsymbol{\chi}\|_\infty > \epsilon$ and $\|\boldsymbol{\chi} - \mathbf{z}\|_\infty > \epsilon$ **do**

1: Solve for all e the local optimization problems

$$\min_{\boldsymbol{\chi}_e} f_e(\boldsymbol{\chi}_e) + \boldsymbol{\lambda}'_e \mathbf{G}_e \boldsymbol{\chi}_e + \frac{\rho}{2}\|\mathbf{G}_e(\boldsymbol{\chi}_e - \mathbf{z}_e)\|^2$$

$$s.t.\ L_e(\boldsymbol{\chi}_e) \leq 0$$

2: Solve the global optimization problem

$$\min_{\mathbf{z}} \sum_e -\boldsymbol{\lambda}'_e \mathbf{G}_e \boldsymbol{\chi}_e + \frac{\rho}{2}\|\mathbf{G}_e(\boldsymbol{\chi}_e - \mathbf{z}_e)\|^2$$

$$s.t.\ \sum_e \mathbf{G}_e \mathbf{z}_e = 0$$

3: Update for all e the Lagrange multipliers

$$\boldsymbol{\lambda}_e \leftarrow \boldsymbol{\lambda}_e + \rho \mathbf{G}_e(\boldsymbol{\chi}_e - \mathbf{z}_e)$$

end while

4.2. ALADIN algorithm

The main idea of the ALADIN algorithm is to form the step of the global variable update in ADMM based on the gradient and Hessian of local problems, aiming to improve the convergence rate and ensure convergence in the case of non-convex problems [56]. Algorithm 2 shows the pseudocode of ALADIN for solving the robust model (60), including four main steps [53]. Step 1 is similar to that of ADMM; the only difference is that ALADIN considers only one global dual variable and formulates the augmented term with the weighted norm $\|*\|^2_\Sigma = *'\Sigma*$, where Σ_e is a symmetric and positive semidefinite weighting matrix. In step 2, the gradient of the objective function, the Hessian approximation of the Lagrange function and the Jacobian approximation of the constraints are calculated for each entity and communicated to the central aggregator to solve a quadratic programming problem in step 3. Once the primal variable increments are computed, they must be broadcast to the entities to update $\mathbf{z}_e$ and $\boldsymbol{\lambda}$ in step 4 before running the individual minimization problems. The algorithm then starts from the beginning and continues until the termination conditions are satisfied.

Algorithm 2

Input: Initial guesses $\boldsymbol{\lambda}, \mathbf{z}_e$; penalty parameter ρ, and tolerance ϵ

While $\|\mathbf{G}\boldsymbol{\chi}\|_\infty > \epsilon$ and $\|\boldsymbol{\chi} - \mathbf{z}\|_\infty > \epsilon$ **do**

1: Solve for all e the local optimization problems

$$\min_{\boldsymbol{\chi}_e} f_e(\boldsymbol{\chi}_e) + \boldsymbol{\lambda}'\mathbf{G}_e \boldsymbol{\chi}_e + \frac{\rho}{2}\|(\boldsymbol{\chi}_e - \mathbf{z}_e)\|^2_{\Sigma_e}$$

$$s.t.\ L_e(\boldsymbol{\chi}_e) \leq 0 \quad |\boldsymbol{\kappa}_e$$

2: Compute for all e the Jacobian, gradient and Hessian matrix

$$\boldsymbol{C}_{e,j} = \begin{cases} \frac{\partial}{\partial y}\left(L_e(\boldsymbol{y})\right)_j\Big|_{y=\boldsymbol{\chi}_e} & \text{if}\left(L_e(\boldsymbol{\chi}_e)\right)_j = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{g}_e = \nabla f_e(\boldsymbol{\chi}_e)$$

$$\boldsymbol{H}_e = \nabla^2\{f_e(\boldsymbol{\chi}_e) + \boldsymbol{\kappa}'_e L_e(\boldsymbol{\chi}_e)\}$$

3: Solve the global optimization problem

$$\min_{\Delta\chi, \boldsymbol{d}} \sum_e \left\{\frac{1}{2}\Delta\boldsymbol{\chi}'_e H_e \Delta\boldsymbol{\chi}_e + g'_e \Delta\boldsymbol{\chi}_e\right\} + \boldsymbol{\lambda}'\boldsymbol{d} + \frac{\mu}{2}\|\boldsymbol{d}\|^2$$

$$s.t.\ \sum_e \boldsymbol{G}_e(\boldsymbol{\chi}_e + \Delta\boldsymbol{\chi}_e) = \boldsymbol{d} \quad |\boldsymbol{\lambda}_Q$$

$$\boldsymbol{C}_e \Delta\boldsymbol{\chi}_e = 0 \quad \forall e$$

4: Update for all e the global variables and Lagrange multiplier

$$\mathbf{z}_e \leftarrow \boldsymbol{\chi}_e + \Delta\boldsymbol{\chi}_e \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda}_Q$$

end while

5. Case study

The proposed data-driven model for the energy scheduling of multiple MGs is tested on a modified IEEE 33-bus DN connected with three MGs [70,71], which is illustrated in Fig. 3. The total active/reactive load for MG1, MG2 and MG3 is 0.21/0.12 (p.u.), 0.105/0.07 (p.u.) and 0.072/0.048 (p.u.), respectively [71]. The load consumption at each bus is equal to others, and the base load is extended to multi-time periods using the load variation coefficients from [72]. The line resistance and reactance are 0.006 and 0.01 (p.u.), respectively, and the base power of the system is 10 MVA [71]. MTs are used as dispatchable generators in this study, with a second-degree polynomial cost function whose coefficients are set at 0.0003 $/(KWh)^2, 0.3 $/KWh and 0 $, respectively [70]. The cost of reserve capacity of MTs is set at 20% of their highest marginal cost of generation [73]. Two types of RGs are considered in this study, including WTs and PVs. The output of WTs and PVs is calculated as the sum of their predicted value and prediction error, where the predicted value is regarded as a fixed nominal value and the prediction error is treated as a random variable fluctuating around the nominal value. The Gaussian distribution that well represents WT and PV prediction errors is utilized to generate 1000 samples [74]. The predicted outputs of WTs and PVs during different time periods are obtained from [72], and the mean and standard deviation of the Gaussian distribution are set at zero and 10% of the predicted outputs, respectively [74]. Moreover, 100 contaminating samples drawn from the uniform distribution are added to the dataset. The selling price ϑ^s_t is set at 0.4 $/kWh, and the buying price ϑ^b_t is set at 0.8 $/kWh for the off-peak periods (between 24:00 and 6:00) and 1 $/kWh during other time periods [68]. The power shortage and surplus for DN and MGs are penalized by 1.1 $/kWh and 0.5 $/kWh, respectively. Each bus is equipped with an ESS whose technical parameters such as charge/discharge efficiency and maximum

