Energy 198 (2020) 117347

Contents lists available at ScienceDirect

Energy

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

Risk-averse real-time dispatch of integrated electricity and heat system using a modified approximate dynamic programming approach

Zhenning Pan [a], Tao Yu [a, *], Jie Li [a], Kaiping Qu [a], Bo Yang [b]

[a] College of Electric Power, South China University of Technology, Guangzhou, 510640, China
[b] Faculty of Electric Power Engineering, Kunming University of Science and Technology, Kunming, 650500, China

ARTICLE INFO

Article history:
Received 25 December 2019
Received in revised form
2 March 2020
Accepted 7 March 2020
Available online 11 March 2020

Keywords:
Integrated electricity and heat system
Real-time optimization
Risk-averse optimization
Stochastic optimization
Approximate dynamic programming

ABSTRACT

Coordinated operation of integrated electricity and heat system can improve operation flexibility and reduce cost. However, multiple uncertainties challenge its optimal operation. This paper aims at developing a risk-averse and computationally efficient policy for real-time stochastic dispatch of integrated electricity and heat system, which improves the economy as well as avoiding the risk of high costs in critical scenarios. First, real-time dispatch of integrated electricity and heat system is formulated as a multistage risk-averse stochastic sequential optimization problem with dynamic risk measure, where combined heat and power unit, energy storage, flexible electricity and heat load are jointly utilized to minimize the risk-adjusted total costs. Next, a risk-averse dynamic programming formulation of the original problem is presented, upon which a data-driven risk-averse approximate dynamic programming is employed to address computational challenge, and develop almost optimal and computationally efficient policy. By exploiting information from training samples in off-line learning, the proposed algorithm can efficiently responses to the stochastic exogenous information. Comparative simulations with different risk-aversion preferences and different methods verify the effectiveness of the proposed algorithm.

1. Introduction

With the advantages of high efficiency and operation flexibility [1], combined heat and power (CHP) units have been gradually deployed [2]. As a result, power system and heat system are strongly coupled into an integrated electricity and heat system (IEHS) [3]. To exploit synergism, coordinated operation of IEHS has attracted enormous attentions since independent operation of subsystem leads to unnecessary costs and renewable energy source (RES) curtailment [4].

Extensive researches have been conducted on the optimal operation of IEHS. Modeling of IEHS was studied in [5,6]. Optimal energy flow of IEHS was analyzed in [7]. Optimal operation of IEHS with independent heat storage was studied in [8]. Market equilibrium in IEHS was investigated in [9]. The electricity-heat coordinated retail energy market framework considering demand response was proposed in [10]. Ref. [11] utilized the thermal inertia of buildings to enhance the collaborative scheduling of IEHS. The heterogeneous decomposition [12], optimality condition decomposition [13], and feasible region method [14] were adopted to achieve the decentralized coordinated operation. The above mentioned researches have verified the benefits of coordinated operation of IEHS in reducing operation cost and accommodating RES. However, their major drawback is that the investigations were conducted under the assumption of the perfect information of IEHS and uncertainties are not considered.

Since multivariate uncertainties, e.g., stochastic loads and intermittent RES [15], have non-negligible influences on the optimal dispatch of IEHS, the stochastic dispatch of IEHS has attracted many studies. A scenario-based stochastic optimization was adopted in [16], but the computational burden is heavy as the number of scenario increases. A two stage robust optimization for optimal IEHS scheduling was proposed in [17], where the day-ahead stage was to optimize price signals and thermal storage operation considering the uncertainties of loads and RES outputs.

* Corresponding author. College of Electric Power, South China University of Technology, Guangzhou, 510640, China.
E-mail address: taoyu1@scut.edu.cn (T. Yu).

https://doi.org/10.1016/j.energy.2020.117347
0360-5442/

Nomenclature

Acronyms

ADP	Approximate dynamic programming
CHP	Combined heat and power
CVaR	Conditional value at risk
IEHS	Integrated electricity and heat system
EDN/HDN	Electric/heat distribution networks
ESS	Energy storage system
RES	Renewable energy source
FHL	Flexible heat load
DRA	Demand side resource aggregator
D/RMPC	Deterministic/risk-averse model predictive control

Sets and Parameters

T	Set of dispatch stages. $t = 1,2, \ldots, \lvert T \rvert$
G	Set of segments of resource state. $g = 1,2, \ldots, \lvert G \rvert$
$N^{\text{pipe}}/N^{\text{node}}$	Set of pipelines/nodes in HDN
$N^{\text{node}}_{\text{in},k}/N^{\text{node}}_{\text{out},k}$	Set of nodes flowing to/from node k in supply pipelines
$N^{\text{bus}}/N^{\text{line}}$	Set of buses/lines in EDN
$N^{\text{CHP}}/N^{\text{HL}}$	Set of CHP units/heat loads
$N^{\text{CHP}}_{\text{P},n}$	Set of extreme points of operation region of CHP n
$N^{\text{ESS}}/N^{\text{RES}}/N^{\text{DRA}}/N^{\text{SC}}$	Set of ESSs/RESs/DRAs/STATCOM
$\varepsilon(t)$	Real-time price
Δt	Time interval
$a_{n,i}$	Cost coefficient of CHP n, $i = 1,2,3,4,5$
$C_{\text{wt}}/C_{\text{air}}$	Specific heat capacity of water/air
$H^{\text{IHL}}_n(t)$	Inflexible heat load of node n
$m^{\text{CHP}}_n/m^{\text{HL}}_n$	Mass flow rate of CHP/heat load
$m^{\text{pipe,S}}_{kl}/m^{\text{pipe,R}}_{lk}$	Mass flow rate of supply pipeline kl/return pipeline lk
$\tau^{\text{AM}}(t)$	Ambient temperature at time-slot t
$\overline{\tau}^{\text{CHP,S}}_n/\ \underline{\tau}^{\text{CHP,S}}_n/\overline{\tau}^{\text{CHP,R}}_n/\ \underline{\tau}^{\text{CHP,R}}_n$	Upper/lower limit of temperature of CHP n at supply/return pipeline
$\overline{\tau}^{\text{HL,S}}_n/\ \underline{\tau}^{\text{HL,S}}_n/\overline{\tau}^{\text{HL,R}}_n/\ \underline{\tau}^{\text{HL,R}}_n$	Upper/lower limit of temperature of heat load n at supply/return pipeline
d^{pipe}_{kl}	Temperature drop coefficient of pipeline kl
$\overline{\tau}^{\text{FHL}}_n/\ \underline{\tau}^{\text{FHL}}_n$	Upper/lower limit of indoor temperature of FHL
r^{FHL}_n	Thermal resistance of FHL n
$P^{\text{EL}}_i(t)/\ Q^{\text{EL}}_i(t)$	Active/reactive power of inflexible electrical load of bus i
$\overline{P}^{\text{ESS}}_{\text{c},n}/\overline{P}^{\text{ESS}}_{\text{d},n}$	Maximum charging/discharging power of ESS n
$\overline{E}^{\text{ESS}}_n/\underline{E}^{\text{ESS}}_n$	Upper/lower limit of energy stored in ESS n
$\overline{P}^{\text{RES}}_n(t)/\ \overline{Q}^{\text{RES}}_n(t)$	Maximum active/reactive power output of RES n
$\overline{E}^{\text{DRA}}_n(t)/\underline{E}^{\text{DRA}}_n(t)$	Upper/lower cumulative energy limit of DRA n
$\overline{P}^{\text{DRA}}_n(t)/\ \underline{P}^{\text{DRA}}_n(t)$	Maximum/minimum power consumption of DRA n
$\overline{Q}^{\text{SC}}_n/\ \underline{Q}^{\text{SC}}_n$	Upper/lower limit of reactive power output of STATCOM n
$\overline{Q}^{\text{CHP}}_n/\ \underline{Q}^{\text{CHP}}_n$	Upper/lower limit of reactive power output of CHP n
$\eta_{\text{ESS,c}}/\eta_{\text{ESS,d}}$	Charging/discharging efficiency of ESS
β_{ESS}	Operation cost coefficient of ESS
$\beta_{\text{RES}}/\beta_{\text{DRA}}$	Penalty coefficient for RES curtailment/demand shedding
$v^{\text{ESS}}_{n,t}(g)/\ v^{\text{DRA}}_{n,t}(g)/\ v^{\text{FHL}}_{n,t}(g)$	Slope of segment g of ESS/DRA/FHL n

