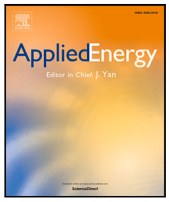
Applied Energy 326 (2022) 119939

Contents lists available at ScienceDirect

Applied Energy

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



Data-driven distributionally robust joint chance-constrained energy management for multi-energy microgrid ☆

Junyi Zhai [a,b], Sheng Wang [c], Lei Guo [b,d], Yuning Jiang [e,*], Zhongjian Kang [a], Colin N. Jones [e]

[a] College of New Energy, China University of Petroleum (East China), Qingdao, China,
[b] State Grid (Suzhou) City & Energy Research Institute, Suzhou, China
[c] State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau, China
[d] State Grid Energy Research Institute, Beijing, China
[e] Automatic Control Laboratory, EPFL, Switzerland

ARTICLE INFO

Keywords:
Multi-energy microgrid (MEMG)
Distributionally robust joint chance-constrained (DRJCC)
Optimized conditional value-at-risk (CVaR) approximation (OCA)
Sequential convex optimization

ABSTRACT

Multi-energy microgrid (MEMG) has the potential to improve the energy utilization efficiency. However, the uncertainty caused by distributed renewable energy resources brings an urgent need for multi-energy co-optimization to ensure secure operation. This paper focuses on the distributionally robust energy management problem for MEMG. Various flexible resources in different energy sectors are utilized for uncertainty mitigation, then, a data-driven Wasserstein distance-based distributionally robust joint chance-constrained (DRJCC) energy management model is proposed. To make the DRJCC model tractable, an optimized conditional value-at-risk (CVaR) approximation (OCA) formulation is proposed to transfer the joint chance-constrained model into a tractable form. Then, an iterative sequential convex optimization algorithm is tailored to reduce the solution conservatism by tuning OCA. Numerical result illustrates the effectiveness of the proposed model.

The main notations and symbols are listed here. The remaining are defined later when they first appear. Boldface lower case and upper case letters represent vectors and matrices, respectively, and $\mathbf{1}_n \in \mathbb{R}^n$ denotes the vector of ones with dimension n. $\mathbb{Z}_{z_1}^{z_2} = \{z \in \mathbb{Z} | z_1 \leq z \leq z_2\}$ denotes integer ranges. And $\mathbb{E}_{\mathbb{P}}$ denotes the expectation over distribution $\mathbb{P}$. For a given set $\mathcal{X}$, we use notation $\mathbb{1}_{\mathcal{X}}(x) = 1$ if $x \in \mathcal{X}$ and $\mathbb{1}_{\mathcal{X}}(x) = 0$, otherwise. For a random vector $\boldsymbol{\xi}$ governed by a distribution $\mathbb{P}$, we define the conditional value-at-risk at level $\epsilon \in (0, 1)$ of a measurable loss function $L(\boldsymbol{\xi})$ by

$$\mathbb{P}\text{-CVaR}_{\epsilon}\left[L(\xi)\right] := \min_{\tau} \left\{\tau + \mathbb{E}_{\mathbb{P}}\left[\max(0, L(\boldsymbol{\xi}) - \tau)\right]/\epsilon\right\}. \tag{1}$$

1. Introduction

In the past decades, distributed energy resources (DERs) have developed rapidly. However, DERs have the characteristics of intermittent and random, which has a great impact on the safe operation of the distribution network [1]. To address this challenge, microgrids are gaining more and more attention as an effective solution to increase the integration rate of renewable energy. Compared with traditional microgrids, multi-energy microgrids (MEMG) aim to integrate multiple energy carriers such as electricity, heat, and cooling energy to achieve higher energy efficiency.

Recently, some methodologies have been conducted to plan and manage MEMG or multi-energy systems. The dynamic economic dispatch model and temporally-coordinated operation model of MEMG aming for improving the operating efficiency is respectively proposed in [2,3]. A stochastic deployment strategy considering demand response for residential MEMG is aproposed in [4]. An optimal configuration with respect to capacity sizes and types of DERs for MEMG is presented in [5]. An interval-based planning model for MEMG with multiple uncertainties is proposed in [6]. The robustly coordinated operation models for grid-connected and islanded MEMG with flexible loads are proposed in [7,8]. Meanwhile, a distributionally robust optimization (DRO) model for electricity and heating networks is presented in [9]. However, the uncertainties propagated from the electricity system to other energy system via the coupling facilities are not fully modeled in these models. This motivates us to design a strategy to fully utilize various flexible resources in different energy sectors of MEMG for uncertainty mitigation.

☆ This work was supported in part from the Natural Science Foundation of Jiangsu Province, China (BK20210103) and the Swiss National Science Foundation under the RISK project (Risk Aware Data Driven Demand Response, grant number 200021 175627)

* Corresponding author.
E-mail addresses: zhaijunyi@upc.edu.cn (J. Zhai), wangsheng_zju@zju.edu.cn (S. Wang), guolei@sgcc.com.cn (L. Guo), yuning.jiang@epfl.ch (Y. Jiang), kangzjzh@163.com (Z. Kang), colin.jones@epfl.ch (C.N. Jones).

https://doi.org/10.1016/j.apenergy.2022.119939
Received 17 July 2022; Received in revised form 27 August 2022; Accepted 3 September 2022
Available online 19 September 2022

Nomenclature

Sets

$\mathcal{A}_i$	Set of parent node of node i
$\mathcal{BS}$	Set of nodes with BSs
$\mathcal{CCHP}$	Set of nodes with CCHPs
$\mathcal{C}_i$	Set of children nodes of node i
$\mathcal{EB}$	Set of nodes with EBs
$\mathcal{EC}$	Set of nodes with ECs
$\mathcal{L}$	Set of power lines, indexed by $\mathcal{N}$
$\mathcal{N}$	Set of non-root nodes
$\mathcal{TS}$	Set of nodes with TSs
$\mathcal{T}$	Set of time periods
$\mathcal{WT}$	Set of nodes with WTs

Parameters

η_i^{loss}	Heat loss factor of MT
η_i^{BS}/η_i^{TS}	Storing/releasing efficiency of BS/TS
$\eta_i^{MT}/\eta_i^{AC}/\eta_i^{EC}/\eta_i^{EB}$	Efficiency of MT/AC/EC/EB
$\hat{P}_{it}^{WT}/\hat{Q}_{it}^{WT}$	Forecasted active/reactive power of WT
$\overline{H}_i^C/\overline{H}_i^D$	Maximum storing/releasing rate of TS
$\overline{P}^T$	Maximum active power exchange with upper-level grid
$\overline{P}_i^{EB}/\overline{P}_i^{EC}/\overline{H}_i^{AC}$	Maximum output of EB/EC/AC
$\overline{u}_i/\overline{d}_i$	Maximum upward/downward reserve adjustment of MT
ρ^{LHV}	Low heating value of natural gas
$\underline{E}_i^{BS}/\overline{E}_i^{BS}$	Minimum/maximum electricity stored in BS
$\underline{E}_i^{TS}/\overline{E}_i^{TS}$	Minimum/maximum heating energy stored in TS
$\underline{P}_{i,t}^{MT}/\overline{P}_{i,t}^{MT}$	Lower/upper active power of MT
$\underline{Q}_{i,t}^{MT}/\overline{Q}_{i,t}^{MT}$	Lower/upper reactive power of MT
$\underline{V}_i/\overline{V}_i$	Lower/upper nodal voltage magnitude
φ	Constant power factor
ϑ^{heat}	Heating coefficient
c^O	Operation and maintenance cost of energy storage facilities
c^{MT}	Fuel cost of MTs
c_t^B/c_t^S	Electricity purchasing/selling price
H_t^{HL}/H_t^{CL}	Heating/cooling load
P_{it}^L/Q_{it}^L	Nodal active/reactive load
r_i/x_i	Resistance/reactance of power line
$V_{0,t}$	Voltage magnitude of root node

Random Variables

$\xi_{j,t}^P/\xi_{j,t}^Q$	Active/reactive power forecasting error of WT, and $\xi_{j,t}^Q = \sqrt{(1-\varphi^2)/\varphi^2}\xi_{j,t}^P$

Decision Variables

$\alpha_{i,t}^{AC}/\alpha_{i,t}^{BS}/\alpha_{i,t}^{TS}$	Adjustment factors of AC/BS/TS
$\alpha_{i,t}^{BSC}/\alpha_{i,t}^{BSD}$	Adjustment factors of storing/releasing rate of BS
$\alpha_{i,t}^{MT}/\alpha_{i,t}^{EB}/\alpha_{i,t}^{EC}$	Adjustment factors of MT/EB/EC/BS
$\alpha_{i,t}^{TSC}/\alpha_{i,t}^{TSD}$	Adjustment factors of storing/releasing rate of TS
$\hat{E}_{i,t}^{BS}/\hat{E}_{i,t}^{TS}$	BS/TS stored energy under nominal state
$\hat{H}_{i,t}^{TSC}/\hat{H}_{i,t}^{TSD}$	TS storing/releasing rate under nominal state
$\hat{P}_{i,t}/\hat{Q}_{i,t}$	Active/reactive power flow of power line under nominal state
$\hat{P}_{i,t}^{BSC}/\hat{P}_{i,t}^{BSD}$	BS storing/releasing rate under nominal state
$\hat{P}_{i,t}^{EB}/\hat{P}_{i,t}^{EC}/\hat{H}_{i,t}^{AC}$	Power/heat consumption of EB/EC/AC under nominal state
$\hat{P}_{i,t}^{MT}/\hat{Q}_{i,t}^{MT}$	MT active/reactive power under nominal state
$\hat{V}_{i,t}/V_{i,t}$	Nodal voltage magnitude under nominal state/under uncertainty
$E_{i,t}^{BS}/E_{i,t}^{TS}$	BS/TS stored energy under uncertainty
$H_{i,t}^{TSC}/H_{i,t}^{TSD}$	TS storing/releasing rate under uncertainty
P_t^B/P_t^S	Electricity buying/selling
$P_{i,t}/Q_{i,t}$	Active/reactive power of distribution line under uncertainty
$P_{i,t}^{BSC}/P_{i,t}^{BSD}$	BS storing/releasing rate under uncertainty
$P_{i,t}^{EB}/P_{i,t}^{EC}/H_{i,t}^{AC}$	Power/heat consumption of EB/EC/AC under uncertainty
$P_{i,t}^{MT}/Q_{i,t}^{MT}$	MT active/reactive power under uncertainty

Uncertainties from renewable energy can pose significant challenges and complicate the energy management of MEMG. DRO is an effective way to handle diverse uncertainties, which incorporates the available probability distribution information into an ambiguity set to characterize the true probability distribution of uncertainties and reduces the solution conservatism of robust optimization (RO) [10,11]. Presently, DRO adopts two main types of ambiguity sets, that is, moment-based [9,12] and metric-based [13–16] ambiguity sets. Recently, the moment-based DRO models for multi-energy systems were presented in [9,12], which comprise all probability distribution informations with an identical mean and covariance. Yet, only the first two moments do not constitute a detailed characterization of the true probability distribution. Especially when we have a lot of data at hand, more probabilistic distribution information can be extracted and utilized than just the first two moments. In contrast, the metric-based DRO models can take full advantage of the existing historial data at the cost of the increased computational burden. In addition, the decision-maker can also regard the radius of the metric-based ambiguity set as a tuning parameter to adjust their risk attitude towards the system operation. Recently, metric-based DRO models have attracted much attention in power systems, including optimal power flow (OPF) [14], unit commitment [15], and home energy management [16] problems. However, the aforementioned moment-based or metric-based DRO models [9,12,14–17] mainly focus on the distributionally robust individual chance-constrained (DRICC) problems.