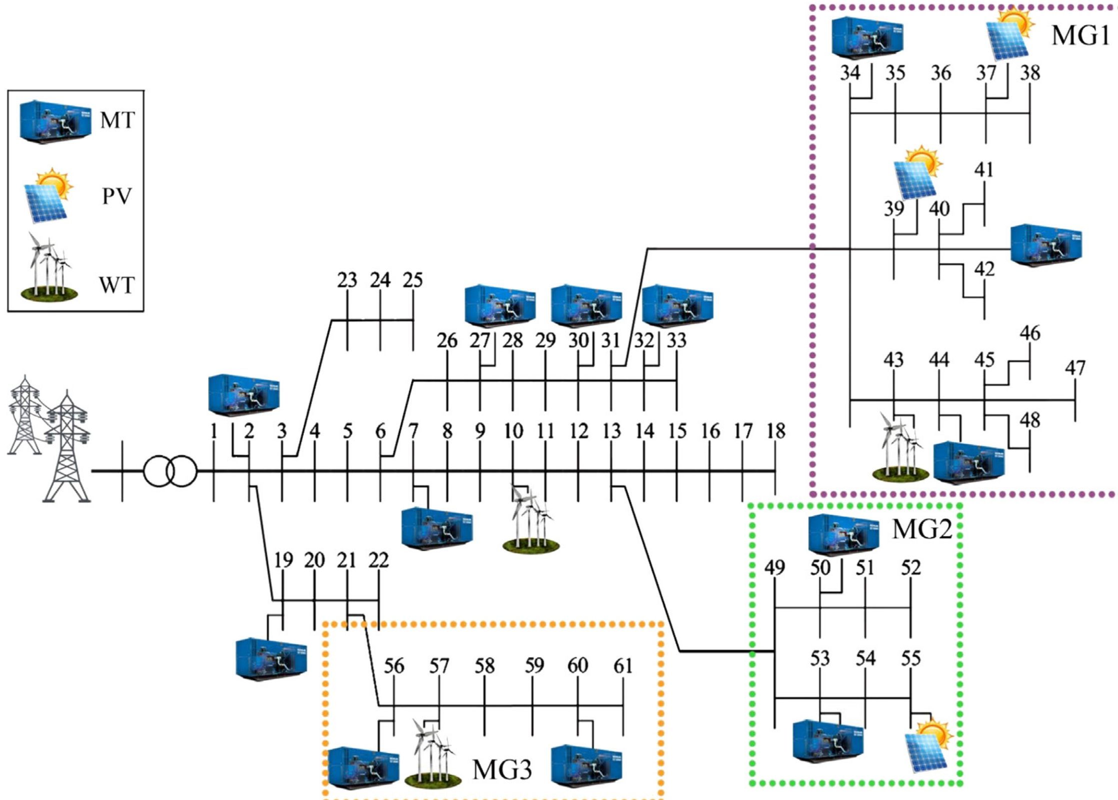

Fig. 3. Modified IEEE-33 bus DN connected with three MGs.

charge/discharge power are obtained from [45]. The upper and lower deviations of bus voltage are bounded within $\pm 10\%$ of a nominal voltage. The details about DN can be found in [75]. The convergence tolerance ϵ is set as 10^{-3}. The initial solution for ALADIN and ADMM is derived by individually solving the robust scheduling model of DN and MGs without considering any power exchange between them. All optimization problems are implemented on a PC with an Intel Core i5 CPU and 8 GB RAM and solved using MATLAB 2016a with the CPLEX solver.

5.1. Comparative cost analysis of robust and deterministic models

This subsection evaluates the economic performance of the proposed data-driven model under the RKDE-induced uncertainty set in comparison with commonly used approaches, including (i) the deterministic scheduling model where all uncertain parameters are set at their nominal values and (ii) the robust scheduling model under the polyhedral uncertainty [76]. Due to the different nature of the polyhedral uncertainty set and the RKDE-induced uncertainty set, it is impossible to similarly adjust their conservatism level. To make a fair comparison, they are assessed under different levels of conservatism and uncertainty data coverage. To be more specific, the confidence level for the RKDE-induced uncertainty set, which is equal to $(1-2\beta)$, is set as 0.99, 0.95, 0.90, 0.85 and 0.80. The parameter φ, which controls the maximum deviation of the overall uncertainty, is chosen to be 0.05, 0.10, 0.15, 0.20 and 0.25. For the polyhedral set, the variation range of uncertain parameters is assumed to be 100%, 95%, 90%, 85% and 80% of the distance between their lower and upper bounds. It is further assumed that 100%, 90%, 80%, 70% and 60% of uncertain parameters take their worst-case values within the variation range while others are set at their nominal values.

Fig. 4 shows the total costs of MGs determined by the three optimization approaches. In all cases, the deterministic approach leads to the lowest operating costs. However, it provides scheduling schemes that might lose their feasibility by small variations in renewable energy generation. To compare the performance of robust solutions, the price of robustness (PR) is one of the main criteria that is often employed. This criterion measures how much optimality must be sacrificed in order to achieve robustness, and is defined as $PR = (obj_r - obj_d)/obj_d$, where obj_d and obj_r are the optimal costs returned by the deterministic and robust models, respectively. The results indicate that the RKDE-induced set protects against uncertainty at a lower cost than the polyhedral uncertainty set. In MG1, for example, the polyhedral set ensures robustness against 60% of the uncertain parameters with a PR of about 14%, while the RKDE-induced set increases the optimal cost by about 13% to safeguard against all the uncertain parameters. As a result, by extracting the probabilistic information from uncertain renewable energy sources, the RKDE-induced set substantially reduces the PR of conventional uncertainty sets such as the polyhedral set, which are constructed without considering the probability of realization of uncertain parameters. It can also be seen from Fig. 4 that the difference between the costs of the two robust approaches is larger for MG1 than for MG2 and MG3, particularly when the confidence level and the variation range increase. This result is expected because MG1 is subject to a greater number of uncertainty sources than the other two MGs. The proposed data-driven framework equipped with the RKDE-induced set guarantees the robustness of scheduling and trading plans while sacrificing only a small part of economic benefits and is therefore more practical for the energy management of MGs under multiple uncertainty sources.

5.2. Robustness verification

For the purpose of verifying the quality of the optimal solutions obtained by the proposed model, a posteriori analysis of