Variables

$H^{\text{FHL}}_n(t)$	Heat power consumption of FHL n
$H^{\text{CHP}}_n(t)$	Heat power output of CHP n
$\tau^{\text{FHL}}_n(t)$	Indoor temperature of FHL n
$\tau^{\text{node,S}}_k(t)/\ \tau^{\text{node,R}}_k(t)$	Flow temperature of node k at supply/return terminal
$\tau^{\text{CHP,S}}_n(t)/\ \tau^{\text{CHP,R}}_n(t)/\ \tau^{\text{HL,S}}_n(t)/\ \tau^{\text{HL,R}}_n(t)$	Flow temperature of CHP/HL at supply/return pipeline
$\tau^{\text{pipe,S,in}}_{kl}(t)/\tau^{\text{pipe,S,out}}_{kl}(t)/\ \tau^{\text{pipe,R,in}}_{kl}(t)/\tau^{\text{pipe,R,out}}_{kl}(t)$	Inlet/outlet flow temperature of supply pipeline kl/return pipeline kl
$\delta^{\text{CHP}}_{n,c}(t)$	Auxiliary variable of the c-th extreme point of CHP n
$P^{\text{ESS}}_{\text{c},n}(t)/\ P^{\text{ESS}}_{\text{d},n}(t)/\ E^{\text{ESS}}_n(t)$	Charging/discharging power/energy stored of ESS n
$P^{\text{RES}}_n(t)/\ Q^{\text{RES}}_n(t)/\ P^{\text{CHP}}_n(t)/Q^{\text{CHP}}_n(t)$	Active/reactive power output of RES/CHP
$P^{\text{DRA}}_n(t)/\ Q^{\text{DRA}}_n(t)$	Active/reactive power consumption of DRA
$E^{\text{DRA}}_n(t)/\ E^{\text{DRA}}_{\text{shed},n}(t)$	Cumulative energy/energy shedding of DRA n
$P^{\text{DRA}}_{\text{shed},n}(t)$	Power shedding of DRA n
$Q^{\text{SC}}_n(t)$	Reactive power output of STATCOM n
$R^{\text{ESS}}_n(t)/R^{\text{DRA}}_n(t)\ /R^{\text{FHL}}_n(t)$	Resource state of ESS/DRA/FHL n
$y^{\text{ESS}}_{n,t}(g)/y^{\text{DRA}}_{n,t}(g)/\ y^{\text{FHL}}_{n,t}(g)$	Auxiliary variable of ESS/DRA/FHL n

Ref. [18] adopted adaptive two-stage robust optimization to co-optimize both energy and reserve. These papers focused on coping with uncertainties in day-ahead stage. However, a scenario-based stochastic optimization solution or a robust optimization solution cannot be adapted to uncertainties [19]. Meanwhile, in real-time stage, the optimal dispatch of IEHS should be made at each specific time-slot after observing any random realizations of uncertainties and the decisions at current stage will affect future dispatch. As a result, the above mentioned methods cannot be directly applied to the real-time dispatch of IEHS.

Note that when exogenous information is subject to uncertainty, real-time optimal dispatch of IEHS leads to a multistage stochastic sequential optimization problem. The optimal solution to such problem can be solved by the Bellman optimality equation [20], but it is usually intractable due to the high dimension state space. For the sake of compromise, model predictive control (MPC) is the most widely used approach to address uncertainty in real-time stage. An MPC with feedback for real-time dispatch of CHP was proposed in [21]. In [22], a risk-averse MPC was applied to real-time energy hub dispatch. MPC is easily applied because it directly employs the forecasting information of uncertainties to evaluate the impact of current decision on future. However, it has drawbacks of high computational effort and dependence on forecasting information. Specifically, a reliable MPC based dispatch requires comprehensive forecasting information. With the increasing penetration of flexible resources in IEHS, such information may be unavailable in real-

time. Besides, solving a multistage problem at each stage causes extra computational burdens, thus the computation time of MPC may not satisfy the real-time requirement. As a result, there remains a new methodology for real-time IEHS dispatch.

Alternatively, without the necessity of forecasting information, ADP provides computationally efficient and almost optimal real-time dispatch which is adapted to the stochastic factors observed overtime. It has been successfully applied in similar problems, e.g., real-time dispatch of ESS [23], microgrid [24], and electric vehicles [25]. However, the major drawback of these studies is that the objective is to derive a risk neutral policy which minimizes the expectation of sequential costs. Note that uncertainties bring risks, although a risk neutral policy is economical in most scenarios, it neglects critical scenarios which rare happen but cause unexpected outcomes [26], e.g., high costs or even networks constraints violations. Hence, it is crucial to consider the risk caused by uncertainties and develop a risk-averse optimal policy, which minimizes total costs as well as reducing the risk of high costs in critical scenarios. To construct a risk-averse policy, risk-averse MPC [22] can be used. However, its high computational burden caused by large number of scenarios is contradictory to the real-time requirement. ADP using cost function approximations for risk-averse energy hub dispatch was proposed in [19]. However, networks constraints were not considered and its performance was sensitive to the choice of basis functions. Another successful implementation is the stochastic dual dynamic programming with risk measure [26,27]. However, the slow convergence and requirement of problem linearity limit its applications. Hence, it is necessary to investigate a computationally efficient and risk-averse policy for real-time stochastic dispatch of IEHS.

This paper aims at filling the gap left by the above mentioned studies, including insufficient consideration of uncertainties, incapability of real-time optimization, and insufficient consideration of risk measure. To achieve this goal, a risk-averse real-time dispatch for IEHS as well as its ADP based solution methodology is proposed, which main contributions are summarized as:

•*Problem formulation*: for the consideration of both economy and risk aversion, a novel risk-averse real-time optimal dispatch of IEHS is proposed. The problem is further reformulated as a risk-averse dynamic programming using a nested conditional value at risk (CVaR) measure [29]. Multivariate uncertainties are incorporated and multiple flexible resources, e.g., flexible heat load (FHL), demand side resource aggregator (DRA), and energy storage system (ESS), are jointly optimized while satisfying networks constraints.

•*Algorithm design*: to cope with the computation complexities brought by the dimensions of original problem, a novel ADP based real-time dispatch framework is established. A data-driven ADP with risk measure (RADP), which is modified from the basis of the method in [23], is then employed to exploit knowledge from training samples. Thus, the almost optimal and computationally efficient policy is achieved. To our best knowledge, it is the first attempt to apply risk-averse dynamic programming to real-time optimal dispatch of IEHS and develop an ADP based real-time solution.

•*Performance evaluation*: effectiveness of the proposed model and RADP in facilitating economy and avoiding risk is comprehensively validated through comparative simulations. The impacts of risk preferences are analyzed. The scalability and computation efficiency of RADP are also verified in comparison with MPC policy.

The remainder of this paper is organized as follows: IEHS modeling is presented in Section 2. In Section 3, risk-averse real-time dispatch of IEHS is formulated. In Section 4, RADP based solution technology is introduced. Section 5 presents simulation results. Finally, conclusions are made in Section 6.

2. IEHS modeling

2.1. System description and assumption

The IEHS studied in this paper comprises CHP, FHL, DRA, ESS, RES [30], and STATCOM, whose constraints are explicitly formulated, as well as the electric distribution networks (EDN) and heat distribution networks (HDN). DRA aggregates shiftable electrical loads, e.g., electric vehicles, in a specific region. IEHS is envisaged to be capable of utilizing flexibilities of DRA and FHL to improve its reliability and economy. HDN is assumed to operate in a mainly used constant flow and variable temperature mode [4,7–14] which means the mass flow rates of CHP, heat load, and HDN are constant. Thus it leads to a linear formulation of HDN instead of a non-convex formulation.