Different from the DRICC model, the distributionally robust joint chance-constrained (DRJCC) model can provide stronger guarantees for overall power systems safety with high probability by enforcing multiple safety constraints to be satisfied simultaneously [18]. That is, DRJCC is more expressive and can overcome the drawback that solutions of DRICC to achieve joint feasibility are overly conservative [19–21]. However, handling DRJCC models are mathematically harder than DRICC models [22]. The standard way for approximating a joint chance constraint is to decompose it into multiple individual chance constraints using Bonferroni's inequality. Moreover, the sufficient condition for ensuring the feasibility of joint chance-constrained problem is to ensure that the total sum of violation probabilities of multiple individual chance constraints is less than that of the joint chance constraint. However, there is no systematic way to choose a

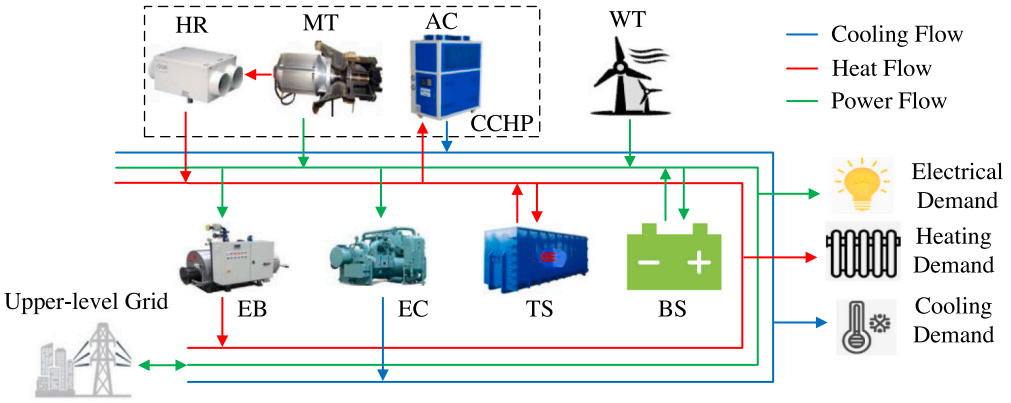


Fig. 1. Multi-energy microgrid demonstration.

better allocation of risk factors among individual chance constraints. The natural choice is to divide the constraint violation probability equally among the multiple individual chance constraints [23]. Unfortunately, even if the individual chance constraints are independent, the Bonferroni's inequality is only an approximation at best. In the events when the individual chance constraints are correlated, the Bonferroni's inequality may be more conservative [22].

Recently, a few up to date moment-based or metric-based DRJCC single-period dispatch models for power transmission system are proposed [18,19,21,24]. In [18,24], a Wasserstein metric-based DRJCC dispatch model based on the Bonferroni approximation (BA) formulation is proposed. In [21], the moment-based DRJCC DC-OPF model is proposed, where the optimized BA method is presented to reduce the solution conservatism of Bonferroni's inequality. In [19], the Bonferroni's inequality and the optimized CVaR tractable approximations for the metric-based joint chance-constrained real-time dispatch problem are proposed. Even so, the DRJCC models applied to power systems are still limited due to the great challenge of handling joint chance constraints. To the best of authors' knowledge, the merits of DRJCC have not been investigated in multi-energy systems.

Additionally, when the samples are drawn independent and identically distributed (i.i.d.) from an underlying probability distribution, some existing researches have established rigorous bounds for Wasserstein metric-based DRO models between the empirical probability distribution and the data-generating probability distribution, such as the cross validation method [14] and the holdout method [25]. However, the renewable energy data is not necessarily being drawn from any underlying probability distribution in an independent manner. Thus, the i.i.d. assumption for renewable energy data may be too strong so that it may be violated in practice [24]. The above motivates us not only to look for a better approximation for joint chance-constrained problems, but also enable further incorporation of spatial–temporal correlations among uncertain renewable energy at different nodes and periods.

To address the research gap, this paper contributes in the following aspects:

(1) A data-driven Wasserstein metric-based DRJCC energy management model for MEMG is proposed to account for complicated spatial–temporal correlations among uncertain renewable energy. Under the joint feature of distributionally robust chance constraints, various flexible resources in different energy sectors is utilized for uncertainty mitigation. Then, a unified linear decision rule model is adopted to facilitate further reformulation of joint chance constraints.
(2) An optimized conditional value at risk (CVaR) approximation (OCA) formulation for joint chance constraints is presented to transfer the data-driven DRJCC energy management model into a tractable form. The tightness of the OCA formulation depends on a set of scaling parameters. Then, an iterative sequential convex algorithm is tailored to tune the OCA to reduce the solution conservatism.

The rest of this paper is organized as follows: Section 2 shows the DRJCC energy management formulation for MEMG. Section 3 describes the data-driven tractable reformulation. Case study and conclusion are presented in Sections 4 and 5.

2. DRJCC energy management formulation

This section describes the structure of MEMG and presents the DRJCC energy management model for MEMG which can utilize various flexible resources in different energy sectors for uncertainty mitigation. A unified linear decision rule (ULDR) model is also presented to facilitate further reformulation of joint chance constraints.

2.1. System structure

A typical MEMG structure is presented in Fig. 1. The MEMG has a radial topology based multi-energy supply system, integrating renewable energy (e.g. wind turbine (WT)), battery storage (BS), thermal storage (TS), electric boiler (EB), electric chiller (EC), and combined cooling, heat, and power (CCHP) plant to satisfy electric, heat, and cooling loads simultaneously. The CCHP plant generally includes three parts: a micro-turbine (MT), an absorption chiller (AC) and the heat recovery (HR) system. Since long-distance thermal transmission will lead to great thermal loss, the thermal energy is supplied locally within a geographical region and heat networks among different thermal groups are not considered [2–4,6–8].

2.2. Unified linear decision rule for uncertain energy flow

In practice, renewable energy is usually hard to predict accurately a day ahead. The forecast errors need to be compensated in the real-time stage by adjusting the flexible resources. Linear decision rule (LDR) is widely used in DRO decision-making, which can offer tractable and equivalent reformulations under quite a few ambiguity sets [14,15,18]. Meanwhile, from application perspective, LDR is applicable in many practical decision-making frameworks of electricity industry due to its simplicity [26,27]. In order to facilitate further reformulation of the following DRJCC energy management model, ULDR model is adopted to exploit various flexible resources in MEMG for uncertainty mitigation, shown in Fig. 2.

The electricity facilities of MT, EB, EC, and BS in MEMG are selected to adjust their power output based on the following ULDR. And (2d) guarantees that the renewable energy forecasting errors are fully mitigated in MEMG.

$$P_{i,t}^{MT} = \hat{P}_{i,t}^{MT} - \alpha_{i,t}^{MT} \sum_{j \in \mathcal{WT}} \xi_{j,t}^{P}, \ i \in \mathcal{CCHP}, \ t \in \mathcal{T} \tag{2a}$$

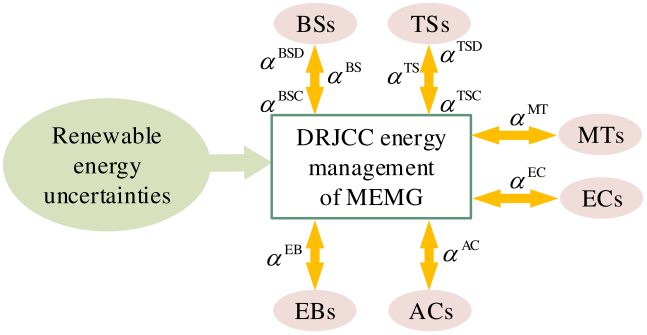


Fig. 2. ULDR framework for DRJCC energy management model.

$$Q_{i,t}^{MT} = \hat{Q}_{i,t}^{MT} - \alpha_{i,t}^{MT} \sum_{j\in\mathcal{WT}} \xi_{j,t}^{Q},\ i \in CCHP,\ t \in \mathcal{T} \tag{2b}$$

$$P_{i,t}^{*} = \hat{P}_{i,t}^{*} + \alpha_{i,t}^{*} \sum_{j\in\mathcal{WT}} \xi_{j,t}^{P},\ *\in \{\mathcal{EB},\mathcal{EC},\mathcal{BSC},\mathcal{BSD}\},$$
$$i \in \mathcal{EB}\cup\mathcal{EC}\cup\mathcal{BS},\ t \in \mathcal{T} \tag{2c}$$

$$1 = \sum_{i\in CCHP} \alpha_{i,t}^{MT} + \sum_{i\in\mathcal{EB}} \alpha_{i,t}^{EB} + \sum_{i\in\mathcal{EC}} \alpha_{i,t}^{EC} + \sum_{i\in\mathcal{BS}} \left(\alpha_{i,t}^{BSC} + \alpha_{i,t}^{BSD}\right),\ t \in \mathcal{T} \tag{2d}$$

$$\alpha_{i,t}^{MT},\ \alpha_{i,t}^{EB},\ \alpha_{i,t}^{EC},\ \alpha_{i,t}^{BSC},\ \alpha_{i,t}^{BSD} \in [0,1] \tag{2e}$$

The linearized DistFlow model for radial distribution network is adopted to describe the AC power flows and ensure the nodal voltage security under uncertainties. The uncertain power line flow affected by the renewable energy injections is derived as

$$P_{i,t} = \hat{P}_{i,t} - \boldsymbol{B}_{i*}\left(\xi_t^{P} - \boldsymbol{\alpha}_t \sum_{j\in\mathcal{N}} \xi_{j,t}^{P}\right) \tag{3a}$$

$$Q_{i,t} = \hat{Q}_{i,t} - \boldsymbol{B}_{i*}\left(\xi_t^{Q} - \boldsymbol{\alpha}_t \sum_{j\in\mathcal{N}} \xi_{j,t}^{Q}\right) \tag{3b}$$

for all $i \in \mathcal{L}$ and $t \in \mathcal{T}$. Here, ξ_t^{P} and ξ_t^{Q} denote the active and reactive power forecasting errors of renewable energy on non-root nodes. $\boldsymbol{\alpha}_t$ denotes the adjustment factors of non-root nodes, if node i has no DERs, then $\alpha_{i,t} = 0$. $\boldsymbol{B}_{i*}$ denotes the ith row of $\boldsymbol{B} \in \mathbb{R}^{|\mathcal{L}|\times|\mathcal{N}|}$ with $b_{(ij)} = 1$ if power line i is part of the path from root node to node j and $b_{(ij)} = 0$, otherwise.