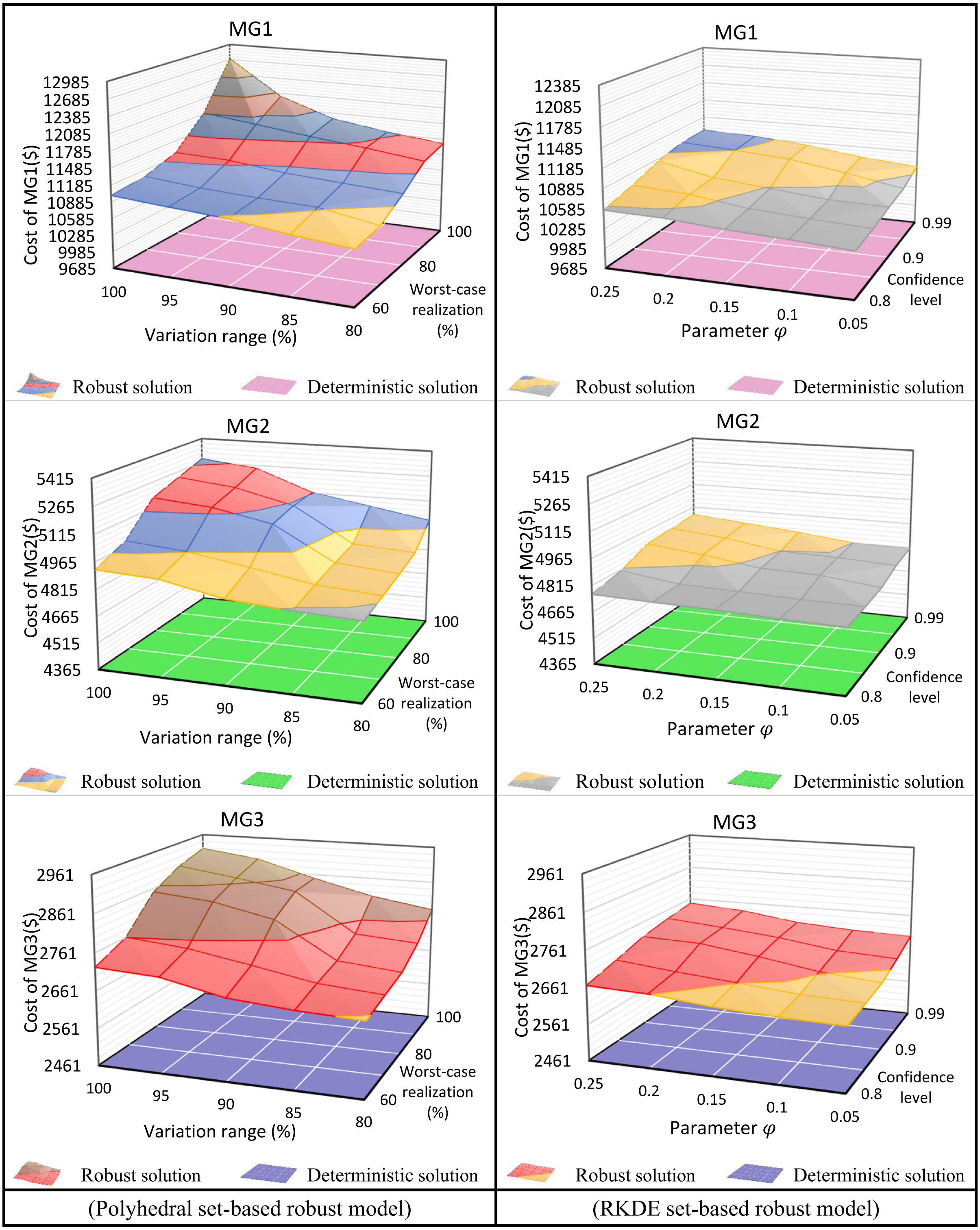

Fig. 4. Comparison of cost of MGs under the deterministic and robust models.

solution robustness is conducted in this subsection to examine how much the robustness guaranteed by the two robust models is achieved in reality [16]. This analysis considers five cases with different uncertainty magnitudes, which are constructed by changing the variance of the generated samples of WT and PV power outputs based on factors of $V = \{0.6, 0.7, 0.8, 0.9, 1\}$. A smaller value of V means the power outputs are confined within a smaller range and thus have less impact on the performance of MGs.

In each case, the optimal first-stage decisions provided by the robust models along with the newly generated samples are entered into the optimization problem (31), which is solved to determine the recourse decisions and total cost. Under this setting, the recourse decisions are allowed to update their values since they can be made after the uncertainty is realized, while the first-stage decisions are treated as fixed parameters since they must be made at the beginning of the scheduling horizon prior to observing the realizations [77,78]. After repeating the above

Table 2
Costs and payments of DN and MGs in the robust models.

Model	Scheme	DN	MG1	MG2	MG3	Total
Polyhedral set-based robust model	Cost (without trading)	16294	14491	5846	3186	39817
	Cost (with trading)	17276	11474	3835	4219	36804
	Payment (with trading)	−1377	1227	1451	−1301	0
	Cost and payment (with trading)	14899	12701	5286	2918	36804
RKDE set-based robust model	Cost (without trading)	13879	12300	5307	2954	34440
	Cost (with trading)	14494	10018	3669	4307	32488
	Payment (with trading)	−619	935	1245	−1561	0
	Cost and payment (with trading)	12875	10953	4914	2746	32488

procedure for all the robust solutions derived under different confidence levels and variation ranges, the mean and standard deviation of the obtained total costs are calculated as two indexes to measure robustness. The process carried out for each case can be outlined as follows:

Step 1. The optimal first-stage decisions obtained by the robust model under a certain confidence level and variation range are fixed as parameters in the problem (31).

Step 2. The new data samples generated by changing the variance of the original samples of WT and PV outputs are individually plugged into the problem (31).

Step 3. For each sample, the problem (31) that now has a single-stage deterministic framework is solved, and the second-stage decisions and the total cost are determined.

Step 4. The above steps are repeated for the robust solutions under other confidence levels and variation ranges, and the mean and standard deviation of all the total costs are calculated as robustness criteria to compare the performance of the two robust optimization approaches.

The results of the posteriori analysis described above are displayed in Fig. 5. When the uncertainty magnitude is larger, the RKDE-based robust model significantly outperforms the polyhedral-based robust model in terms of the expected cost. As the uncertainty magnitude decreases, almost the same performance is achieved by both models. The reason is that both models hedge against the worst-case uncertainty realizations, while the random realizations generated for the performance evaluation shrink around the nominal values when the level of uncertainty is reduced. It implies that the superiority of the RKDE-based model is more pronounced in cases where renewable power supply is accompanied by high fluctuations and forecasting errors.

It can also be observed from Fig. 4 that the RKDE-based robust model achieves better performance in terms of optimality robustness, which means the operational costs of MGs under possible realizations of WT and PV outputs remain close to the optimal cost determined by that model. Comparing the results of the three MGs shows that the gap between the costs of the robust models is wider for MG1 than for MG2 and MG3. Specifically, when the magnitude factor is set at 1, the economic benefit of using the RKDE-based robust model for MG1 is 28%, while that for MG2 and MG3 are 21% and 19%, respectively. This stems from the fact that MG1 is confronted with greater uncertainty since it includes more WT and PV units. The above findings suggest that the RKDE-based model is a preferable option for MG operators who seek to enhance the worst-case performance of MGs influenced by multiple sources of uncertainty.

5.3. Power scheduling results of robust models

The optimal scheduling schemes for MG1, MG2 and MG3 are shown in Fig. 6. These schemes are determined by the robust models, which are both set at their maximum confidence level and variation range to better investigate the difference between them. Compared to the polyhedral set-based model, the RKDE set-based model reduces the purchased power from MGs and the generated and reserved power of MTs. The difference becomes more pronounced when renewable power generation and associated uncertainty increase. The reason is that the polyhedral set ensures robustness against the worst-case values of renewable power outputs with extremely low or even no probability of being realized, which makes MGs buy more power from others or produce more power from MTs to compensate for the deviation of renewable energy sources from their expected outputs. By contrast, the RKDE set hedges against those worst-case scenarios that are likely to take place, avoiding further increases in power production and reserve capacity. By taking into account realistic worst-case uncertainty realizations, the RKDE set results in obtaining more power from WTs and PVs. This advantage allows MGs to reduce power generation from MTs with high operating costs. In addition, MGs can sell excess renewable power to make more revenue from power trading. Using the RKDE set-based model, MG1 and MG3 have higher wind power outputs, which can be used locally to reduce the amount of power purchased from other entities or can be sold after meeting internal loads. With more power generated by PVs between 9:00 and 18:00, MG1 and MG2 have the opportunity to consume less non-renewable power and sell more power for additional revenue. The polyhedral set-based model, on the other hand, charges more power into ESSs to compensate for the decreased renewable power production and to meet peak loads. However, the operating cost of ESSs is higher than that of WTs and PVs, which increases the total costs of MGs.