2.2. IEHS modeling

1) ***CHP***: The operation region of a CHP can be illustrated by a convex combination of its extreme operation points, as:

$$H_n^{\mathrm{CHP}}(t)=\sum_{c\in N_{\mathrm{P},n}^{\mathrm{CHP}}}\delta_{n,c}^{\mathrm{CHP}}(t)H_{n,c}^{\mathrm{CHP}},\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T, \tag{1a}$$

$$P_n^{\mathrm{CHP}}(t)=\sum_{c\in N_{\mathrm{P},n}^{\mathrm{CHP}}}\delta_{n,c}^{\mathrm{CHP}}(t)P_{n,c}^{\mathrm{CHP}},\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T, \tag{1b}$$

$$\sum_{c\in N_{\mathrm{P},n}^{\mathrm{CHP}}}\delta_{n,c}^{\mathrm{CHP}}(t)=1,\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T,\ \forall c\in N_{\mathrm{P},n}^{\mathrm{CHP}}, \tag{1c}$$

$$0\le\delta_{n,c}^{\mathrm{CHP}}(t),\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T,\ \forall c\in N_{\mathrm{P},n}^{\mathrm{CHP}}, \tag{1d}$$

$$\underline{Q}_n^{\mathrm{CHP}}\le Q_n^{\mathrm{CHP}}(t)\le\overline{Q}_n^{\mathrm{CHP}},\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T. \tag{1e}$$

The operation cost of a CHP $C_n^{\mathrm{CHP}}(t)$ is formulated as:

$$\begin{aligned}C_n^{\mathrm{CHP}}t\Big)=a_{n,5}P_n^{\mathrm{CHP}}t\Big)\Big)(t))^2+a_{n,4}H_n^{\mathrm{CHP}}(t))^2\\+a_{n,3}P_n^{\mathrm{CHP}}\Big(t\Big)H_n^{\mathrm{CHP}}\Big(t\Big)+a_{n,2}P_n^{\mathrm{CHP}}\Big(t\Big)+a_{n,1}H_n^{\mathrm{CHP}}\Big(t\Big)\\+a_{n,0},\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T.\end{aligned} \tag{2}$$

The approximation method [31] is adopted to approximate the non-convex Eq. (2), whose details are shown in Appendix A.

The relationship between heat power of CHP and its flow temperature are expressed in (3a), and the flow temperature is bounded in (3 b) and (3c)

$$H_n^{\mathrm{CHP}}(t)=C_{\mathrm{wt}}m_n^{\mathrm{CHP}}\Big(\tau_n^{\mathrm{CHP,S}}(t)-\tau_n^{\mathrm{CHP,R}}(t)\Big),\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T, \tag{3a}$$

$$\underline{\tau}_n^{\mathrm{CHP,S}}\le\tau_n^{\mathrm{CHP,S}}(t)\le\overline{\tau}_n^{\mathrm{CHP,S}},\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T, \tag{3b}$$

$$\underline{\tau}_n^{\mathrm{CHP,R}}\le\tau_n^{\mathrm{CHP,R}}(t)\le\overline{\tau}_n^{\mathrm{CHP,R}},\ \forall n\in N^{\mathrm{CHP}},\ \forall t\in T. \tag{3c}$$

2) Heat load: The operation constraints of heat load are formulated as

$$H_n^{\mathrm{IHL}}(t)+H_n^{\mathrm{FHL}}(t)=C_{\mathrm{wt}}m_n^{\mathrm{HL}}\left(\tau_n^{\mathrm{HL,S}}(t)-\tau_n^{\mathrm{HL,R}}(t)\right),\ \forall n\in N^{\mathrm{HL}},\ \forall t\in T, \tag{4a}$$

$$\underline{\tau}_n^{\mathrm{HL,S}}\le\tau_n^{\mathrm{HL,S}}(t)\le\overline{\tau}_n^{\mathrm{HL,S}},\ \forall n\in N^{\mathrm{HL}},\ \forall t\in T, \tag{4b}$$

$$\underline{\tau}_n^{\mathrm{HL,R}}\le\tau_n^{\mathrm{HL,R}}(t)\le\overline{\tau}_n^{\mathrm{HL,R}},\ \forall n\in N^{\mathrm{HL}},\ \forall t\in T. \tag{4c}$$

The heat power consumption of FHL is adjustable by using indoor temperature control [12,17]. Thus its heat power consumption is affected by indoor temperature setpoint, ambient temperature, and thermal resistance of building. The discrete thermodynamic equation [12,17] is used to formulate the temperature transition and the heat demand of FHL, as

$$H_n^{\mathrm{FHL}}(t)=C_{\mathrm{air}}\left(\tau_n^{\mathrm{FHL}}(t)-\tau_n^{\mathrm{FHL}}(t-1)\right)-\left(\tau^{\mathrm{AM}}(t)-\tau_n^{\mathrm{FHL}}(t)\right)\Delta t/r_n^{\mathrm{FHL}},\ \forall n\in N^{\mathrm{HL}},\ \forall t\in T. \tag{5}$$

The indoor temperature should be within a range to ensure the user's comfortability, as

$$\underline{\tau}_n^{\mathrm{FHL}}\le\tau_n^{\mathrm{FHL}}(t)\le\overline{\tau}_n^{\mathrm{FHL}},\ \forall n\in N^{\mathrm{HL}},\ \forall t\in T. \tag{6}$$

3) ***HDN***: an HDN consists of a supply pipeline network and a return pipeline network. Mass flow rate of pipeline should satisfy mass flow equation, as

$$\sum_{l\in N_{\mathrm{out},k}^{\mathrm{node}}}m_{kl}^{\mathrm{pipe,S}}-\sum_{j\in N_{\mathrm{in},k}^{\mathrm{node}}}m_{jk}^{\mathrm{pipe,S}}=m_k^{\mathrm{CHP}}-m_k^{\mathrm{HL}},\ \forall k\in N^{\mathrm{node}}, \tag{7a}$$

$$\sum_{l\in N_{\mathrm{in},k}^{\mathrm{node}}}m_{kl}^{\mathrm{pipe,R}}-\sum_{j\in N_{\mathrm{out},k}^{\mathrm{node}}}m_{jk}^{\mathrm{pipe,R}}=m_k^{\mathrm{HL}}-m_k^{\mathrm{CHP}},\ \forall k\in N^{\mathrm{node}}. \tag{7b}$$

Heat loss inevitably occurs as water flows in pipeline, which results in different inlet and outlet flow temperature of pipeline. Since the spatial scale of HDN studied in this paper is small, a thermal equilibrium is assumed to be achieved at the end of each time-slot [8]. Thus heat storage effect of pipeline is not considered and inlet/outlet flow temperature of pipeline should follow (8a) and (8 b),

$$\tau_{kl}^{\mathrm{pipe,S,out}}(t)=d_{kl}^{\mathrm{pipe}}\left(\tau_{kl}^{\mathrm{pipe,S,in}}(t)-\tau^{\mathrm{AM}}(t)\right)+\tau^{\mathrm{AM}}(t), \tag{8a}$$

$$\tau_{kl}^{\mathrm{pipe,R,out}}(t)=d_{kl}^{\mathrm{pipe}}\left(\tau_{kl}^{\mathrm{pipe,R,in}}(t)-\tau^{\mathrm{AM}}(t)\right)+\tau^{\mathrm{AM}}(t), \tag{8b}$$

where d_{kl}^{pipe} depends on heat transfer properties and mass flow rate of pipeline [8], and always smaller than 1.