Accordingly, the nodal voltage magnitude is derived as

$$\begin{aligned} V_{i,t} &= V_{\mathcal{A}_i} - (r_i P_{i,t} + x_i Q_{i,t})/V_{0,t} \\ &= \hat{V}_{i,t} + \boldsymbol{B}_{*i}^{\top}\left[\boldsymbol{R}\boldsymbol{B}(\xi_t^{P} - \boldsymbol{\alpha}_t \sum_{j\in\mathcal{N}} \xi_{j,t}^{P}) + \boldsymbol{X}\boldsymbol{B}(\xi_t^{Q} - \boldsymbol{\alpha}_t \sum_{j\in\mathcal{N}} \xi_{j,t}^{Q})\right]/V_{0,t} \end{aligned} \tag{4}$$

for all $i \in \mathcal{N}$ and $t \in \mathcal{T}$. $\boldsymbol{R}$ and $\boldsymbol{X}$ are $|\mathcal{L}|\times|\mathcal{L}|$ matrices with diagonal entries consisting of power line resistances and reactances respectively: $R_{(ii)} = r_i$, $R_{(ij,i\neq j)} = 0$, $\boldsymbol{X}$ in analogy.

Electricity stored in BS under uncertainty is modeled by

$$E_{i,t}^{BS} = \hat{E}_{i,t}^{BS} + \alpha_{i,t}^{BS} \sum_{j\in\mathcal{WT}} \xi_{j,t}^{P},\ i \in \mathcal{BS},\ t \in \mathcal{T} \tag{5}$$

while $E_{i,t}^{BS}$ by definition is given by

$$E_{i,t}^{BS} = E_{i,t-1}^{BS} + P_{i,t}^{BSC}\eta_i^{BS}\Delta t - P_{i,t}^{BSD}/\eta_i^{BS}\Delta t,\ i \in \mathcal{BS},\ t \in \mathcal{T} \tag{6}$$

Such that combining (2) and (5) yields

$$\hat{E}_{i,t}^{BS} = \hat{E}_{i,t-1}^{BS} + \hat{P}_{i,t}^{BSC}\eta_i^{BS}\Delta t - \hat{P}_{i,t}^{BSD}/\eta_i^{BS}\Delta t,\ i \in \mathcal{BS},\ t \in \mathcal{T} \tag{7a}$$

$$\alpha_{i,t}^{BS} = \alpha_{i,t-1}^{BS} + \alpha_{i,t}^{BSC}\eta_i^{BS} + \alpha_{i,t}^{BSD}/\eta_i^{BS},\ i \in \mathcal{BS},\ t \in \mathcal{T} \tag{7b}$$

It is clear that making (6) be satisfied is equivalent to enforcing (7). According to [28], the binary variables reflecting that the energy storage system cannot be charged and discharged simultaneously can be omitted.

Uncertainties from renewable energy will propagate from the power network to the heat and cooling network via HR, EB, and EC facilities. Then, the TS and AC within the heat and cooling system can be utilized for uncertainty mitigation.

The heat consumption of AC and heat storing and releasing rates of TS under uncertainty are modeled by

$$H_{i,t}^{*} = \hat{H}_{i,t}^{*} + \alpha_{i,t}^{*} \sum_{j\in\mathcal{WT}} \xi_{j,t}^{P},\ *\in \{\mathcal{AC},\mathcal{TSC},\mathcal{TSD}\},$$
$$i \in \mathcal{AC}\cup\mathcal{TSC}\cup\mathcal{TSD},\ t \in \mathcal{T} \tag{8}$$

Similarly, the heat energy stored in the TS under uncertainty is modeled by

$$E_{i,t}^{TS} = \hat{E}_{i,t}^{TS} + \alpha_{i,t}^{TS} \sum_{j\in\mathcal{WT}} \xi_{j,t}^{P},\ i \in \mathcal{TS},\ t \in \mathcal{T} \tag{9}$$

while $E_{i,t}^{TS}$ is defined by

$$E_{i,t}^{TS} = E_{i,t-1}^{TS} + H_{i,t}^{TSC}\eta_i^{TS}\Delta t - H_{i,t}^{TSD}/\eta_i^{TS}\Delta t,\ i \in \mathcal{TS},\ t \in \mathcal{T} \tag{10}$$

Such that combining (8) and (9) yields

$$\hat{E}_{i,t}^{TS} = \hat{E}_{i,t-1}^{TS} + \hat{H}_{i,t}^{TSC}\eta_i^{TS}\Delta t - \hat{H}_{i,t}^{TSD}/\eta_i^{TS}\Delta t,\ i \in \mathcal{TS},\ t \in \mathcal{T} \tag{11a}$$

$$\alpha_{i,t}^{TS} = \alpha_{i,t-1}^{TS} + \alpha_{i,t}^{TSC}\eta_i^{TS} - \alpha_{i,t}^{TSD}/\eta_i^{TS},\ i \in \mathcal{TS},\ t \in \mathcal{T} \tag{11b}$$

Thus, enforcing (10) is equivalent to enforcing (11).

The heat balance under uncertainty of each thermal group (omit the index) can be expressed as

$$\begin{aligned} &\sum_{i\in CCHP}\left(H_{i,t}^{MT} - H_{i,t}^{AC}\right) + \sum_{i\in\mathcal{TS}}\left(H_{i,t}^{TSD} - H_{i,t}^{TSC}\right) \\ &+ \sum_{i\in\mathcal{EB}} H_{i,t}^{EB} = H_t^{HL},\ t \in \mathcal{T} \end{aligned} \tag{12}$$

where the heat output of MT can be expressed as

$$H_{i,t}^{MT} = \frac{P_{i,t}^{MT}\left(1 - \eta_i^{MT} - \eta_i^{\text{loss}}\right)\vartheta^{\text{heat}}}{\eta_i^{MT}},\ i \in CCHP,\ t \in \mathcal{T} \tag{13}$$

The heat output of EB can be expressed as

$$H_{i,t}^{EB} = \eta_i^{EB} P_{i,t}^{EB}, i \in \mathcal{EB}, t \in \mathcal{T} \tag{14}$$

Combining (2a), (2c), (8), (13), and (14) yields

$$\begin{aligned} &\sum_{i\in CCHP}\left[\frac{\hat{P}_{i,t}^{MT}\left(1 - \eta_i^{MT} - \eta_i^{\text{loss}}\right)\vartheta^{\text{heat}}}{\eta_i^{MT}} - \hat{H}_{i,t}^{AC}\right] \\ &\qquad + \sum_{i\in\mathcal{EB}}\eta_i^{EB}\hat{P}_{i,t}^{EB} + \sum_{i\in\mathcal{TS}}\left(\hat{H}_{i,t}^{TSD} - \hat{H}_{i,t}^{TSC}\right) \\ &= H_t^{HL}, t \in \mathcal{T} \end{aligned} \tag{15a}$$

$$\begin{aligned} &\sum_{i\in CCHP}\left[\frac{-\alpha_{i,t}^{MT}\left(1 - \eta_i^{MT} - \eta_i^{\text{loss}}\right)\vartheta^{\text{heat}}}{\eta_i^{MT}} - \alpha_{i,t}^{AC}\right] \\ &\qquad + \sum_{i\in\mathcal{EB}}\eta_i^{EB}\alpha_{i,t}^{EB} + \sum_{i\in\mathcal{TS}}\left(\alpha_{i,t}^{TSD} - \alpha_{i,t}^{TSC}\right) = 0, t \in \mathcal{T} \end{aligned} \tag{15b}$$

Thus, enforcing (12) is equivalent to enforcing (15).

The cooling balance under uncertainty of each thermal group is expressed as

$$\sum_{i\in\mathcal{EC}} H_{i,t}^{EC} + \sum_{i\in CCHP}\eta_i^{AC} H_{i,t}^{AC} = H_t^{CL}, t \in \mathcal{T} \tag{16}$$

where the cooling output of EC can be expressed as

$$H_{i,t}^{EC} = \eta_i^{EC} P_{i,t}^{EC}, i \in \mathcal{EC}, t \in \mathcal{T} \tag{17}$$

Combining (8) and (17) yields

$$\sum_{i\in\mathcal{EC}} \eta_i^{EC} \hat{P}_{i,t}^{EC} + \sum_{i\in\mathcal{CCHP}} \eta_i^{AC} \hat{H}_{i,t}^{AC} = H_t^{CL}, t \in \mathcal{T} \tag{18a}$$

$$\sum_{i\in\mathcal{EC}} \eta_i^{EC} \alpha_{i,t}^{EC} + \sum_{i\in\mathcal{CCHP}} \eta_i^{AC} \alpha_{i,t}^{AC} = 0, t \in \mathcal{T} \tag{18b}$$

Thus, enforcing (16) is equivalent to enforcing (18).

2.3. DRJCC energy management model

The DRJCC energy management model for MEMG is formulated as

$$\min_{\mathbb{P}\in D} \max \mathbb{E}_{\mathbb{P}} \sum_{t\in\mathcal{T}} \left[\sum_{i\in\mathcal{CCHP}} \frac{c^{MT} P_{i,t}^{MT}}{\rho^{LHV} \eta_i^{MT}} + \sum_{i\in\mathcal{BS}} c^O \left(P_{i,t}^C + P_{i,t}^D \right) + \sum_{i\in\mathcal{TS}} c^O \left(H_{i,t}^C + H_{i,t}^D \right) + c_t^B P_t^B - c_t^S P_t^S \right] \tag{19}$$

subject to

$$0 \le P_t^B \le \overline{P}^T,\ 0 \le P_t^S \le \overline{P}^T,\ t \in \mathcal{T} \tag{20a}$$

$$\hat{P}_{i,t} = \sum_{j\in C_i} \hat{P}_{j,t} - \hat{P}_{i,t}^{MT} - \hat{P}_{i,t}^{WT} + \hat{P}_{i,t}^{EB} + \hat{P}_{i,t}^{EC} + \hat{P}_{i,t}^{C} - \hat{P}_{i,t}^{D} + P_{i,t}^{L}$$

$$\text{with } \hat{P}_{1,t} = P_t^B - P_t^S,\ i \in \mathcal{N},\ t \in \mathcal{T} \tag{20b}$$

$$\hat{Q}_{i,t} = \sum_{j\in C_i} \hat{Q}_{j,t} - \hat{Q}_{i,t}^{MT} - \hat{Q}_{i,t}^{WT} + Q_{i,t}^{L},\ i \in \mathcal{N},\ t \in \mathcal{T} \tag{20c}$$