5.4. P2P power trading results of robust models

Table 2 shows the operating costs and transaction payments of DN and MGs determined by the robust models with the maximum confidence level and variation range. P2P power trading in the polyhedral set-based model reduces the operating cost of the entire system (DN and MGs) by 13.5% from 39817 to 36804, which is higher than the reduction caused by it in the RKDE set-based model. However, the combined cost and payment of the RKDE set-based model is lower than that of the polyhedral set-based model, making it more economically viable for the energy transaction management of networked MGs. Comparing the results of the RKDE set-based model for different entities shows that MG1 and MG2 have lower operating costs after participating in the trading, while DN and MG3 have higher operating costs. The reason is that, even with the increased operation costs, DN and MG3 can make more profit by generating more power and selling it to MG1 and MG2. When the operating costs and transaction payments are taken into account simultaneously, the individual costs of DN, MG1, MG2 and MG3 are reduced by 7.8%, 12.3%, 8.2% and 7.6%, respectively. Therefore, all entities are benefited by P2P power trading and the effectiveness of the proposed trading and payment scheme is verified.

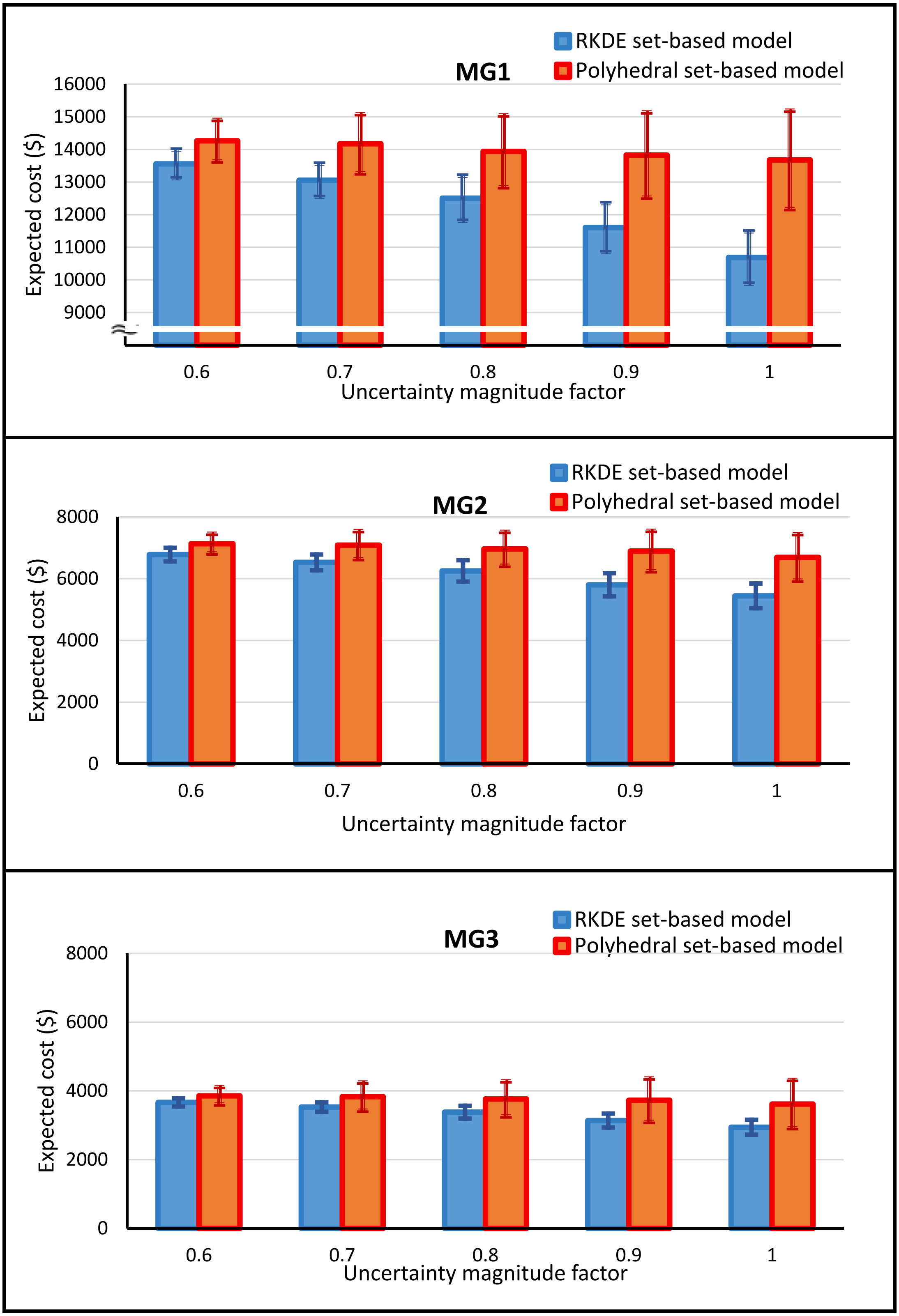

Fig. 5. Mean and standard deviation of MG costs under the realization of uncertain renewable generation.

5.5. Computational performance of ADMM and ALADIN algorithms

The purpose of this subsection is to examine the computational costs of solving the RKDE set-based model (with a confidence level of 0.99 and uncertainty deviation of 0.25) using the ADMM and ALADIN algorithms. In addition to the IEEE 33-bus DN with three MGs (i.e., MG1, MG2 and MG3), a modified IEEE 123-bus DN is analyzed to further illustrate the convergence behavior of the algorithms as the system size increases. This network is connoted to nine MGs, denoted by MGA to MGC with the same configuration as MG1, MGD to MGF with the same configuration as MG2 and MGG to MGI with the same configuration as