Water from different pipelines is mixed at a confluence node, the mixed flow temperature can be obtained according to the energy conversion law, as

$$\sum_{l\in N_{\mathrm{in},k}^{\mathrm{node}}}\tau_{kl}^{\mathrm{pipe,S,out}}(t)m_{kl}^{\mathrm{pipe,S}}+\tau_k^{\mathrm{CHP,S}}(t)m_k^{\mathrm{CHP}}=\tau_k^{\mathrm{node,S}}(t)\left(\sum_{l\in N_{\mathrm{in},k}^{\mathrm{node}}}m_{kl}^{\mathrm{pipe,S}}+m_k^{\mathrm{CHP}}\right),\ \forall k\in N^{\mathrm{node}}, \tag{9a}$$

$$\sum_{l\in N_{\mathrm{out},k}^{\mathrm{node}}}\tau_{kl}^{\mathrm{pipe,R,out}}(t)m_{kl}^{\mathrm{pipe,R}}+\tau_k^{\mathrm{HL,R}}(t)m_k^{\mathrm{HL}}=\tau_k^{\mathrm{node,R}}(t)\left(\sum_{l\in N_{\mathrm{out},k}^{\mathrm{node}}}m_{kl}^{\mathrm{pipe,R}}+m_k^{\mathrm{HL}}\right),\ \forall k\in N^{\mathrm{node}}. \tag{9b}$$

The inlet flow temperature of pipeline should be equal to the mixed flow temperature. Besides, flow temperature of supply pipeline of heat load/return pipeline of CHP should be equal to the mixed flow temperature at supply/return terminal, as

$$\tau_{kl}^{\mathrm{pipe,S,in}}(t)=\tau_k^{\mathrm{node,S}}(t)=\tau_k^{\mathrm{HL,S}}(t),\ \forall k\in N^{\mathrm{node}},\ \forall l\in N_{\mathrm{out},k}^{\mathrm{node}}, \tag{10a}$$

$$\tau_{kl}^{\mathrm{pipe,R,in}}(t)=\tau_k^{\mathrm{node,R}}(t)=\tau_k^{\mathrm{CHP,R}},\ \forall k\in N^{\mathrm{node}},\ \forall l\in N_{\mathrm{out},k}^{\mathrm{node}}. \tag{10b}$$

4) ***DRA***: demand side resources, e.g., electric vehicles, provide extra flexibilities for IEHS dispatch. In this paper coordinated operation in DRA level, e.g., an electric vehicle fleet, is considered. According to [25,32], the flexibility of a shiftable DRA can be modeled by its aggregated power limits ($\overline{P}_n^{\mathrm{DRA}}(t)$, $\underline{P}_n^{\mathrm{DRA}}(t)$) and cumulative energy limits ($\overline{E}_n^{\mathrm{DRA}}(t)$, $\underline{E}_n^{\mathrm{DRA}}(t)$). Demand shedding occurs as cumulative energy trajectory $E_n^{\mathrm{DRA}}(t)$ is beneath $\underline{E}_n^{\mathrm{DRA}}(t)$. The operation constraints of DRA can be formulated by

$$E_n^{\mathrm{DRA}}(t)=\sum_{m=1}^{t}P_n^{\mathrm{DRA}}(m)\Delta t,\ \forall t\in T,\ \forall n\in N^{\mathrm{DRA}}, \tag{11a}$$

$$E_{\mathrm{shed},n}^{\mathrm{DRA}}(t)=\sum_{m=1}^{t}P_{\mathrm{shed},n}^{\mathrm{DRA}}(m)\Delta t,\ \forall t\in T,\ \forall n\in N^{\mathrm{DRA}}, \tag{11b}$$

$$\underline{E}_n^{\mathrm{DRA}}(t)\le E_n^{\mathrm{DRA}}(t)+E_{\mathrm{shed},n}^{\mathrm{DRA}}(t)\le\overline{E}_n^{\mathrm{DRA}}(t),\ \forall t\in T,\ \forall n\in N^{\mathrm{DRA}}, \tag{11c}$$

$$\underline{P}_n^{\mathrm{DRA}}(t)\le P_n^{\mathrm{DRA}}(t)\le\overline{P}_n^{\mathrm{DRA}}(t),\ \forall t\in T,\ \forall n\in N^{\mathrm{DRA}}, \tag{11d}$$

$$0\le P_{\mathrm{shed},n}^{\mathrm{DRA}}(t)\le\overline{P}_n^{\mathrm{DRA}}(t),\ \forall t\in T,\ \forall n\in N^{\mathrm{DRA}}. \tag{11e}$$

DRA is assumed to be operated at a constant power factor φ_n, thus its reactive power consumption can be formulated as

$$Q_n^{\mathrm{DRA}}(t)=P_n^{\mathrm{DRA}}(t)\sqrt{(1-\varphi_n^2)/\varphi_n^2},\ \forall t\in T,\ \forall n\in N^{\mathrm{DRA}}. \tag{12}$$

DRA will receive a compensation for demand shedding, which is described by a positive and linear function, as follows

$$C_n^{\mathrm{DRA}}(\mathrm{t})=\beta_{\mathrm{DRA}}P_{\mathrm{shed},n}^{\mathrm{DRA}}(t)\Delta t,\ \forall t\in T,\ \forall n\in N^{\mathrm{DRA}}. \tag{13}$$

5) ***RES***: active/reactive power outputs of RES [33] are assumed to be controllable, as

$$0\le P_n^{\mathrm{RES}}(t)\le\overline{P}_n^{\mathrm{RES}}(t),\ \forall t\in T,\ \forall n\in N^{\mathrm{RES}}, \tag{14a}$$

$$0\le Q_n^{\mathrm{RES}}(t)\le\overline{Q}_n^{\mathrm{RES}}(t),\ \forall t\in T,\ \forall n\in N^{\mathrm{RES}}. \tag{14b}$$

The operation cost of RES is assumed to be zero, but RES

curtailment will be penalized [34], as

$$C_n^{\mathrm{RES}}(t)=\beta_{\mathrm{RES}}\left(\overline{P}_n^{\mathrm{RES}}(t)-P_n^{\mathrm{RES}}(t)\right)\Delta t,\ \forall t\in T,\ \forall n\in N^{\mathrm{RES}}. \tag{15}$$

6) ***ESS***: the constraints of ESS are shown in (16a) to (16f)

$$0\le P_{\mathrm{c},n}^{\mathrm{ESS}}(t)\le \overline{P}_{\mathrm{c},n,}^{\mathrm{ESS}},\ \forall t\in T,\ \forall n\in N^{\mathrm{ESS}}, \tag{16a}$$

$$0\le P_{\mathrm{d},n}^{\mathrm{ESS}}(t)\le \overline{P}_{\mathrm{d},n,}^{\mathrm{ESS}},\ \forall t\in T,\ \forall n\in N^{\mathrm{ESS}}, \tag{16b}$$

$$\underline{E}_n^{\mathrm{ESS}}\le E_n^{\mathrm{ESS}}(t)\le \overline{E}_n^{\mathrm{ESS}},\ \forall t\in T,\ \forall n\in N^{\mathrm{ESS}}, \tag{16c}$$

$$E_n^{\mathrm{ESS}}(t)=E_n^{\mathrm{ESS}}(t-1)+\left(\eta_{\mathrm{ESS,c}}P_{\mathrm{c},n}^{\mathrm{ESS}}(t)-P_{\mathrm{d},n}^{\mathrm{ESS}}(t)\times\Big/\eta_{\mathrm{ESS,d}}\right)\Delta t,\ \forall t\in T,\ \forall n\in N^{\mathrm{ESS}}, \tag{16d}$$

$$E_n^{\mathrm{ESS}}(0)=E_n^{\mathrm{ESS}}(|T|)=0.5\overline{E}_n^{\mathrm{ESS}},\ \forall t\in T,\ \forall n\in N^{\mathrm{ESS}}, \tag{16e}$$

$$P_{\mathrm{c},n}^{\mathrm{ESS}}(t)P_{\mathrm{d},n}^{\mathrm{ESS}}(t)=0,\ \forall t\in T,\ \forall n\in N^{\mathrm{ESS}}. \tag{16f}$$

The operation cost of ESS is formulated as a linear function for charging and discharging [35], as

$$C_n^{\mathrm{ESS}}(t)=\beta_{\mathrm{ESS}}\left(P_{\mathrm{c},n}^{\mathrm{ESS}}(t)+P_{\mathrm{d},n}^{\mathrm{ESS}}(t)\right)\Delta t,\ \forall t\in T,\ \forall n\in N^{\mathrm{ESS}}. \tag{17}$$

Although Eq. (17) is an increasing positive function for charging and discharging, complementarity constraint (16f) cannot be exactly relaxed since RES curtailment penalty is considered. To avoid involving this non-convex constraint, $\eta_{\mathrm{ESS,c}}$ and $\eta_{\mathrm{ESS,d}}$ are assumed to be 1.