$$\hat{V}_{i,t} = \hat{V}_{\mathcal{A}_i} - \left(r_i \hat{P}_{i,t} + x_i \hat{Q}_{i,t} \right) / V_{0,t},\ i \in \mathcal{N},\ t \in \mathcal{T} \tag{20d}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[\underline{P}_i^{MT} \le P_{i,t}^{MT} \le \overline{P}_i^{MT},\ i \in \mathcal{CCHP},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_1 \tag{20e}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[-\overline{d}_i \le -\alpha_{i,t}^{MT} \sum_{j\in\mathcal{WT}} \xi_{j,t}^P \le \overline{u}_i,\ i \in \mathcal{CCHP},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_2 \tag{20f}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[0 \le P_{i,t}^{BSC} \le \overline{P}_i^{BSC},\ i \in \mathcal{BS},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_3 \tag{20g}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[0 \le P_{i,t}^{BSD} \le \overline{P}_i^{BSD},\ i \in \mathcal{BS},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_4 \tag{20h}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[\underline{E}_i^{BS} \le E_{i,t}^{BS} \le \overline{E}_i^{BS},\ i \in \mathcal{BS},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_5 \tag{20i}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[\underline{V}_i \le V_{i,t} \le \overline{V}_i,\ i \in \mathcal{N},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_6 \tag{20j}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[0 \le H_{i,t}^{AC} \le \overline{H}_i^{AC},\ i \in \mathcal{CCHP},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_7 \tag{20k}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[\underline{Q}_i^{MT} \le Q_{i,t}^{MT} \le \overline{Q}_i^{MT},\ i \in \mathcal{CCHP},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_8 \tag{20l}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[0 \le P_{i,t}^{EC} \le \overline{P}_i^{EC},\ i \in \mathcal{EC},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_9 \tag{20m}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[0 \le P_{i,t}^{EB} \le \overline{P}_i^{EB},\ i \in \mathcal{EB},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_{10} \tag{20n}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[0 \le H_{i,t}^{TSC} \le \overline{H}_i^{TSC},\ i \in \mathcal{TS},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_{11} \tag{20o}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[0 \le H_{i,t}^{TSD} \le \overline{H}_i^{TSD},\ i \in \mathcal{TS},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_{12} \tag{20p}$$

$$\min_{\mathbb{P}\in D} \mathbb{P} \left[\underline{E}_i^{TS} \le E_{i,t}^{TS} \le \overline{E}_i^{TS},\ i \in \mathcal{TS},\ t \in \mathcal{T} \right] \ge 1 - \epsilon_{13} \tag{20q}$$

$$(2),\ (4),\ (7)\text{–}(9),\ (11), (15),\ (18) \tag{20r}$$

with prescribed violation probabilities $\epsilon_1, \ldots, \epsilon_{13} \in (0, 1)$. Here, we consider the set of time periods given by $\mathcal{T} = \mathbb{Z}_1^T$. In the following, we stack all uncertainties by $\boldsymbol{\xi} = [\boldsymbol{\xi}_1^\top, \ldots, \boldsymbol{\xi}_T^\top]^\top \in \mathbb{R}^{n_s \cdot T}$ with $\boldsymbol{\xi}_t \in \mathbb{R}^{n_s}$ for all $t \in \mathcal{T}$. As the uncertainty distribution $\mathbb{P}$ is in general, not precisely known, this DRJCC model restricts it to lie in the ambiguity set D that defines a family of probability distribution supported by $\Omega \subseteq \mathbb{R}^{n_s \cdot T}$.

Objective function (19) aims to find an optimal solution that minimize the worst-case expected operation costs. (20b)–(20d) denote the DistFlow model under nominal state. (20a) introduces the buying and selling power limit with the upper-level grid. Chance constraints (20e)–(20q) ensure that even under the worst case distribution, MT active power, MT power adjustment, BS charging rate, BS discharging rate, BS stored energy, nodal voltage magnitude, MT reactive power, AC output, EC output, EB output, TS storing rate, TS releasing rate, and TS stored energy are still satisfied with the prescribed probability.

3. Data-driven tractable reformulation

This section proposes an OCA formulation for joint chance constraints to transfer the data-driven Wasserstein metric-based DRJCC energy management model into a tractable form, where complicated spatial–temporal correlations among uncertain renewable energy are considered. Then a sequential convex optimization algorithm is tailored to tune the OCA to reduce the solution conservatism.

To this end, we first rewrite the problem into a dense form. In the following, we use shorthands $\boldsymbol{\alpha}_t \in \mathbb{R}^{n_c}$ and $\boldsymbol{x}_t \in \mathbb{R}^{n_x}$ to stack all adjustment factors and the other decision variables in the Nomenclature for all $t \in \mathcal{T}$, respectively. In a result, substituting the proposed ULDR in Section 2.2 into the DRJCC model in Section 2.3 yields

$$\min_{(\boldsymbol{x},\boldsymbol{\Lambda})\in\Theta} \ \boldsymbol{c}_{0,1}^\top \boldsymbol{x} + \max_{\mathbb{P}\in D} \mathbb{E}_{\mathbb{P}} \left[\boldsymbol{c}_{0,2}^\top \boldsymbol{\Lambda} \boldsymbol{C}_0 \boldsymbol{\xi} \right] \tag{21a}$$

$$\text{s.t.} \quad \min_{\mathbb{P}\in D} \mathbb{P} \left[\boldsymbol{C}_\ell(\boldsymbol{\Lambda})\boldsymbol{\xi} \le \boldsymbol{c}_\ell(\boldsymbol{x}) \right] \ge 1 - \epsilon_\ell,\ \ell \in \mathbb{Z}_1^{13} \tag{21b}$$

with coefficients $\boldsymbol{c}_{0,1}$, $\boldsymbol{c}_{0,2}$ and $\boldsymbol{C}_0$. We stack all decision variables by $\boldsymbol{\Lambda} = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_T] \in \mathbb{R}^{n_c \times T}$ and $\boldsymbol{x} = [\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_T^\top]^\top \in \mathbb{R}^{n_x \cdot T}$ over time $t \in \mathcal{T}$ that satisfy the polyhedral constraints Θ including deterministic constraints (20a)–(20d) and (20r). Here we consider the complicated spatial–temporal correlations among uncertain renewable energy $\xi_{j,t}$ at different nodes and periods, i.e. correlations over nodes $j \in \mathcal{WT}$ and time periods $t \in \mathcal{T}$, but not assume they are i.i.d..

The objective (19) is the sum of a linear function at the reference output under nominal state and a linear function to denote the incremental operation cost under renewable energy uncertainties. Therefore, we separate it from the deterministic term

$$\boldsymbol{c}_{0,1}^\top \boldsymbol{x} := \sum_{t\in\mathcal{T}} \left[\sum_{i\in\mathcal{CCHP}} \frac{c^{MT} \hat{P}_{i,t}^{MT}}{\rho^{LHV} \eta_i^{MT}} + \sum_{i\in\mathcal{BS}} c^O \left(\hat{P}_{i,t}^C + \hat{P}_{i,t}^D \right) + \sum_{i\in\mathcal{TS}} c^O \left(\hat{H}_{i,t}^C + \hat{H}_{i,t}^D \right) + c_t^B P_t^B - c_t^S P_t^S \right], \tag{22}$$

and the worse-case term

$$\boldsymbol{c}_{0,2}^\top \boldsymbol{\Lambda} \boldsymbol{C}_0 \boldsymbol{\xi} := \sum_{t\in\mathcal{T}} \left\{ \left[\sum_{i\in\mathcal{BS}} c^O \left(\alpha_{i,t}^{BSC} - \alpha_{i,t}^{BSD} \right) + \sum_{i\in\mathcal{TS}} c^O \left(\alpha_{i,t}^{TSC} + \alpha_{i,t}^{TSD} \right) - \sum_{i\in\mathcal{CCHP}} \frac{c^{MT} \alpha_{i,t}^{MT}}{\rho^{LHV} \eta_i^{MT}} \right] \times \sum_{j\in\mathcal{WT}} \xi_{j,t}^P \right\} \tag{23}$$

for facilitating further reformulation. Here, coefficients $\boldsymbol{C}_0$ have elements either 0 or 1. Chance constraints (21b) represent (20e)–(20q), where $\boldsymbol{C}_\ell(\cdot)$ and $\boldsymbol{c}_\ell(\cdot)$ for all $\ell \in \mathbb{Z}_1^{13}$ denote a linear matrix and vector operator of $\boldsymbol{\Lambda}$ and $\boldsymbol{x}$, respectively.

3.1. Wasserstein-metric-based ambiguity set

Definition 1 (*Wasserstein Metric*). For two distributions $\mathbb{P}_a$ and $\mathbb{P}_b$ on $\mathbb{R}^n$, the type-1 Wasserstein distance is defined by

$$\mathbb{W}(\mathbb{P}_a, \mathbb{P}_b) = \min_{\Psi} \left\{ \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\chi_a - \chi_b\|_1 \Psi(d\xi_a, d\xi_b) \right\},$$

where Ψ is a joint distribution on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals $\mathbb{P}_a$ and $\mathbb{P}_b$.

The Wasserstein metric between $\mathbb{P}_a$ and $\mathbb{P}_b$ can be viewed as the cost of an optimal mass transportation plan Π that minimizes the cost of moving $\mathbb{P}_a$ to $\mathbb{P}_b$ with $\|\chi_a - \chi_b\|_1$ the cost of moving a unit mass from χ_a to χ_b.

It is assumed that a finite number of samples $\hat{\Xi} = \{\hat{\xi}_1, \ldots, \hat{\xi}_N\} \subseteq \Omega$ can be drawn independently from the unknown distribution $\mathbb{P}$ of ξ. Accordingly, we can denote the ambiguity set based on Definition 1

$$D := \left\{\mathbb{P} \in \mathcal{P}(\Omega) : \mathbb{W}(\mathbb{P}, \hat{\mathbb{P}}) \le \theta\right\} \tag{24}$$

where $\mathcal{P}(\Omega)$ is the set of all probability distributions on Ω and the radius $\theta > 0$. For the power system operator, the radius θ is a tuning parameter to adjust the risk attitude.

3.2. Reformulation of objective function

Without loss of generality, this paper considers the uncertainty ξ supported by a polytope $\Omega = \{\xi : L\xi \le l\}$. According to [25, Corollary 5.1], the worst-case expectation term $\max_{\mathbb{P}\in D} \mathbb{E}_{\mathbb{P}}\left[c_{0,2}^{\top}\Lambda C_0 \xi\right]$ is equivalent to the conic program

$$\min_{\kappa^o, \beta^o, \gamma^o} \ \kappa^o\theta + \frac{1}{N}\mathbf{1}_N^{\top}\beta^o \tag{25a}$$

$$\text{s.t.} \quad c_{0,2}^{\top}\Lambda C_0\hat{\xi}_i + \tau_i^{o\top}(l - L\hat{\xi}_i) \le \beta_i^o, i \in \mathbb{Z}_1^N \tag{25b}$$

$$\left\|L^{\top}\tau_i^o + C_0^{\top}\Lambda^{\top}c_{0,2}\right\|_{\infty} \le \kappa^o, i \in \mathbb{Z}_1^N \tag{25c}$$

$$\kappa^o \ge 0,\ \gamma_i^o \ge 0,\ i \in \mathbb{Z}_1^N \tag{25d}$$

for a given radius $\theta > 0$. As the Wasserstein metric defined by Definition 1 uses the ℓ_1-norm, Problem (25) is a linear program by rewriting (25c) element-wise [29].