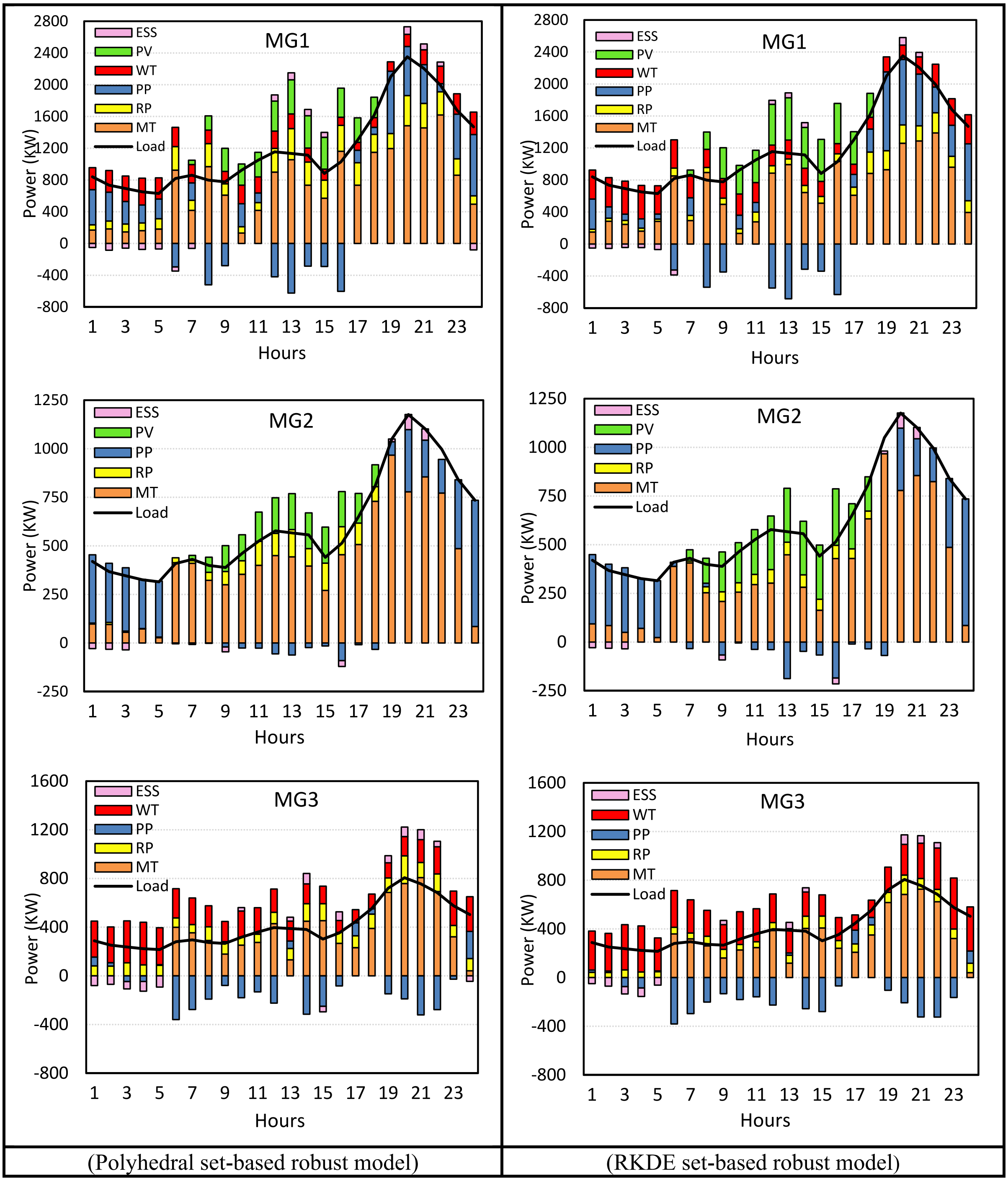

Fig. 6. Comparison of power scheduling schemes determined by the robust models.

MG3. The diagram and related data of this system can be found in [70]. Since both algorithms perform differently depending on the penalty parameter ρ, they are run with different values of ρ in order to ensure a fair comparison. The convergence process of ADMM and ALADIN for solving subproblem S1 of MG1 and MGA is shown in Fig. 7. The computation time of subproblem S2 is less than 1 s in all experiments, which has no impact on the total computation time. The cost of MGs in the first iteration, which can be referred to as the non-exchange solution, is lower than that in the final iterations. This is because when the coupling constraints are not met at the beginning, each MG can achieve its lowest cost thanks to enough power provided by other entities. As the iteration proceeds, more information is exchanged between the entities and the cost increases progressively until all constraints are met in the final iteration and agreement is obtained between them. In comparison to MG1, MGA requires more iterations and a longer run time to reach the converged solution. Due to the involvement of more MGs in the negotiation process of the IEEE 123-bus system, more iterations and computational effort are needed to form a consensus among all of them.

As proved in [52,53], the convexity of the proposed models ensures that ADMM and ALADIN converge to the optimal solution regardless of how the penalty parameters and initial solution are selected. Although both algorithms yield the same optimal cost eventually, ALADIN converges faster than ADMM in terms of the number of iterations and computation time. The superiority of

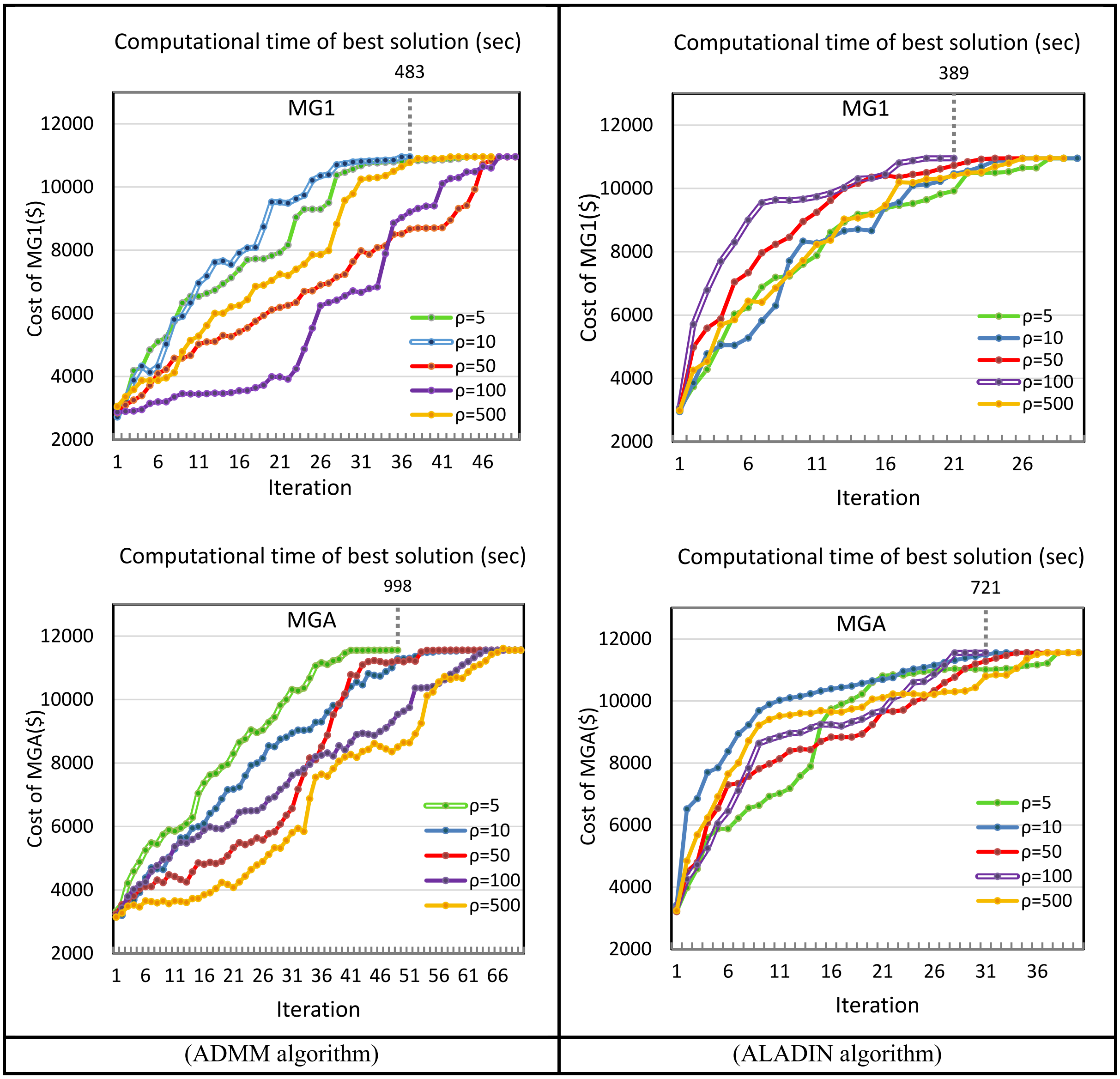

Fig. 7. Computational performance of ADMM and ALADIN under different penalty parameters.