7) ***STATCOM***: IEHS is equipped with STATCOM for reactive power support. Its operation constraint are given as

$$\underline{Q}_n^{\mathrm{SC}}\le Q_n^{\mathrm{SC}}(t)\le \overline{Q}_n^{\mathrm{SC}},\ \forall t\in T,\ \forall n\in N^{\mathrm{SC}}. \tag{18}$$

8) ***EDN***: The linearized AC power flow model [36] is employed to describe power flow in EDN, yields

$$P_i(t)=\sum_{j=1,j\ne i}^{|N^{\mathrm{bus}}|}\left(\frac{x_{ij}}{r_{ij}^2+x_{ij}^2}\left(\theta_i(t)-\theta_j(t)\right)+\frac{r_{ij}}{r_{ij}^2+x_{ij}^2}\left(u_i(t)-u_j(t)\right)\right),\ \forall t\in T,\ \forall i\in N^{\mathrm{bus}}, \tag{19a}$$

$$Q_i(t)=\sum_{j=1,j\ne i}^{|N^{\mathrm{bus}}|}\left(\frac{x_{ij}}{r_{ij}^2+x_{ij}^2}\left(u_i(t)-u_j(t)\right)-\frac{r_{ij}}{r_{ij}^2+x_{ij}^2}\left(\theta_i(t)-\theta_j(t)\right)\right),\ \forall t\in T,\ \forall i\in N^{\mathrm{bus}}, \tag{19b}$$

$$P_{ij}(t)=\frac{r_{ij}\left(u_i(t)-u_j(t)\right)+x_{ij}\left(\theta_i(t)-\theta_j(t)\right)}{r_{ij}^2+x_{ij}^2},\ \forall t\in T,\ \forall ij\in N^{\mathrm{line}}, \tag{19c}$$

$$Q_{ij}(t)=\frac{x_{ij}\left(u_i(t)-u_j(t)\right)-r_{ij}\left(\theta_i(t)-\theta_j(t)\right)}{r_{ij}^2+x_{ij}^2},\ \forall t\in T,\ \forall ij\in N^{\mathrm{line}}, \tag{19d}$$

$$-\overline{P}_{ij}\le P_{ij}(t)\le \overline{P}_{ij},\ \forall t\in T,\ \forall ij\in N^{\mathrm{line}}, \tag{19e}$$

$$\underline{u}\le u_i(t)\le \overline{u},\ \forall t\in T,\ \forall i\in N^{\mathrm{bus}}. \tag{19f}$$

where r_{ij}/x_{ij} represents resistance/reactance of line ij; $\overline{P}_{ij}$ is the flow limit; $P_i(t)/\ Q_i(t)$ is the active/reactive power injection; $P_{ij}(t)/\ Q_{ij}(t)$ is the active/reactive branch flow; $\theta_i(t)$ is the voltage angle of bus i; Finally, $u_i(t)$ is the voltage magnitude of bus i and $\overline{u}/\underline{u}$ are its upper/lower bounds.

The active/reactive power injections except the slack bus (numbered as bus 1) should follow

$$P_i(t)=P_i^{\mathrm{CHP}}(t)+P_i^{\mathrm{RES}}(t)-P_{\mathrm{c},i}^{\mathrm{ESS}}(t)+P_{\mathrm{d},i}^{\mathrm{ESS}}(t)-P_i^{\mathrm{DRA}}(t)-P_i^{\mathrm{EL}}(t),i\ne 1, \tag{20a}$$

$$Q_i(t)=Q_i^{\mathrm{CHP}}(t)+Q_i^{\mathrm{RES}}(t)+Q_i^{\mathrm{SC}}(t)-Q_i^{\mathrm{CON}}(t)-Q_i^{\mathrm{EL}}(t),i\ne 1. \tag{20b}$$

9) ***Stage cost***: The stage cost of IEHS at time-slot t C_t includes operational cost and penalty cost. Specifically, the operation cost includes cost of purchasing electricity from upper level power system, operation cost of CHP and ESS. Penalty cost includes demand shedding penalty and RES curtailment penalty.

$$C(t)=\varepsilon(t)P_1(t)+\sum_{n\in N^{\mathrm{CHP}}}C_n^{\mathrm{CHP}}(t)+\sum_{n\in N^{\mathrm{DRA}}}C_n^{\mathrm{DRA}}(t)+\sum_{n\in N^{\mathrm{ESS}}}C_n^{\mathrm{ESS}}(t)+\sum_{n\in N^{\mathrm{RES}}}C_n^{\mathrm{RES}}(t),\ \forall t\in T. \tag{21}$$

3. Problem formulation

3.1. Objective for risk-averse real-time dispatch of IEHS

Here, electrical and heat loads, maximum RES output, ambient temperature, real-time price, and cumulative energy boundaries and power limits of DRA are assumed to be stochastic [32] and their actual values are unknown to IEHS in advance. At time-slot t, IEHS operator makes decisions based on the stochastic factors realized at current stage and their future uncertainties. Thus real-time dispatch of IEHS is a multistage stochastic sequential optimization problem. Typically, the objective is to derive a policy π which minimizes the expectation of sequential costs in risk neutral case, as

$$\min_{\pi\in\Pi}\mathbf{E}\left(C_1,\ldots,C_t,\ldots,C_{|T|}\right), \tag{22}$$

However, uncertainties bring risks. Although a risk neutral policy is economical in most scenarios, it usually neglects the

possible high costs in critical scenarios which rare happen. For example, IEHS operator may decide to postpone the energy consumption until a future certain stage where the real-time price may be low or RES may has large output. If RES output suddenly decreases at that stage, a high penalty occurs because demand shedding has to be implemented in order to satisfy networks constraints. Thus, for a risk-sensitive IEHS operator, a risk-averse policy, which minimizes costs as well as avoid the risk of high costs in critical scenarios, is more attractive.

To fully consider the risk-aversion in sequential optimization, the framework of dynamic programming with dynamic risk measure proposed in [28] is adopted. By replacing the expectation operator with dynamic risk measure $\rho_{1,|T|}$, the risk neutral problem (22) becomes a risk-averse dynamic programming problem, as

$$\min_{\pi\in\Pi}\rho_{1,|T|}\left(C_1,\ldots,C_t(t),\ldots,C_{|T|}\right). \tag{23}$$

According to [28], problem (23) can be rewritten as problem (24) by using the nested single stage coherent risk measures ρ_t if $\rho_{1,|T|}$ is time-consistent

$$\min_{\pi\in\Pi} C_1+\rho_1\left(C_2\ldots+\rho_{t-1}\left(C_t+\ldots+\rho_{|T|-1}\left(C_{|T|}\right)\right)\right), \tag{24}$$

where "coherent" means ρ_t satisfies convexity, monotonicity, positive homogeneity, and translation equivariance [27].

Meanwhile, CVaR is adopted as a widely used coherent risk measure. By given a random stage cost C_t and percentile α_t, Value-at-Risk (VaR) is the smallest cost after excluding worse scenarios whose probabilities are less that (1-α_t), while CVaR is the expectation of loss exceeding VaR. CVaR of C_t can be obtained by (25).

$$\mathrm{CVaR}_{\alpha_t}(C_t)=u+\frac{\int_{x\in C_t}[(x-u)]^+p(x)\mathrm{d}x}{1-\alpha_t}, \tag{25}$$

where u is the Value-at-Risk of C_t with α_t, $p(x)$ is the possibility of $C_t=x$, and $[y]^+=\max(y,0)$.