In practice, if there is no prior knowledge about the polytopic support, as discussed in [30], one can simply choose $\Omega = \mathbb{R}^{n_c \cdot T}$ such that (25) is reduced to

$$\min_{\kappa^o, \beta^o, \gamma^o} \kappa^o\theta + \frac{1}{N}\mathbf{1}_N^{\top}\beta^o \ \text{ s.t.} \begin{cases} c_{0,2}^{\top}\Lambda C_0\hat{\xi}_i \le \beta_i^o, i \in \mathbb{Z}_1^N, \\ \left\|C_0^{\top}\Lambda^{\top}c_{0,2}\right\|_{\infty} \le \kappa^o, i \in \mathbb{Z}_1^N, \\ \kappa^o \ge 0,\ \gamma_i^o \ge 0,\ i \in \mathbb{Z}_1^N. \end{cases}$$

3.3. OCA for joint chance constraints

In order to transfer the joint chance constraints approximately into a tractable form, we omit the index in (21b) for notational simplification in the following discussion, i.e.,

$$\Gamma := \left\{(x, \Lambda) : \min_{\mathbb{P}\in D} \mathbb{P}\left[C(\Lambda)\xi \le c(x)\right] \ge 1 - \epsilon\right\} \tag{26a}$$

$$= \left\{(x, \Lambda) : \min_{\mathbb{P}\in D} \mathbb{P}\left[C^k(\Lambda)\xi \le c^k(x),\ k \in \mathbb{Z}_1^K\right] \ge 1 - \epsilon\right\} \tag{26b}$$

with C^k and c^k the kth row of C and c, respectively. The classical method to reformulate (26) based on the ambiguity set (24) first utilizes Bonferroni's inequality to transfer it into individual chance constraints, i.e.,

$$\Gamma_{\mathrm{B}} := \left\{(x, \Lambda) : \min_{\mathbb{P}\in D} \mathbb{P}\left[C^k(\Lambda)\xi \le c^k(x)\right] \ge 1 - \epsilon^k,\ k \in \mathbb{Z}_1^K\right\}$$

and then, conservatively approximate the individual chance constraints above by the worst-case CVaR constraints.

$$\Gamma_{\mathrm{BC}} := \left\{(x, \Lambda) : \max_{\mathbb{P}\in D} \mathbb{P}\text{-CVaR}_{\epsilon^k}\left[C^k(\Lambda)\xi - c^k(x)\right] \le 0,\ k \in \mathbb{Z}_1^K\right\}$$

Here, the definition of $\mathbb{P}$-CVaR$_{\epsilon^k}$ follows (1). One can show the inclusion $\Gamma \subseteq \Gamma_{\mathrm{B}} \subseteq \Gamma_{\mathrm{BC}}$ as discussed in [31]. Moreover, [32, Corollary 2] shows that if we choose $\epsilon^k \le N^{-1}$ for all $k \in \mathbb{Z}_1^K$, we can have $\Gamma_{\mathrm{BC}} = \Gamma_{\mathrm{B}}$. This implies that Γ_{BC} constitutes the best convex inner approximation of Γ_{B}.

However, the choice of ϵ_k greatly affects the performance of BA. The standard way for BA is to set $\epsilon_k = \epsilon/K$, which is overly conservative [33]. An open question related to BA is how to choose ϵ_k. It would be attractive to treat ϵ_k as decision variables (subject to $\epsilon_k > 0$ and $\sum_k \epsilon_k \le \epsilon$). Unfortunately, such an attempt destroys the convexity and thus, makes the approximation intractable. Recent works [21,34] treat ϵ_k as decision variables and develop the optimized BA (OBA) algorithm for moment-based DRJCC problem. However, it is of huge challenge to incorporate OBA to the metric-based DRJCC model [21].

Moreover, the classical Bonferroni's inequality approximation is insufficient when the sets of violations of renewable energy scenarios for different individual chance constraints in Γ_{B} have significant overlap [35]. Therefore, we present the OCA formulation that provides an intuitive dual interpretation and is provably tighter than the BA.

The main idea of OCA is first to equivalently reformulate (26) as a distributionally individual chance constraint

$$\min_{\mathbb{P}\in D}\left[\max_{k\in\mathbb{Z}_1^K}\left\{\delta^k\left[C^k(\Lambda)\xi \le c^k(x)\right]\right\} \le 0\right] \ge 1 - \epsilon \tag{27}$$

for any $\delta = [\delta^1, \ldots, \delta^K]^{\top} \in \Pi^{++} := \left\{\delta \in \mathbb{R}_{++}^K : \mathbf{1}_K^{\top}\delta = 1\right\}$. Here, set Π^{++} denotes the relative interior of the probability simplex. Then, directly applying the CVaR approximation for (27) yields

$$\Gamma_{\mathrm{C}}(\delta) := \left\{(x, \Lambda) : \max_{\mathbb{P}\in D} \mathbb{P}\text{-CVaR}_{\epsilon}\left[\max_{k\in\mathbb{Z}_1^K}\left\{\delta^k\left[C^k(\Lambda)\xi - c^k(x)\right]\right\}\right] \le 0\right\}.$$

One can show the inclusion $\Gamma_{\mathrm{C}}(\delta) \subseteq \Gamma$ for all $\delta \in \Pi^{++}$ as discussed in [33, Section 4.2]. Moreover, the approximation $\Gamma_{\mathrm{C}}(\delta)$ becomes essentially exact when the scaling parameters δ are chosen optimally and the worst-case CVaR constraints can be evaluated efficiently if the scaling parameters are kept constant [36].

Based on the definition of $\mathbb{P}$-CVaR given by (1), [30, Proposition 2] shows that the left-hand side of CVaR approximation $\Gamma_{\mathrm{C}}(\delta)$ for a given $\delta \in \Pi^{++}$ coincides with the optimal value of the following conic program

$$\min_{\kappa^c, \sigma^c, \beta^c, \gamma^c} \kappa^c\theta + \frac{1}{N}\mathbf{1}_N^{\top}\beta^c \tag{28a}$$

$$\text{s.t.} \quad \sigma^c \le \beta_i^c,\ \gamma_{i,k}^c \ge 0,\ i \in \mathbb{Z}_1^N, k \in \mathbb{Z}_1^K, \tag{28b}$$

$$\delta_k\left[C^k(\Lambda)\hat{\xi}_i - c^k(x)\right] + (\epsilon - 1)\sigma^c + \epsilon(l - L\hat{\xi}_i)^{\top}\gamma_{i,k}^c \le \epsilon\beta_i,\ i \in \mathbb{Z}_1^N, k \in \mathbb{Z}_1^K, \tag{28c}$$

$$\left\|\epsilon\gamma_{i,k}^c L - \delta_k C^k(\Lambda)\right\|_{\infty} \le \epsilon\kappa^c,\ i \in \mathbb{Z}_1^N, k \in \mathbb{Z}_1^K. \tag{28d}$$

Similar to the reformulation of the objective function, we can rewrite the infinity norm in (28d) element-wise such that (28) becomes a linear program.

3.4. Sequential convex algorithm for OCA

Let us recap the original problem (21). We denote the number of inequalities in the ℓth chance constraint by K_ℓ and define Π_ℓ^{++} as the associated relative interior of the K_ℓ-dimensional probability simplex. Moreover, for any $\delta_\ell \in \Pi_\ell^{++}$, the discussion in Section 3.3 implies that

$$\Phi(\Delta) := \min_{(x,\Lambda)\in\Theta} c_{0,1}^{\top}x + \max_{\mathbb{P}\in D} \mathbb{E}_{\mathbb{P}}\left[c_{0,2}^{\top}\Lambda C_0\xi\right] \tag{29}$$

$$\text{s.t.} \quad \max_{\mathbb{P}\in D} \mathbb{P}\text{-CVaR}_{\epsilon}\left[\max_{k\in\mathbb{Z}_1^{K_\ell}}\left\{\delta_\ell^k\left[C_\ell^k(\Lambda)\xi - c_\ell^k(x)\right]\right\}\right] \le 0,\ \forall \ell$$

constitutes a tractable problem and provides an upper bound on (21) for a given $\Delta = [\delta_1^{\top}, \ldots, \delta_{13}^{\top}]^{\top}$. Therefore, we can optimize over all $\delta_\ell \in \Pi_\ell^{++}$ to find the best bound $\Phi^\star$, i.e,

$$\Phi^\star = \min_{\Delta} \Phi(\Delta) \quad \text{s.t. } \delta_\ell \in \Pi_\ell^{++},\ \ell \in \mathbb{Z}_1^{13}. \tag{30}$$

Unfortunately, treating δ^j as additional decision variables makes (30) nonconvex. Therefore, the iterative algorithm proposed in [30,37] is tailored to deal with it by sequentially optimizing (x, Λ) and Δ in an alternating direction style. There are two main steps outlined in

Algorithm 1 Sequential convex algorithm for OCA

Input: $\varsigma > 0$, a sufficient large $\phi > 0$, $\boldsymbol{\delta}_\ell = \mathbf{1}_{K_\ell}/K_\ell$, $\ell \in \mathbb{Z}_1^{13}$.
Repeat:

S1 Solve

$$\phi^+ = \min_{(\boldsymbol{x},\boldsymbol{\Lambda})\in\Theta,\boldsymbol{\nu}\geq 0} \boldsymbol{c}_{0,1}^\top \boldsymbol{x} + \omega \mathbf{1}_{13}^\top \boldsymbol{\nu} + \max_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}}\left[\boldsymbol{c}_{0,2}^\top \boldsymbol{\Lambda} \boldsymbol{C}_0 \boldsymbol{\xi}\right] \tag{31}$$

$$\text{s.t. } \max_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\text{-CVaR}_\epsilon\left[\max_{k\in\mathbb{Z}_1^{K_\ell}} \left\{\delta_\ell^k \left[\boldsymbol{C}_\ell^k(\Lambda)\boldsymbol{\xi} - \boldsymbol{c}_\ell^k(\boldsymbol{x})\right]\right\}\right] \leq \nu_\ell, \forall \ell$$

by replacing the objective and the joint chance constraints according to the tractable reformulation (25) and (28). If $|(\phi^+ - \phi)/\phi| < \varsigma$ or the maximum number of iteration has been executed, then stop and return the solution $(\boldsymbol{x}, \boldsymbol{\Lambda}, \boldsymbol{\nu})$ of (31), else go to Step 2.