ALADIN is more obvious for MGA that belongs to a larger network with more MGs. A more detailed comparison reveals that using ALADIN rather than ADMM results in a 19% and 28% reduction in run-time for MG1 and MGA, respectively. As expected, each iteration of ALADIN takes a longer time because it involves solving a quadratic programming problem to update the global variables, which is computationally more expensive than the simple averaging problem of ADMM. The strength of ALADIN can be seen in the first iterations after which it achieves near-optimality, while ADMM takes more iterations and time to come to that point. Therefore, ALADIN has desirable convergence properties, making it more practical for large-scale systems than ADMM.

6. Conclusions

This paper proposes a game theoretic-based collaborative framework for the distributed power management of DN and multiple MGs, each of which solves its own optimization problem for the co-optimization of MG' internal operations and P2P energy trading without disclosing its private information or internal operation to others. The main contributions are threefold. First, a combined Nash barraging game and DDR optimization model is developed to not only optimize the robustness cost of scheduling and trading decisions but also to incentivize power trading among MGs through a fair benefit allocation mechanism. Previous studies either focus on P2P energy trading problems with over-conservative robust optimization approaches under fixed-shape uncertainty sets or formulate DDR-based energy management models without designing a proper mechanism for the allocation of trading benefits. Second, the RKDE is used to construct a self-adaptive uncertainty set and improve the performance of the proposed DDR optimization model in handling messy and large-scale uncertainty data. Third, the individual problems of DN and MGs are coordinated and solved by developing two distributed optimization models based on the ADMM and ALADIN algorithms.

To make a fair comparison, the proposed model is tested on the 33-bus DS with three MGs and the 123-bus DS with nine MGs that are commonly used in the previous studies [18,40,70,71]. The results indicate that the trading and payment mechanism reduces the operating costs of all MGs and incentivizes them to participate in the trading. The RKDE uncertainty set outperforms conventional fixed-shape sets in hedging against uncertainty since it ensures the robustness of scheduling and trading decisions while providing more economic benefits. This superiority becomes more pronounced as the number of WTs and PVs in MGs increases. Therefore, the RKDE set-based energy management model is a reliable option to cope with uncertainty in future multi-MGs systems with high penetration of renewable energy resources. The comparative analysis of ADMM and ALADIN shows that although both take a finite number of iterations to successfully converge to the optimum, ALADIN is much faster than ADMM, particularly as the number of MGs connected to DN and the complexity of the consensus process increase. Compared to standard ADMM algorithms, the ALADIN-based distributed energy trading model provides higher computational efficiency and better scalability for application in large-scale systems.

The current study can be extended in the following ways. First, the developed model can be formulated in a more accurate way by considering the startup/shutdown status of MT generators and their up/down ramping rates [79]. Second, although ALADIN demonstrates an impressive improvement in convergence rate, it requires MGs to communicate more information to the central aggregator, which may raise privacy concerns. It is of significant interest to design a mechanism that allows MG operators to control the trade-off between the amount of information exchanged and the speed of convergence. Third, despite the fact that all experiments in this study are performed on a single machine, parallel computation on multiple machines connected together may provide a better picture of the performance of the scheduling model.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Shayan Mohseni: Conceptualization, Methodology, Software, Writing – original draft. **Mir Saman Pishvaee:** Supervision, Conceptualization, Writing – review & editing. **Reza Dashti:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix

Property 1. *In the optimal solution of the energy management model with minimization of ESSs' operating costs in objective function (1) and constraints (20)–(23), each ESS cannot charge and discharge at the same time.*

Proof (*By Contradiction*)**.** *Let us assume that the charging and discharging of each ESS happen simultaneously during period t in the optimal solution. Then,* $P^{c}_{i,t} = a > 0$ *and* $P^{dc}_{i,t} = b > 0$*. Because of its charge and discharge efficiencies, the ESS loses* b/η^{dc} *units of power and stores* $a.\eta^{c}$ *units of power, and the net stored power is equal to* $s = a.\eta^{c} - b/\eta^{dc}$ *with the operating cost* $C^{es}(a+b)$*. However, s units of power can be provided in another way by just charging* $s/\eta^{c} = a - b/\eta^{dc}.\eta^{c}$ *units. Since* $\eta^{c}, \eta^{dc} < 0$*, we have* $a - b/\eta^{dc}.\eta^{c} < a - b$*. Therefore, the same level of power storage can be obtained by generating less power (less generation cost). Moreover, the cost of ESS in this case is equal to* $C^{es}\left(a - b/\eta^{dc}.\eta^{c}\right)$*, which is lower than* $C^{es}(a+b)$*. This is in contradiction with our assumption that the solution with simultaneous charge and discharge is optimal.*