The single stage coherent risk measures ρ_t for random stage cost is formulated as a convex combination of its expectation and CVaR value, which has been successfully used in other risk-averse dynamic programming problems, as

$$\rho_{\omega_t}(C_t)=(1-\lambda_t)\mathbf{E}[C_t]+\lambda_t\mathrm{CVaR}_{\alpha_t}(C_t), \tag{26}$$

where λ_t is the weighted coefficient and ω_t=(α_t, λ_t) represents the risk-aversion preference. The larger α_t and λ_t, a more risk-sensitive RADP is. Note that if $\lambda_t=0$, problem (24) degenerates into a risk neutral dynamic programming problem.

The objective of risk-averse real-time dispatch of IEHS is determined by (23) or (24), which can be interpreted as minimizing the risk-adjusted sequential costs incurred by policy π using a nested CVaR measure.

3.2. Dynamic programming based problem reformulation

In order to obtain the risk-averse optimal policy for (24), a dynamic programming based reformulation is firstly presented. The advantage of the formulation in (24) is a similar structure with traditional dynamic programming, which can be solved by Bellman recursion. Some important notations of the dynamic programming based reformulation [37] are defined as follows.

Decision variable: decision variables X_t of IEHS at time-slot t are given as

$$\begin{aligned}X_t=\Big\{&H_n^{\mathrm{CHP}}(t),P_n^{\mathrm{CHP}}(t),\tau_n^{\mathrm{FHL}}(t),P_n^{\mathrm{DRA}}(t),P_{\mathrm{shed},n}^{\mathrm{DRA}},\\&P_{\mathrm{c},n}^{\mathrm{ESS}}(t),P_{\mathrm{d},n}^{\mathrm{ESS}}(t),P_n^{\mathrm{RES}}(t),Q_n^{\mathrm{RES}}(t),Q_n^{\mathrm{SC}}(t)\Big\}\end{aligned} \tag{27}$$

Exogenous information: exogenous information W_t contains the stochastic factors observed at time-slot t

$$\begin{aligned}W_t=\Big\{&\varepsilon(t),P_i^{\mathrm{EL}}(t),Q_i^{\mathrm{EL}}(t),H_n^{\mathrm{IHL}}(t),\tau^{\mathrm{AM}}(t),\\&\overline{P}_n^{\mathrm{RES}}(t),\overline{Q}_n^{\mathrm{RES}}(t),\overline{E}_n^{\mathrm{DRA}}(t),\underline{E}_n^{\mathrm{DRA}}(t),\overline{P}_n^{\mathrm{DRA}}(t),\underline{P}_n^{\mathrm{DRA}}(t)\Big\}\end{aligned} \tag{28}$$

System state: system state contains information for IEHS operator to make decision, which includes exogenous information and resource state R_t. Resource state of IEHS are defined as

$$R_n^{\mathrm{FHL}}(t)=\tau_n^{\mathrm{FHL}}(t)-\underline{\tau}_n^{\mathrm{FHL}},\ \forall n\in N^{\mathrm{FHL}}, \tag{29a}$$

$$R_n^{\mathrm{DRA}}(t)=\overline{E}_n^{\mathrm{DRA}}(t)-E_n^{\mathrm{DRA}}(t)-E_{\mathrm{shed},n}^{\mathrm{DRA}}(t),\ \forall n\in N^{\mathrm{DRA}}, \tag{29b}$$

$$R_n^{\mathrm{ESS}}(t)=E_n^{\mathrm{ESS}}(t)-\underline{E}_n^{\mathrm{ESS}},\ \forall n\in N^{\mathrm{ESS}}, \tag{29c}$$

$$R_t=\left\{R_n^{\mathrm{FHL}}(t),R_n^{\mathrm{DRA}}(t),R_n^{\mathrm{ESS}}(t)\right\}. \tag{29d}$$

Pre-decision states S_t is the system state after W_t update but before decision X_t is made, while post-decision state S_t^X is the system state after decision X_t but before W_{t+1} arrives [23]. Pre-decision state and post-decision state of IEHS at time-slot t are expressed in (30a)-(30 b)

$$S_t=\{W_t,R_t\}, \tag{30a}$$

$$S_t^X=\left\{W_t,R_t^X\right\}. \tag{30b}$$

State transition: state transition describes the evolution of IEHS by given the state and decision variables, which is formulated in (31a)-(31c)

$$R_n^{\mathrm{FHL},X}(t)=R_n^{\mathrm{FHL}}(t)+\tau_n^{\mathrm{FHL}}(t)-\tau_n^{\mathrm{FHL}}(t-1),\ \forall n\in N^{\mathrm{FHL}}, \tag{31a}$$

$$\begin{aligned}R_n^{\mathrm{ESS},X}(t)=R_n^{\mathrm{ESS}}(t)+\Big(&\eta_{\mathrm{ESS,c}}P_{\mathrm{c},n}^{\mathrm{ESS}}(t)-P_{\mathrm{d},n}^{\mathrm{ESS}}(t)\\&\times\Big/\eta_{\mathrm{ESS,d}}\Big)\Delta t,\ \forall n\in N^{\mathrm{ESS}},\end{aligned} \tag{31b}$$

$$R_n^{\mathrm{DRA},X}(t)=R_n^{\mathrm{DRA}}(t)-\left(P_n^{\mathrm{DRA}}(t)+P_{\mathrm{shed},n}^{\mathrm{DRA}}(t)\right)\Delta t,\ \forall n\in N^{\mathrm{DRA}}. \tag{31c}$$

It has been proven in [28] that problem (24) has the optimal policy which can be obtained by the risk-averse Bellman recursion in (32), yields

$$V_t(S_t)=\min_{X_t\in\Pi_t}\left(C_t(S_t,X_t)+\rho_{\omega_t}(V_{t+1}(S_{t+1}))\right), \tag{32}$$

where $V_t(S_t)$ is the optimal value function of S_t. The optimal decision at time-slot t is given by

$$X_t=\underset{X_t\in\Pi_t}{\mathrm{argmin}}\left(C_t(S_t,X_t)+\rho_{\omega_t}(V_{t+1}(S_{t+1}))\right), \tag{33}$$

where Π_t represents the feasible region of decision variables.

4. Solution technique

Note that although the risk-averse optimal policy to real-time IEHS dispatch can be obtained by solving (33) theoretically, the complexities of the problem make it almost intractable for traditional approaches, e.g., dynamic programming, due to the following two aspects: first, problem (32) is multistage and multi-dimensional which results in enormous computational burden; second, the accuracy of the nested CVaR measure in (32) relies on the knowledge of the distributions of uncertainties which is usually inaccessible. To address this, a data driven risk-averse ADP (RADP) is employed to develop a computationally efficient and almost optimal policy.

4.1. Risk-averse ADP

In RADP, the post-decision state of IEHS S_t^X is used to approximate $\rho_{\omega t}(V_{t+1}(S_{t+1}))$ in (33) [23], thus Eq. (33) can be reformulated as

$$X_t = \underset{X_t \in \Pi_t}{\operatorname{argmin}} \left(C_t(S_t, X_t) + V_t^X\left(S_t^X\right)\right), \tag{34}$$

where V_t^X is the post-decision value function. The advantage of using the post-decision state is that it only depends on the decisions once pre-decision state is known, thus the operator $\rho_{\omega t}$ can be removed and the policy is more tractable.

Remark 1. V_t^X is convex in S_t^X for every W_{t+1} since the stage cost in (21) is a convex function, CVaR is a convex risk measure [29], and R_t^X appears as right-hand side parameters of linear constraints [27].