S2 Update $\boldsymbol{\delta}_\ell^+$ for all $\ell \in \mathbb{Z}_1^{13}$ by solving

$$\min_{\boldsymbol{\Delta}} \sum_{\ell=1}^{13} \max_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\text{-CVaR}_\epsilon\left[\max_{k\in\mathbb{Z}_1^{K_\ell}} \left\{\delta_\ell^k \left[\boldsymbol{C}_\ell^k(\Lambda)\boldsymbol{\xi} - \boldsymbol{c}_\ell^k(\boldsymbol{x})\right]\right\}\right]$$

$$\text{s.t. } \boldsymbol{\delta}_\ell \in \Pi_\ell^{++},\ \ell \in \mathbb{Z}_1^{13} \tag{32}$$

with substituting the objective by the tractable reformulation (28). Then, return to Step 1.

Algorithm 1, Step 1 solving (31) with a fixed $\boldsymbol{\delta}$, which is a relaxed version of (29) while Step 2 updating $\boldsymbol{\delta}$ based on the solution of (31).

In Algorithm 1, ς denotes a prescribed minimum relative improvement during an iteration. Moreover, we introduce the auxiliary slackness via $\boldsymbol{\nu} \geq 0$ in Problem (31) and penalize it in the objective weighted by $\omega > 0$. They ensure feasibility in case of poor initialization of scaling parameters. If ω is chosen large enough, then the algorithm is guaranteed to terminate at $\boldsymbol{\nu} = 0$, so the output is feasible w.r.t (29).

Remark 1. The main idea is optimizing $\boldsymbol{\Delta}$ at each iteration to achieve a desired improvement. This follows the fact that $\boldsymbol{\Delta}$ constitutes a vector of scaling parameters, which can be tuned to optimize the quality of the CVaR approximation. The sequence of objective values ϕ generated by the algorithm is non-increasing and thus guaranteed to converge.

4. Numerical results

In this paper, an IEEE 33-bus distribution network is modified as a MEMG, shown in Fig. 3. The network topology and power line parameters were acquired from `Matpower`. The risk parameters ϵ of all joint chance constraints is set to be 5%. For the sequential convex optimization Algorithm 1, the maximum number of iterations is set to 10, the minimum relative improvement $\varsigma = 0.1$. The proposed method is tested in `MATLAB R2021a` on an Intel Core i7-8565U, 1.8 GHz, 8 GB RAM PC using `Gurobi 9.0`.

The relative wind output data provided by the Australian Energy Market Operator [38] is utilized to establish the data-driven metric-based ambiguity set. Noted that the hourly wind power samples of each WT are converted into a wind power curve over 24 h manually. The energy management is conducted for a 24 h horizon with the time resolution as 1 h. The WT capacity, daily electricity load profiles, multi-energy load profiles of each thermal group, and other parameters for DERs are all made available online [39]. We select $N \in \{25, 50, 100, 200\}$ samples and $S = 1000$ test samples $\hat{\boldsymbol{\xi}}_{N+i}$, $i \leq S$ to illustrate the out-of-sample performance. Here, the out-of-sample cost $\widehat{\mathcal{O}}^C$ is estimated under the assumption that only the nominal state decisions $\hat{\boldsymbol{x}}$ is executed and that the output adjustments of DERs are determined by solving a deterministic operation model without considering uncertainties. Moreover, drastic measures such as load shedding and renewable energy curtailment must be included in the deterministic operating model to ensure the feasibility of the model. Here we take into account a penalty cost of 0.5 \$/kWh for load shedding and 0 \$/kWh for renewable energy curtailment. The out-of-sample constraint violation probability $\widehat{\mathcal{O}}^V$ denotes the average violations of 13 chance constraints (20e)–(20q), assessed by

$$\widehat{\mathcal{O}}^V = \sum_{\ell=1}^{13} \left\{ \frac{1}{S} \sum_{i=1}^{S} \mathbb{1}_{\Gamma_\ell^c(\hat{\boldsymbol{\xi}}_{N+i})}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\Lambda}}) \right\}$$

with $\Gamma_\ell^c(\boldsymbol{\xi}) := \left\{(\boldsymbol{x}, \boldsymbol{\Lambda}) \,|\boldsymbol{C}_\ell(\boldsymbol{\Lambda})\boldsymbol{\xi} > \boldsymbol{c}_\ell(\boldsymbol{x})\right\}$.

Here, $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\Lambda}})$ is the output of Algorithm 1.

4.1. Out-of-sample performance of BA and OCA for DRJCC model

4.1.1. Impact of wasserstein radius

Since the renewable energy output data is not necessarily being drawn from any underlying distribution in an independent manner, the guarantees to establish rigorous bounds [14,25] on the Wasserstein distance are not necessarily applicable for the multi-period operation problem of MEMG. As a result, we empirically evaluate the out-of-sample performance of both BA and OCA for varying radii, averaged over 20 independent simulation runs to increase statistical robustness. The non-i.i.d. samples and test samples are selected intentionally from historical data of different time windows. Taking the sample set $N = 100$ as an example, an empirical study on $\widehat{\mathcal{O}}^C$ and $\widehat{\mathcal{O}}^V$ as a function of radii θ is carried out, shown in Fig. 4. With the increasing of radius, $\widehat{\mathcal{O}}^V$ tends to decrease while $\widehat{\mathcal{O}}^C$ tends to increase. This is because a larger Wasserstein ball results in higher robustness to sampling errors.

4.1.2. Impact of sample size

The out-of-sample cost $\widehat{\mathcal{O}}^C$ under different sample sizes for both BA and OCA is shown in Fig. 5. It is shown that $\widehat{\mathcal{O}}^C$ is negatively correlated with the sample size N, that is, the larger the available samples, the smaller the out-of-sample cost. Therefore, the conservatism of the proposed DRJCC energy management model for MEMG is mitigated with the increase of sample size. The reason is that when more historical data is available, the more probabilistic information of the true probability distribution about the uncertainty will be revealed.

From Figs. 4 and 5, we can see that the out-of-sample cost $\widehat{\mathcal{O}}^C$ of the BA is larger than the OCA for all radii, and there is a sharp increase in the BA when the radius is close to 10^{-3}. It can be concluded that the solution obtained from the BA tends to be more conservative (with a larger $\widehat{\mathcal{O}}^C$ and a lower $\widehat{\mathcal{O}}^V$) than the OCA when the radius θ is relatively large. This is because that the choice of $\epsilon^k = \epsilon/K$ for BA is overly conservative. Moreover, a closer inspection reveals that when radius is smaller than 10^{-2} (this critical radius is different for different sample sets N), there is no significant change in the out-of-sample performance for OCA. It means that there is a critical radius corresponding to a better out-of-sample performance (with smaller $\widehat{\mathcal{O}}^C$ and lower $\widehat{\mathcal{O}}^V$) for OCA. We can conclude that the out-of-sample performance of the proposed DRJCC model can be improved by acknowledging the ambiguity for sophisticated system operators. However, the optimal ex-ante selection of this critical radius in the OCA formulation is not straightforward and may require learning from experience, i.e., the past outcome of the DRJCC model. Accordingly, we suggest that the decision-maker may select a radius ranging from this critical radius that provides the best out-of-sample performance to infinity, leading to similar decisions of robust optimization.

In short, radius is a principled way to evaluate the trade-off between the solution robustness and performance. In a robust formulation, it can be cumbersome to directly adjust the multiple different bounds of the uncertain parameters, while the radius in the metric-based DRJCC model is a scalar parameter that can control the degree of distributional robustness. For a larger radius, the joint chance constraints are required

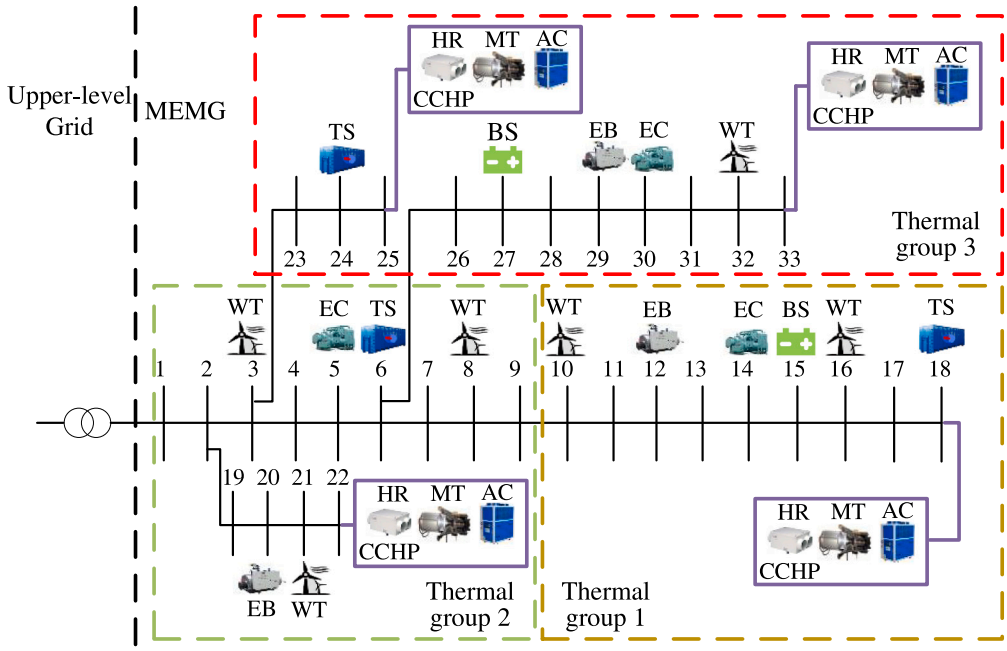


Fig. 3. IEEE 33-bus radial distribution network based MEMG.

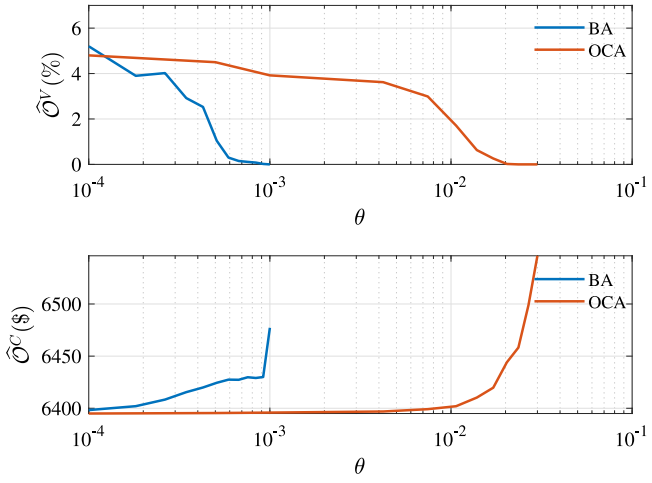


Fig. 4. Out-of-sample performance under different radius (N = 100).