References

[1] E. Noghreian, H.R. Koofigar, Power control of hybrid energy systems with renewable sources (wind-photovoltaic) using switched systems strategy, Sustain. Energy Grids Netw. 21 (2020) 100280.
[2] R. Wu, G. Sansavini, Active distribution networks or microgrids? Optimal design of resilient and flexible distribution grids with energy service provision, Sustain. Energy Grids Netw. 26 (2021) 100461.
[3] J. Mei, C. Chen, J. Wang, J.L. Kirtley, Coalitional game theory based local power exchange algorithm for networked microgrids, Appl. Energy 239 (2019) 133–141.
[4] J. Li, M.E. Khodayar, J. Wang, B. Zhou, Data-driven distributionally robust co-optimization of P2P energy trading and network operation for interconnected microgrids, IEEE Trans. Smart Grid 12 (2021) 5172–5184.
[5] W. Huang, W. Zheng, D.J. Hill, Distributionally robust optimal power flow in multi-microgrids with decomposition and guaranteed convergence, IEEE Trans. Smart Grid 12 (2020) 43–55.
[6] C. Zhang, T. Yang, Y. Wang, Peer-to-peer energy trading in a microgrid based on iterative double auction and blockchain, Sustain. Energy Grids Netw. 27 (2021) 100524.
[7] A. Ghasemi, M. Enayatzare, Optimal energy management of a renewable-based isolated microgrid with pumped-storage unit and demand response, Renew. Energy 123 (2018) 460–474.
[8] C. Marino, M.A. Quddus, M. Marufuzzaman, M. Cowan, A.E. Bednar, A chance-constrained two-stage stochastic programming model for reliable microgrid operations under power demand uncertainty, Sustain. Energy Grids Netw. 13 (2018) 66–77.
[9] B. Sang, T. Zhang, Y. Liu, L. Liu, Z. Shi, Two-stage robust optimal scheduling of cooperative microgrids based on expected scenarios, IET Gener. Transm. Distribution 14 (2021) 6741–6753.
[10] M.S. Misaghian, M. Saffari, M. Kia, M.S. Nazar, A. Heidari, M. Shafie-khah, et al., Hierarchical framework for optimal operation of multiple microgrids considering demand response programs, Electr. Power Syst. Res. 165 (2018) 199–213.
[11] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, Robust optimization, in: Robust Optimization, Princeton University Press, 2009.
[12] H. Karimi, S. Jadid, Two-stage economic, reliability, and environmental scheduling of multi-microgrid systems and fair cost allocation, Sustain. Energy Grids Netw. 28 (2021) 100546.
[13] S. Bairamzadeh, M. Saidi-Mehrabad, M.S. Pishvaee, Modelling different types of uncertainty in biofuel supply network design and planning: A robust optimization approach, Renew. Energy 116 (2018) 500–517.
[14] P. Li, Q. Wu, M. Yang, Z. Li, N.D. Hatziargyriou, Distributed distributionally robust dispatch for integrated transmission-distribution systems, IEEE Trans. Power Syst. 36 (2020) 1193–1205.
[15] S. Mohseni, M.S. Pishvaee, A robust programming approach towards design and optimization of microalgae-based biofuel supply chain, Comput. Ind. Eng. 100 (2016) 58–71.
[16] S. Mohseni, M.S. Pishvaee, H. Sahebi, Robust design and planning of microalgae biomass-to-biodiesel supply chain: A case study in Iran, Energy 111 (2016) 736–755.
[17] M.S. Pishvaee, S. Mohseni, S. Bairamzadeh, Chapter 4 - uncertainties in biofuel supply chain, in: M.S. Pishvaee, S. Mohseni, S. Bairamzadeh (Eds.), Biomass to Biofuel Supply Chain Design and Planning under Uncertainty, Academic Press, 2021, pp. 65–93.
[18] X. Zhou, Q. Ai, M. Yousif, Two kinds of decentralized robust economic dispatch framework combined distribution network and multi-microgrids, Appl. Energy 253 (2019) 113588.
[19] Y. Fu, Z. Zhang, Z. Li, Y. Mi, Energy management for hybrid AC/DC distribution system with microgrid clusters using non-cooperative game theory and robust optimization, IEEE Trans. Smart Grid 11 (2020) 1510–1525.
[20] M. Movahednia, H. Karimi, S. Jadid, Optimal hierarchical energy management scheme for networked microgrids considering uncertainties,demand response, and adjustable power, IET Gener. Transm. Distribution 14 (2020) 4352–4362.
[21] F.S. Gazijahani, J. Salehi, Reliability constrained two-stage optimization of multiple renewable-based microgrids incorporating critical energy peak pricing demand response program using robust optimization approach, Energy 161 (2018) 999–1015.