By exploiting the convexity of $V_t^X(S_t^X)$, a convex and pieces-wise linear function is separately generated for R_t^X to approximate V_t^X. Then, the optimal policy in RADP is given by

$$X_t = \underset{X_t \in \Pi_t}{\operatorname{argmin}} \Bigg(C_t(S_t, X_t) + \sum_{n \in N^{\mathrm{ESS}}} \sum_{g \in G} v_{n,t}^{\mathrm{ESS}}(g) y_{n,t}^{\mathrm{ESS}}(g) + \sum_{n \in N^{\mathrm{DRA}}} \sum_{g \in G} v_{n,t}^{\mathrm{DRA}}(g) y_{n,t}^{\mathrm{DRA}}(g) + \sum_{n \in N^{\mathrm{FHL}}} \sum_{g \in G} v_{n,t}^{\mathrm{FHL}}(g) y_{n,t}^{\mathrm{FHL}}(g) \Bigg), \tag{35}$$

subject to (1)–(15), (16a)-(16e), (17)–(21), (31), and (36).

$$\sum_{g \in G} y_{n,t}^{\mathrm{FHL}}(g) = R_n^{\mathrm{FHL},X}(t), \tag{36a}$$

$$0 \le y_{n,t}^{\mathrm{FHL}}(g) \le \left(\overline{\tau}_n^{\mathrm{FHL}}(t) - \underline{\tau}_n^{\mathrm{FHL}}\right) \Big/ |G|, \tag{36b}$$

$$0 \le y_{n,t}^{\mathrm{DRA}}(g) \le \left(\overline{E}_n^{\mathrm{DRA}}(t) - \underline{E}_n^{\mathrm{DRA}}(t)\right) \Big/ |G|, \tag{36c}$$

$$0 \le y_{n,t}^{\mathrm{ESS}}(g) \le \left(\overline{E}_n^{\mathrm{ESS}}(t) - \underline{E}_n^{\mathrm{ESS}}\right) \Big/ |G|, \tag{36d}$$

$$\sum_{g \in G} y_{n,t}^{\mathrm{DRA}}(g) = R_n^{\mathrm{DRA},X}(t), \tag{36e}$$

$$\sum_{g \in G} y_{n,t}^{\mathrm{ESS}}(g) = R_n^{\mathrm{ESS},X}(t). \tag{36f}$$

4.2. Data-driven based learning process of RADP

Like most ADP and reinforcement learning methods, RADP should be trained to obtain the exact slopes of segments before application. A data-driven based off-learning is employed for RADP to learn how to dispatch IEHS optimally from training samples. The training samples contain different realizations of exogenous information which can be generated by empirical data or Monte Carlo method. The proposed off-line learning originates from the risk neutral value function approximation method [23]. The method in [23] uses a single training sample per iteration to observe the expectation of cost. However, it cannot be directly applicable in risk-averse case because it is hard to obtain the risk-adjusted cost from a single sample. Instead, our modified method uses a small batch of samples and thus the observation of risk-adjusted objective value is available. The learning process of RADP can be summarized as three steps: given a set of training samples, RADP make decision based on current knowledge and the uncertainties realized at time-slot t. Then, it observes the slope of risk-adjusted value function of time-slot t-1. Finally, a projection operation is performed to retain the value function convexity. The detailed descriptions of learning process are as follows:

Step 0: Initialization

0.1) generate a set of training samples Φ by using empirical data or Monte Carlo method, set the initial segment slope v, and set maximum iteration M.

0.2) divide Φ into small batches of samples Φ_m and the number of training sample in Φ_m is Φ/M.

0.3) set $m = 1$.

Step 1: Decision making

1.1) set $t = 1$

1.2) observe the realized exogenous information in each single training sample $\varphi \in \Phi_m$.

1.3) for each φ, obtain the optimal decision $X_{\varphi,t}$ by solving (35) and update the post-decision state $R_{\varphi,t}^X$ by using (31a)-(31c).

1.4) compute the risk-adjusted objective value by using (37).

$$\widehat{F}_{t,m} = \frac{(1-\lambda_{t-1})}{|\Phi_m|} \sum_{\varphi \in \Phi_m} F_{t,m}(\varphi) + \lambda_{t-1} \mathrm{CVaR}\alpha_{t-1}(F_{t,m}), \tag{37}$$

where $F_{t,m}(\varphi) = C_{\varphi,t} + V_{\varphi,t}^X$ and its CVaR value can be computed by solving the linear problem in (38a) [29].

$$\mathrm{CVaR}\alpha_{t-1}(F_{t,m}) = \min_{u,\varpi} \left(u + \frac{1}{\alpha_{t-1}|\Phi_m|} \sum_{\varphi \in \Phi_m} \varpi(\varphi) \right), \tag{38a}$$

subject to

$$\begin{cases} \varpi(\varphi) \ge 0, \forall \varphi, \\ u - F_{t,m}(\varphi) + \varpi(\varphi) \ge 0, \forall \varphi, \end{cases} \tag{38b}$$

where ϖ is the auxiliary variable.

Step 2: Slope observation

If $t > 1$, then execute Step 2.1 to 3.2.

$$\widehat{v}_{m,n,t-1}^{i} = \left(\widehat{F}_{t,m} - \widehat{F}_{t,m}^{i,n}\right) \Big/ \delta R_{i,n}, i = \mathrm{FHL}, \mathrm{DRA}, \mathrm{ESS}, \tag{39}$$

where $\widehat{F}_{t,m}^{i,n}$ is the risk-adjusted objective after imposing a perturbation to FHL/DRA/ESS n; $\delta R_{i,n}$ is the perturbation.

Step 3: Slope update
3.1) update the slope of each FHL, DRA, and ESS, as:

$$v^{i}_{m,n,t-1} = (1-\gamma_m)v^{i}_{m-1,n,t-1} + \gamma_m \hat{v}^{i}_{m,n,t-1},\ i = \text{FHL, DRA, ESS}, \tag{40}$$

where γ_m is the step size and the harmonic step size rule is adopted in this paper.

2.1) for each FHL, DRA, and ESS, impose a perturbation to its post-decision state at t-1, respectively, and then obtain the optimal decision by solving (35).
2.2) compute the risk-adjusted objective value after perturbation, which is similar with Step 1.3.
2.3) compute the observation of the slope by using (39).
3.2) perform projection operation to retain the value function convexity by using the successive projective approximation routine method in [23].
Step 4
4.1) set $t = t+1$, return to Step 1.2) if $t<|T|$.
4.2) set $m = m+1$, return to Step 1.1) if $m < M$.

Note that if λ_t is set as 0, RADP becomes a risk neutral ADP. Besides, parallel computing is applicable when execute Step 2.1 to 3.2, by which a lot of computation time will be saved.

4.3. Overall procedures of RADP

As shown in Fig. 1, like most of the ADP approaches, RADP contains off-line learning phase and on-line application phase. In the former one, by exploiting information from training samples, RADP learns to operate IEHS under multiple uncertainties and response to any realization of exogenous information. Then, in the latter one, the risk-averse real-time optimal dispatch of IEHS is available by solving (35) stage by stage. Note that RADP only requires the newly updated system state and pre-learned knowledge for decision making, i.e., forecasting information is unnecessary and only a single stage problem is solved, which is quite computationally efficient in practical applications.

5. Case studies

5.1. Case settings

A 24-h real-time operation is conducted in a modified IEHS [8] with 33 electricity buses and 21 heat nodes, as shown in Fig. 2. Its parameters and uncertainty settings are listed in Appendix-B. To facilitate analysis of the impacts of risk-aversion preferences, α_t and λ_t of different stages are set to be identical, but they can be set differently in practical applications. 10000 training samples generated by Monte Carlo method are used to train RADP, the maximum iteration of off-line learning is set as 100, and each batch