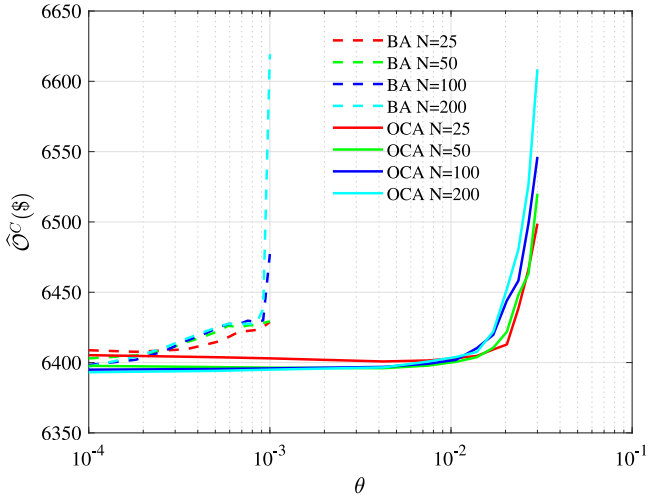


Fig. 5. Out-of-sample cost under different sample size.

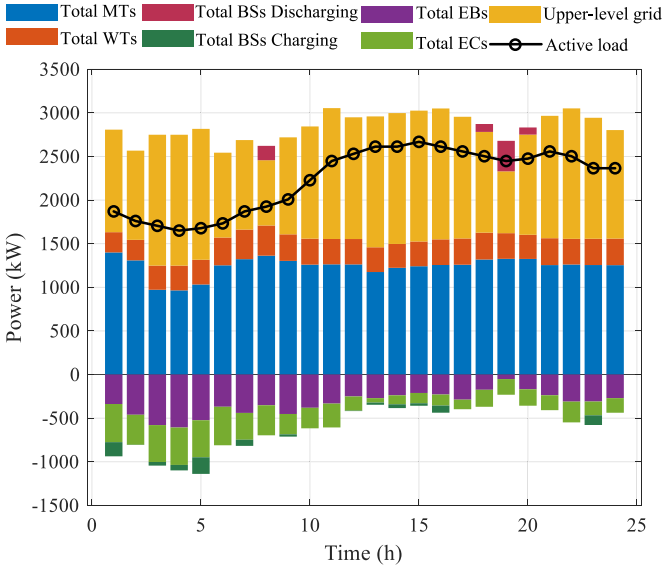


Fig. 6. Electricity dispatch.

to be holded for a larger set of distributions. This is useful when the number of samples available is small. The decision-maker can choose a larger radius to improve the robustness of the solution to uncertain parameters that have not yet been realized. When large amounts of historical data are available, the empirical distribution tends to represent uncertain parameters well, and the decision-maker can choose a smaller radius to reduce operation cost.

4.2. Energy scheduling results

The energy scheduling results obtained from the metric-based DR-JCC model using OCA with the radius $\theta = 10^{-2}$ and the sample size $N = 100$ are illustrated in Figs. 6 and 7, as OCA outperforms BA. It can be seen that due to the low electricity prices at 1–6 h and 13–16 h, MEMG purchases more power from the upper-level grid to supply local

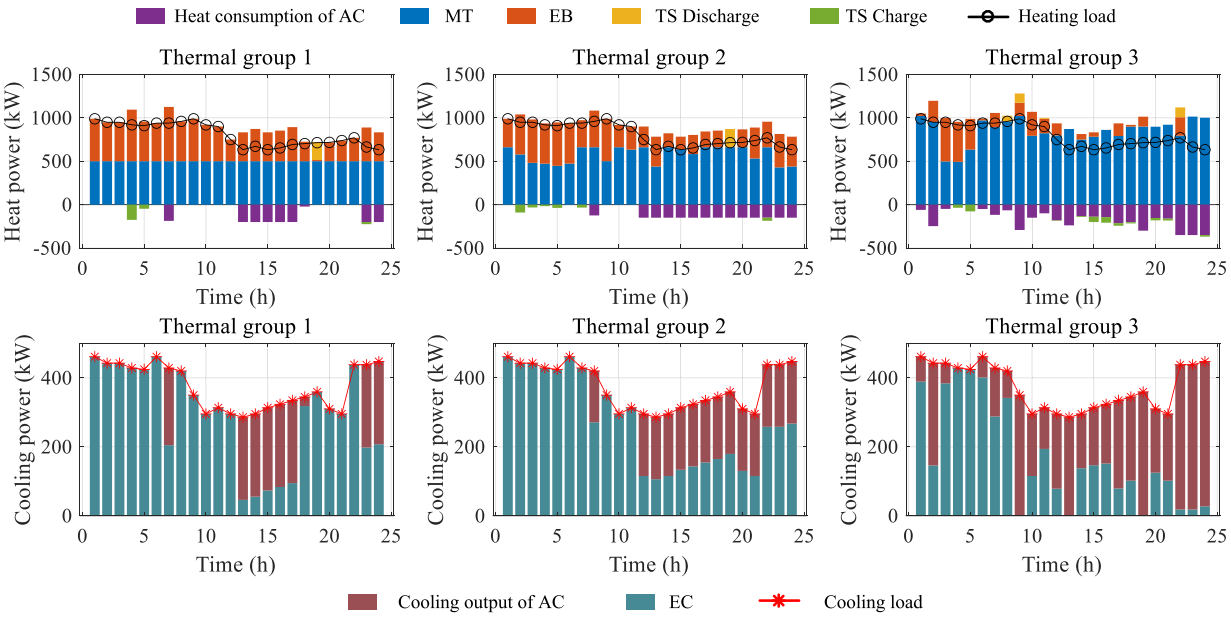


Fig. 7. Heat and cooling energy dispatch.

Table 1
Adjustment factors of electricity facilities.

No.		Time periods (h)																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
α^{MT}	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	0.26	0.50	0.51	0	0.47	0	0	0.41	0	0.07	0	0.60	0	0.06	0.21	0	0	0	0	0.33	0	0.60	0.57
	3	0	0	0.14	0.09	0.20	0.14	0	0	0	0.09	0	0.08	0	0.63	0	0	0.09	0.34	0	0	0.28	0.08	0	0
	4	0	0.21	0	0	0.50	0	0.30	0.14	0	0.52	0.53	0.53	0	0	0.54	0.39	0.51	0	0.32	0.34	0	0.52	0	0.05
α^{EB}	1	0	0	0.26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	0.38	0	0.34	0	0.31	0.06	0	0.27	0	0.05	0	0.40	0	0.04	0.14	0	0	0	0	0.22	0	0.40	0.38
	3	0	0.14	0	0.06	0.30	0	0.20	0	0.18	0.39	0.17	0.26	0	0.12	0.10	0	0.27	0.04	0.21	0	0	0.40	0	0
α^{EC}	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	3	0	0	0.08	0	0	0.08	0	0.10	0	0	0.18	0.13	0	0.25	0.26	0.26	0.13	0.17	0	0.23	0.17	0	0	0
α^{BSC}	1	0.37	0.01	0	0	0	0	0.44	0	0.13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	0.63	0.01	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
α^{BSD}	1	0	0	0	0	0	0	0	0.52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	0.24	0	0	0	0	0	0	0	0	0	0.45	0.47	0.43	0	0	0	0

power demand and generate heat and cooling energy via EBs and ECs. Meanwhile, the extra electricity is used to charge the BSs, and the extra heat is used to charge the TSs. When the load or electricity price is high, BSs and TSs will discharge. Moreover, cooling energy is mainly supplied by the ECs using electricity and supplemented by the ACs of CCHP system. Clearly, the thermal energy is stored and released for several times during the whole day to decouple the electric, heat, and cooling outputs of DERs. We can conclude that through CCHP plants, EBs, BSs, ECs, and TSs, multiple energies are coordinated and CCHP operation is more flexible: when the electricity prices are low, EBs and ECs outputs more to transfer power to heat or cooling energy; when the electricity prices are high, thermal energy is mainly served by the CCHP plants and TSs. In fact, the operation of MEMG is tightly coupled and the proposed model achieves flexible coordination of different energies. For simplicity, only the adjustment factors of electricity facilities MT, EB, EC, and BS are demonstrated in Table 1. The sum of these adjustment factors are all equal to 1 in all time periods, guaranteeing that the renewable energy forecasting errors are fully mitigated.

4.3. Impact of risk parameters

Table 2 shows the total operation cost (in-sample cost) obtained from the proposed metric-based DRJCC model using OCA with the same radius $\theta = 10^{-2}$ and the same sample size $N = 100$ under different risk parameters ϵ. Since a larger ϵ will allow to tolerate some constraint violation probability, the total operation cost will decrease. That is, the choice of ϵ also has an impact on the solution conservatism. This confirms that DRJCC model allows to reduce the total operation cost by allowing constraint violation to some extent. This also implies that the DRJCC model can find an optimal trade-off between the total operation cost and operational risk. In fact, the system operators can reduce the total operation cost by increasing ϵ as long as they can tolerate a relatively high operational risk.

Table 2
Comparison of different risk parameters for DRJCC using OCA.

ϵ	3%	5%	10%	15%	20%
Total operation cost ($)	6696	6482	6343	6247	6201

4.4. Comparison with SO, RO, & moment-based DRJCC model

To comprehensively assess the performance of the metric-based DRJCC model using OCA formulation, four different models are compared. The first is the proposed model with radius $\theta = 10^{-2}$ and the sample size $N = 100$. The second is an ULDR-based adjustable RO model,

Table 3
Comparisons of different optimization models.

Model	Total operation cost ($)	Out-of-sample performance	
		$\widehat{\mathcal{O}}^C$ ($)	$\widehat{\mathcal{O}}^V$
Metric-based DRJCC	6482	6396	1.9%
Moment-based DRJCC	6772	6765	1.1%
ULDR-based adjustable RO	6961	6902	0
Scenario-based SO	5849	6028	7.4%

Table 4
Comparisons for different scheduling time periods of OCA-based DRJCC Model.

Time periods	Out-of-sample performance		Iterations	Solution time (s)
	$\widehat{\mathcal{O}}^V$	$\widehat{\mathcal{O}}^C$ ($)		
24 h	1.9%	6396	3	1416
18 h	1.9%	4702	3	657
12 h	1.9%	3039	3	61
6 h	1.7%	1408	3	21

which requires the security constraints satisfied in a box uncertainty set obtained by the largest samples $N = 200$. The third is the scenario-based SO model, which minimizes the expectation operation cost under the largest samples $N = 200$. The fourth is the moment-based DRJCC model formulating as a second-order cone program **P1** in [21], whose ambiguity set is constructed with the sample mean and covariance obtained by the largest samples $N = 200$.