[22] H. Qiu, B. Zhao, W. Gu, R. Bo, Bi-level two-stage robust optimal scheduling for AC/DC hybrid multi-microgrids, IEEE Trans. Smart Grid 9 (2018) 5455–5466.
[23] D. Bertsimas, M. Sim, The price of robustness, Oper. Res. 52 (2004) 35–53.
[24] S. Mohseni, M.S. Pishvaee, Data-driven robust optimization for wastewater sludge-to-biodiesel supply chain design, Comput. Ind. Eng. (2020) 139.
[25] C. Ning, F. You, Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming, Comput. Chem. Eng. 125 (2019) 434–448.
[26] H. Rahimian, S. Mehrotra, Distributionally robust optimization: A review, 2019, arXiv preprint arXiv:190805659.
[27] F. Lin, X. Fang, Z. Gao, Distributionally robust optimization: A review on theory and applications, Num. Algebra, Control Optim. 12 (2022) 159.
[28] F.H. Aghdam, N.T. Kalantari, B. Mohammadi-Ivatloo, A chance-constrained energy management in multi-microgrid systems considering degradation cost of energy storage elements, J. Energy Storage 29 (2020) 101416.
[29] Z. Bagheri, M. Doostizadeh, F. Aminifar, A receding horizon data-driven chance-constrained approach for energy flexibility trading in multi-microgrid distribution network, IET Renew. Power Gener. 15 (2021) 2860–2877.
[30] L. Li, X. Cao, P. Wang, Optimal coordination strategy for multiple distributed energy systems considering supply demand, and price uncertainties, Energy 227 (2021) 120460.
[31] J. Zhai, Y. Jiang, Y. Shi, C.N. Jones, X.-P. Zhang, Distributionally robust joint chance-constrained dispatch for integrated transmission-distribution systems via distributed optimization, IEEE Trans. Smart Grid 13 (2022) 2132–2147.
[32] C.-S. Karavas, G. Kyriakarakos, K.G. Arvanitis, G. Papadakis, A multi-agent decentralized energy management system based on distributed intelligence for the design and control of autonomous polygeneration microgrids, Energy Convers. Manage. 103 (2015) 166–179.
[33] B. Zhang, Q. Li, L. Wang, W. Feng, Robust optimization for energy transactions in multi-microgrids under uncertainty, Appl. Energy 217 (2018) 346–360.
[34] Z. Liang, Q. Alsafasfeh, W. Su, Proactive resilient scheduling for networked microgrids with extreme events, IEEE Access 7 (2019) 112639-52.
[35] R.S. Pinto, C. Unsihuay-Vila, F.H. Tabarro, Coordinated operation and expansion planning for multiple microgrids and active distribution networks under uncertainties, Appl. Energy 297 (2021) 117108.
[36] R. Nourollahi, P. Salyani, K. Zare, R. Razzaghi, A two-stage hybrid robust-stochastic day-ahead scheduling of transactive microgrids considering the possibility of main grid disconnection, Int. J. Electr. Power Energy Syst. 136 (2022) 107701.
[37] M. Jalali, K. Zare, H. Seyedi, M. Alipour, F. Wang, Distributed model for robust real-time operation of distribution systems and microgrids, Electr. Power Syst. Res. 177 (2019) 105985.
[38] X. Chang, Y. Xu, H. Sun, I. Khan, A distributed robust optimization approach for the economic dispatch of flexible resources, Int. J. Electr. Power Energy Syst. 124 (2021) 106360.
[39] Y. Liu, Y. Li, H.B. Gooi, Y. Jian, H. Xin, X. Jiang, et al., Distributed robust energy management of a multimicrogrid system in the real-time energy market, IEEE Trans. Sustain. Energy 10 (2017) 396–406.
[40] M. Mohiti, H. Monsef, A. Anvari-moghaddam, J. Guerrero, H. Lesani, A decentralized robust model for optimal operation of distribution companies with private microgrids, Int. J. Electr. Power Energy Syst. 106 (2019) 105–123.
[41] K. Zhang, S. Troitzsch, Robust scheduling for networked microgrids under uncertainty, Front. Energy Res. (2021) 9.
[42] L. Wang, C. Jiang, K. Gong, R. Si, H. Shao, W. Liu, Data-driven distributionally robust economic dispatch for distribution network with multiple microgrids, IET Gener. Transm. Distribution 14 (2020) 5712–5719.
[43] Q. Xu, T. Zhao, Y. Xu, Z. Xu, P. Wang, F. Blaabjerg, A distributed and robust energy management system for networked hybrid AC/DC microgrids, IEEE Trans. Smart Grid 11 (2020) 3496–3508.
[44] Y. Liu, L. Guo, C. Wang, A robust operation-based scheduling optimization for smart distribution networks with multi-microgrids, Appl. Energy 228 (2018) 130–140.
[45] L. Wang, Z. Zhu, C. Jiang, Z. Li, Bi-level robust optimization for distribution system with multiple microgrids considering uncertainty distribution locational marginal price, IEEE Trans. Smart Grid 12 (2020) 1104–1117.
[46] Z. Wu, Z. Xu, W. Gu, S. Zhou, X. Yang, Decentralized game-based robustly planning scheme for distribution network and microgrids considering bilateral energy trading, IEEE Trans. Sustain. Energy 13 (2022) 803–817.
[47] C. Wei, Z. Shen, D. Xiao, L. Wang, X. Bai, H. Chen, An optimal scheduling strategy for peer-to-peer trading in interconnected microgrids based on RO and Nash bargaining, Appl. Energy 295 (2021) 117024.
[48] M. Staib, S. Jegelka, Distributionally robust optimization and generalization in kernel methods, Adv. Neural Inf. Process. Syst. (2019) 32.
[49] H. Gao, J. Liu, L. Wang, Z. Wei, Decentralized energy management for networked microgrids in future distribution systems, IEEE Trans. Power Syst. 33 (2018) 3599–3610.
[50] T. Mühlpfordt, X. Dai, A. Engelmann, V. Hagenmeyer, Distributed power flow and distributed optimization—Formulation solution, and open source implementation, Sustain. Energy Grids Netw.. 26 (2021) 100471.
[51] Z. Li, S. Su, X. Jin, H. Chen, Distributed energy management for active distribution network considering aggregated office buildings, Renew. Energy 180 (2021) 1073–1087.
[52] S. Boyd, N. Parikh, E. Chu, Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers, Now Publishers Inc, 2011.
[53] B. Houska, J. Frasch, M. Diehl, An augmented Lagrangian based algorithm for distributed nonconvex optimization, SIAM J. Optim. 26 (2016) 1101–1127.
[54] C. Ning, F. You, A data-driven multistage adaptive robust optimization framework for planning and scheduling under uncertainty, AIChE J. 63 (2017) 4343–4369.
[55] C. Ning, F. You, Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods, Comput. Chem. Eng. 112 (2018) 190–210.
[56] N. Meyer-Huebner, M. Suriyah, T. Leibfried, Distributed optimal power flow in hybrid AC–DC grids, IEEE Trans. Power Syst. 34 (2019) 2937–2946.
[57] Q. Xu, Y. Xu, Z. Xu, L. Xie, F. Blaabjerg, A hierarchically coordinated operation and control scheme for DC microgrid clusters under uncertainty, IEEE Trans. Sustain. Energy 12 (2021) 273–283.
[58] A. Hussain, V.-H. Bui, H.-M. Kim, Robust optimization-based scheduling of multi-microgrids considering uncertainties, Energies 9 (2016) 278.
[59] A. Hussain, V.-H. Bui, H.-M. Kim, Robust optimal operation of AC/DC hybrid microgrids under market price uncertainties, IEEE Access 6 (2017) 2654–2667.
[60] M.E. Baran, F.F. Wu, Optimal capacitor placement on radial distribution systems, IEEE Trans. Power Deliv. 4 (1989) 725–734.
[61] O. Arslan, O.E. Karasan, Cost and emission impacts of virtual power plant formation in plug-in hybrid electric vehicle penetrated networks, Energy 60 (2013) 116–124.
[62] J. Kim, C.D. Scott, Robust kernel density estimation, J. Mach. Learn. Res. 13 (2012) 2529–2565.
[63] P.J. Huber, Robust Statistics, John Wiley & Sons, 2004.
[64] F.R. Hampel, The influence curve and its role in robust estimation, J. Amer. Statist. Assoc. 69 (1974) 383–393.
[65] A. Ben-Tal, A. Goryashko, E. Guslitzer, A. Nemirovski, Adjustable robust solutions of uncertain linear programs, Math. Program. 99 (2004) 351–376.
[66] D. Xu, B. Zhou, N. Liu, Q. Wu, N. Voropai, C. Li, et al., Peer-to-peer multi-energy and communication resource trading for interconnected microgrids, IEEE Trans. Ind. Inform. 17 (2020) 2522–2533.
[67] W. Thomson, Cooperative models of bargaining, Handb. Game Theory Econ. Appl. 2 (1994) 1237–1284.
[68] J. Li, C. Zhang, Z. Xu, J. Wang, J. Zhao, Y-JA. Zhang, Distributed transactive energy trading framework in distribution networks, IEEE Trans. Power Syst. 33 (2018) 7215–7227.
[69] E. Münsing, J. Mather, S. Moura, Blockchains for decentralized optimization of energy resources in microgrid networks, in: 2017 IEEE conference on control technology and applications, Vol. 216, CCTA, IEEE, 2017, pp. 4–71.
[70] Z. Liu, L. Wang, L. Ma, A transactive energy framework for coordinated energy management of networked microgrids with distributionally robust optimization, IEEE Trans. Power Syst. 35 (2019) 395–404.
[71] Z. Wang, B. Chen, J. Wang, Decentralized energy management system for networked microgrids in grid-connected and islanded modes, IEEE Trans. Smart Grid 7 (2015) 1097–1105.
[72] H. Gao, J. Liu, L. Wang, Z. Wei, Decentralized energy management for networked microgrids in future distribution systems, IEEE Trans. Power Syst. 33 (2017) 3599–3610.
[73] M. Mazidi, A. Zakariazadeh, S. Jadid, P. Siano, Integrated scheduling of renewable generation and demand response programs in a microgrid, Energy Convers. Manage. 86 (2014) 1118–1127.
[74] J. Li, M.E. Khodayar, J. Wang, B. Zhou, Data-driven distributionally robust co-optimization of P2P energy trading and network operation for interconnected microgrids, IEEE Trans. Smart Grid (2021).
[75] M.E. Baran, F.F. Wu, Network reconfiguration in distribution systems for loss reduction and load balancing, IEEE Power Eng. Rev. 9 (1989) 101–102.

[76] A.L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, Oper. Res. 21 (1973) 1154–1157.
[77] C. Shang, F. You, Distributionally robust optimization for planning and scheduling under uncertainty, Comput. Chem. Eng. 110 (2018) 53–68.
[78] M.S. Pishvaee, M. Rabbani, S.A. Torabi, A robust optimization approach to closed-loop supply chain network design under uncertainty, Appl. Math. Model. 35 (2011) 637–649.
[79] H. Qiu, W. Gu, W. Sheng, L. Wang, Q. Sun, Z. Wu, Resilience-oriented multistage scheduling for power grids considering nonanticipativity under tropical cyclones, IEEE Trans. Power Syst. (2022).