The total operation cost (in-sample cost) and the out-of-sample performance is compared in Table 3. It can be seen that the $\widehat{\mathcal{O}}^V$ in RO, moment DRJCC, and metric DRJCC models are all lower then the required constraint violation probability 5% since these three models are quite robust. However, the $\widehat{\mathcal{O}}^V$ in SO model is 7.4%, which cannot guarantee the required constraint violation probability. This is because the SO model is solved under a series of scenarios, and thus, some extreme scenarios could inevitably be neglected. The total operation cost of moment and metric DRJCC models are smaller than RO model while larger than SO model. In contrast to RO, metric DRJCC offers a risk-aware framework that provides performance guarantees when the distribution of uncertainty is not perfectly known. In metric DRJCC, system operator may view the radius θ as a tuning parameter to impose his risk attitude by varying the size of ambiguity set. It can also be seen as a proxy representing the confidence level of the operator to his knowledge about the underlying uncertainty. Under two extreme cases, i.e., the smallest ambiguity set containing a unique distribution and the largest one containing all potential distributions, the outcomes of metric DRJCC will be similar to those in a SO model and in a RO model, respectively. The choice for the size of ambiguity set between those two extreme cases enables the system operator to take a risk attitude in between. Metric DRJCC generally outperforms SO and RO due to their inherent shortcomings. On the one hand, scenario-based SO provides poor out-of-sample performance unless the number of scenarios is very high, which in turn, increases the computational burden. On the other hand, RO provides a conservative solution for a given uncertainty set. It is also challenging to describe all potential distributions using a single uncertainty set.

It can also be seen that the solution obtained by the moment DRJCC model is more conservative than the metric DRJCC model. This is because the first two moments are adopted to establish the ambiguity set in the moment DRJCC model. Once the moment information is determined, the ambiguity set in moment-based model is fixed and the solution conservativeness is determined. However, the proposed metric-based DRJCC energy management model for MEMG can flexibly balance the system economic and constraint reliability. For metric-based DRJCC, the more statistical data are available, the more trustworthy and smaller the ambiguity set is, which generates less conservative solution.

4.5. Scalability of Algorithm 1 with Respect to the Number of Constraints

In order to assess the impact of the number of constraints in each joint chance constraint, we test different scheduling time periods with 24 h, 18 h, 12 h, and 6 h under fixing the radius $\theta = 10^{-2}$ and sample size $N = 100$. For simplicity, the adopted 18 h, 12 h, and 6 h data is part of the 24 h data. As the results shown in Table 4, the tailored sequential convex algorithm 1 for OCA formulation usually terminates after 3 iterations, which verifies its fast convergence in practice. And the longest solution time of 24 h scheduling time periods is less than 0.5 h, which is acceptable for a day-ahead energy management problem. Therefore, the proposed OCA formulation and its associated solution algorithm 1 is scalable and the effectiveness is not affected by the number of constraints in each joint chance constraint.

5. Conclusions

This paper proposes a data-driven Wasserstein metric-based DRJCC energy management model for MEMG to account for complicated spatial–temporal correlations among uncertain renewable energy. Various flexible resources in different energy sectors is utilized for uncertainty mitigation. Then, a ULDR model is adopted to facilitate further reformulation of joint chance constraints. Based on a tailored sequential convex algorithm, an OCA formulation for joint chance constraints is proposed to make the data-driven DRJCC model tractable. Numerical results illustrate that the OCA formulation for joint chance constraints is less conservative than the standard BA algorithm. The tailored sequential convex algorithm for solving OCA can achieve high computational efficiency. Moreover, a critical Wasserstein radius exists corresponding to a better out-of-sample performance (with smaller out-of-sample cost and lower constraint violation probability) for the DRJCC energy management model. This illustrates that the out-of-sample performance can be improved by acknowledging the ambiguity for sophisticated system operators.

Future works can be investigated in the following directions. First, extending the proposed model to the nonlinear AC-OPF problem with joint chance constraints is an interesting topic. Second, exploring the OBA formulation for metric-based DRJCC problem is a valuable direction. Then, OBA and OCA can be compared in terms of out-of-sample performance to verify their superiority in joint chance-constrained modeling.

CRediT authorship contribution statement

Junyi Zhai: Conceptualization, Software, Investigation, Data curation, Funding acquisition, Writing – original draft. **Sheng Wang:** Data curation, Writing – original draft. **Lei Guo:** Formal analysis, Writing – review & editing. **Yuning Jiang:** Methodology, Validation, Writing – review & editing, Writing – original draft. **Zhongjian Kang:** Validation, Supervision. **Colin N. Jones:** Supervision, Funding acquisition.

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.

References

[1] Lu S, Gu W, Meng K, Dong Z. Economic dispatch of integrated energy systems with robust thermal comfort management. IEEE Trans Sustain Energy 2021;12(1):222–33.

[2] Li Z, Yan X. Optimal coordinated energy dispatch of a multi-energy microgrid in grid-connected and islanded modes - sciencedirect. Appl Energy 2018;210:974–86.

[3] Li Z, Xu Y. Temporally-coordinated optimal operation of a multi-energy microgrid under diverse uncertainties. Appl Energy 2019;240.

[4] Li Z, Xu Y, Feng X, Wu Q. Optimal stochastic deployment of heterogeneous energy storage in a residential multienergy microgrid with demand-side management. IEEE Trans Ind Inf 2021;17(2):991–1004.

[5] Zidan A, Gabbar H, Eldessouky A. Optimal planning of combined heat and power systems within microgrids. Energy 2015;93:235–44.

[6] Yang D, Jiang C, Cai G, Yang D, Liu X. Interval method based optimal planning of multi-energy microgrid with uncertain renewable generation and demand. Appl Energy 2020;277:115491.

[7] Zhang C, Xu Y, Li Z, Dong ZY. Robustly coordinated operation of a multi-energy microgrid with flexible electric and thermal loads. IEEE Trans Smart Grid 2019;10(3):2765–75.

[8] Zhang C, Xu Y, Dong ZY. Robustly coordinated operation of a multi-energy micro-grid in grid-connected and islanded modes under uncertainties. IEEE Trans Sustain Energy 2020;11(2):640–51.

[9] Zhou Y, Shahidehpour M, Wei Z, Li Z, Sun G, Chen S. Distributionally robust unit commitment in coordinated electricity and district heating networks. IEEE Trans Power Syst 2020;35(3):2155–66.

[10] Zhai J, Zhou M, Li J, Zhang X, Li G, Ni C, Zhang W. Hierarchical and robust scheduling approach for vsc-mtdc meshed ac/dc grid with high share of wind power. IEEE Trans Power Syst 2021;36(1):793–805.

[11] Zhai J, Jiang Y, Li J, Jones C, Zhang XP. Distributed adjustable robust optimal power-gas flow considering wind power uncertainty. Int J Electr Power Energy Syst 2022;139.

[12] Yz A, Wei LA, Zh A, Feng ZA, Jian LB, Shu ZB. Distributionally robust coordination optimization scheduling for electricity-gas-transportation coupled system considering multiple uncertainties. Renew Energy 2021;163:2037–52.

[13] Ji R, Lejeune MA. Data-driven distributionally robust chance-constrained optimization with wasserstein metric. J Global Optim 2021;79(4):779–811.

[14] Yao L, Wang X, Li Y, Duan C, Wu X. Distributionally robust chance-constrained ac-opf for integrating wind energy through multi-terminal vsc-hvdc. IEEE Trans Sustain Energy 2020;11(3):1414–26.

[15] Zhu R, Wei H, Bai X. Wasserstein metric based distributionally robust approximate framework for unit commitment. IEEE Trans Power Syst 2019;34(4):2991–3001.

[16] Saberi H, Zhang C, Dong ZY. Data-driven distributionally robust hierarchical coordination for home energy management. IEEE Trans Smart Grid 2021;12(5):4090–101.

[17] Zhou Y, Shahidehpour M, Wei Z, Li Z, Sun G, Chen S. Distributionally robust co-optimization of energy and reserve for combined distribution networks of power and district heating. IEEE Trans Power Syst 2020;35(3):2388–98.

[18] Zhai J, Jiang Y, Shi Y, Jones CN, Zhang XP. Distributionally robust joint chance-constrained dispatch for integrated transmission-distribution systems via distributed optimization. IEEE Trans Smart Grid 2022;13(3):2132–47.

[19] Ordoudis C, Nguyen VA, Kuhn D, Pinson P. Energy and reserve dispatch with distributionally robust joint chance constraints. Oper Res Lett 2021.

[20] Peña-Ordieres A, Molzahn DK, Roald LA, Wächter A. Dc optimal power flow with joint chance constraints. IEEE Trans Power Syst 2021;36(1):147–58.

[21] Yang L, Xu Y, Sun H, Wu W. Tractable convex approximations for distributionally robust joint chance constrained optimal power flow under uncertainties. IEEE Trans Power Syst 2021;1.

[22] Chen W, Sim M, Sun J, Teo CP. From cvar to uncertainty set: Implications in joint chance-constrained optimization. Oper Res 2010;58(2):470–85.

[23] Xie W, Ahmed S, Jiang R. Optimized bonferroni approximations of distributionally robust joint chance constraints. 2017.

[24] Poolla BK, Hota AR, Bolognani S, Callaway DS, Cherukuri A. Wasserstein distributionally robust look-ahead economic dispatch. IEEE Trans Power Syst 2021;36(3):2010–22.

[25] Mohajerin Esfahani P, Kuhn D. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. Math Program 2018;171(1):115–66.

[26] Kuhn D, Wiesemann W, Georghiou A. Primal and dual linear decision rules in stochastic and robust optimization. Math Program 2011;130(1):177–209.

[27] Arrigo A, Ordoudis C, Kazempour J, De Grève Z, Toubeau JF, Vallée F. Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation. European J Oper Res 2022;296(1):304–22.

[28] Li Z, Xu Y, Fang S, Zheng X, Feng X. Robust coordination of a hybrid ac/dc multi-energy ship microgrid with flexible voyage and thermal loads. IEEE Trans Smart Grid 2020;11(4):2782–93.

[29] Boyd S, Vandenberghe L. Convex optimization. Cambridge University Press; 2004.

[30] Ordoudis C, Pinson P, Morales JM. An integrated market for electricity and natural gas systems with stochastic power producers. Eur J Oper Res 2019;272(2):642–54.

[31] Shapiro A, Dentcheva D, Ruszczyński A. Lectures on stochastic programming. Society for Industrial and Applied Mathematics; 2009.

[32] Zhi G, Kuhn D, Wiesemann W. Data-driven chance constrained optimization under wasserstein ambiguity sets. 2018, arXiv preprint arXiv:1809.00210v1.

[33] Nemirovski A, Shapiro A. Convex approximations of chance constrained programs. SIAM J Optim 2007;17(4):969–96.

[34] Xie W, Ahmed S, Jiang R. Optimized bonferroni approximations of distributionally robust joint chance constraints. Math Program 2019;1–34.

[35] Chen W, Sim M, Sun J, Teo CP. From CVaR to uncertainty set: Implications in joint chance-constrained optimization. Oper Res 2010;58(2):470–85.

[36] Chen Z, Kuhn D, Wiesemann W. Data-driven chance constrained programs over wasserstein balls, optimization online. 2018.

[37] Zymler S, Kuhn D, Rustem B. Distributionally robust joint chance constraints with second-order moment information. Math Program 2013;137(1–2):167–98.

[38] Dowell J, Pinson P. Very-short-term probabilistic wind power forecasts by sparse vector autoregression. IEEE Trans Smart Grid 2016;7(2):763–70.

[39] https://github.com/JunyiZhai1990/MEMG.