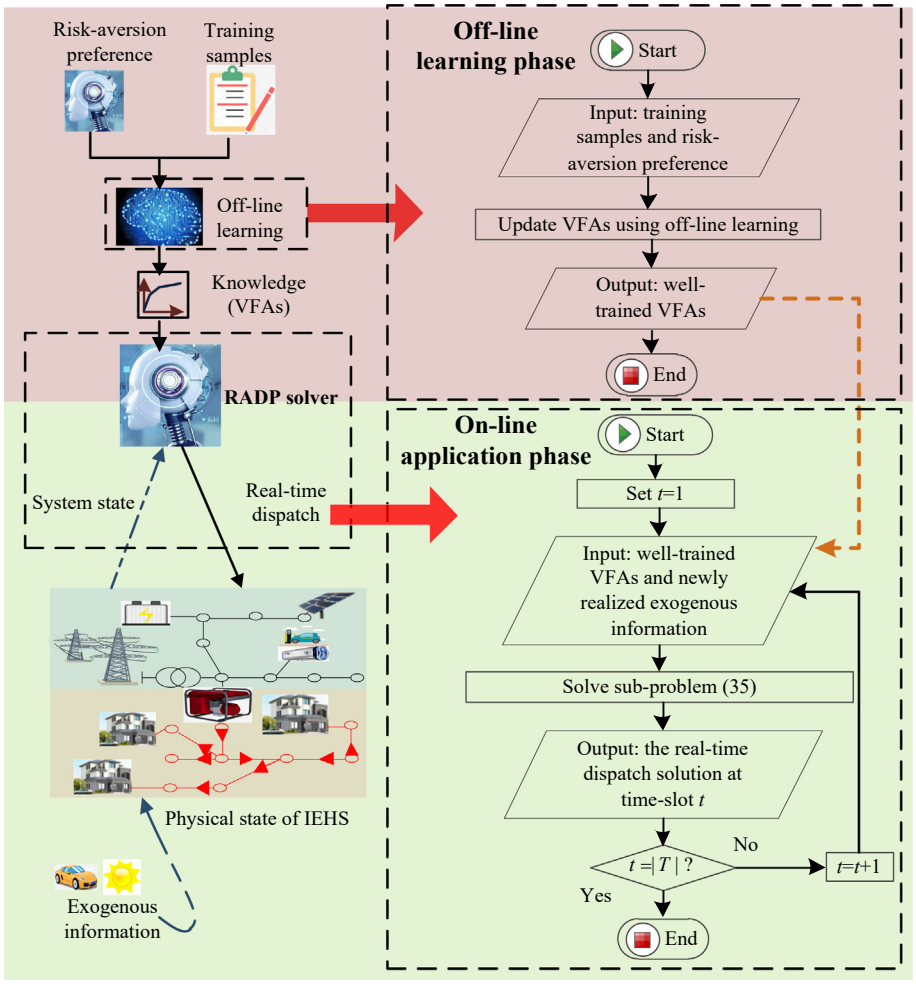


Fig. 1. Decision flows of RADP.

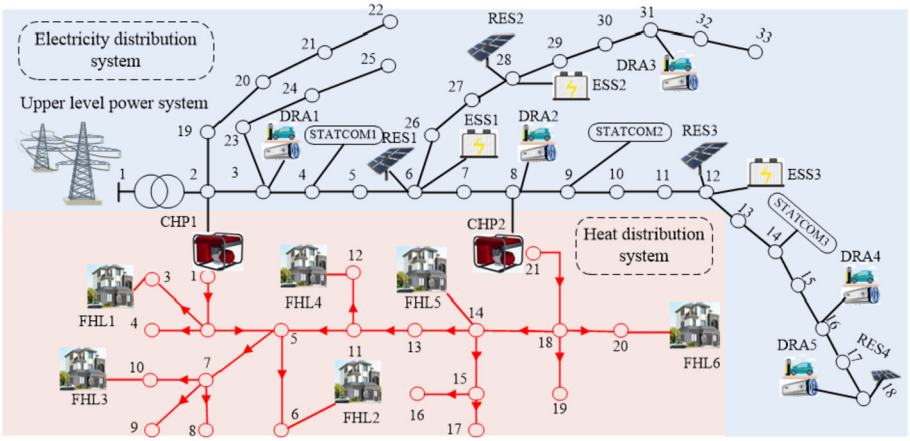


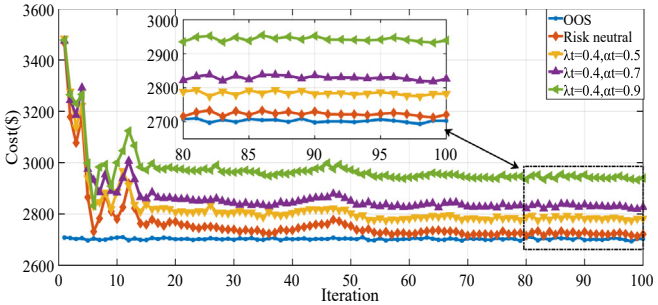
Fig. 2. IEHS topology.

contains 100 samples. The number of segment of resource state is set as 80 and the initial segment slopes are set as 0. Simulations are carried out in MATLAB R2018a and each sub-problem is solved by CPLEX.

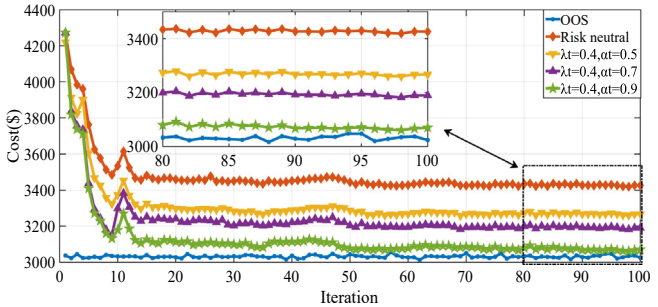
5.2. Learning process of RADP

The performance of RADP in off-line learning phase is firstly investigated. Note that the optimal solution to problem (24) should be obtained by dynamic programming, but it is intractable due to the dimension complexities. Therefore, the performance of RADP is compared with the optimal offline solution (OOS) which minimizes the total cost by assuming the perfect information of uncertainties. Fig. 3 plots the average cost and the 95%-VaR of cost (the smallest cost after excluding worse scenarios whose probabilities are less that 5%) of each batch of samples in off-line learning.

It can be found that RADP shows great convergence in off-line learning. Although RADP performs badly at first since the correct value function approximations have not been built, it learns to operate IEHS under multiple uncertainties after about 70 iterations. This is because by exploiting the convexity of the value function, there is no need to explore all the state space, thus the optimal policy is tractable. It is obvious that RADP with different risk-aversion preferences converges to different policies. Fig. 3(a) demonstrates that a risk neutral policy ($\lambda_t = 0$) achieves almost the same average cost with OOS after learning. While for a more risk-averse RADP, it leads to higher average cost. In contrast, Fig. 3(b) illustrates that the 95%-VaR of cost of a more risk-averse RADP is closer to OOS. While a risk neutral policy leads to higher cost variability.



(a)Average cost.



(b) 95%-VaR of cost.

Fig. 3. Performance of RADP in off-line learning.

5.3. Impacts of different risk-aversion preferences on RADP

After sufficient learning, RADP can be used in practical applications. In this section, 1000 random samples are used to investigate the impacts of different risk-aversion preferences on the optimal dispatch of RADP. β_{RES} and β_{DRA} are set as \$0.01/kWh and \$0.5/kWh. Fig. 4 illustrates that IEHS becomes more risk-sensitive and conservative as α_t increases, which tends to store more energy in ESSs, increases the energy consumptions of DRAs at time-slots 8–10 and 17–21, and increases the indoor temperature of FHLs. This is because such policy can reduce the possible high cost in critical scenarios, e.g., RES power shortage at time-slot 11–15 or high real-time price at time-slot 21–24.

Table 1 lists the average value and 95%-VaR of cost under different α_t and λ_t. It illustrates a more aggressive policy incurs a smaller mean cost but causes much higher risks. Although a more risk-sensitive RADP, i.e., with larger α_t and λ_t, leads to a higher mean cost, the cost variability is significantly reduced. For example, when $\alpha_t = 0.8$ and $\lambda_t = 0.4$, RADP incurs an 8.5% growth in mean cost but a 10.3% reduction in 95%-VaR value in comparison to the risk neutral case. These facts indicate that different risk-aversion preferences lead to different dispatch decisions and RADP offers policies which allow IEHS operator to make a compromise between expected cost and risk.

5.4. Impacts of different penalty coefficients on RADP

Fig. 5 presents the dispatch results of RADP under different β_{RES}. It can be found that the larger β_{RES}, the less energy stored in ESSs at time-slot 6 to 15 and the less energy consumptions of DRAs at time-slot 6 to 9. This is because IEHS tends to reserve more capacity to

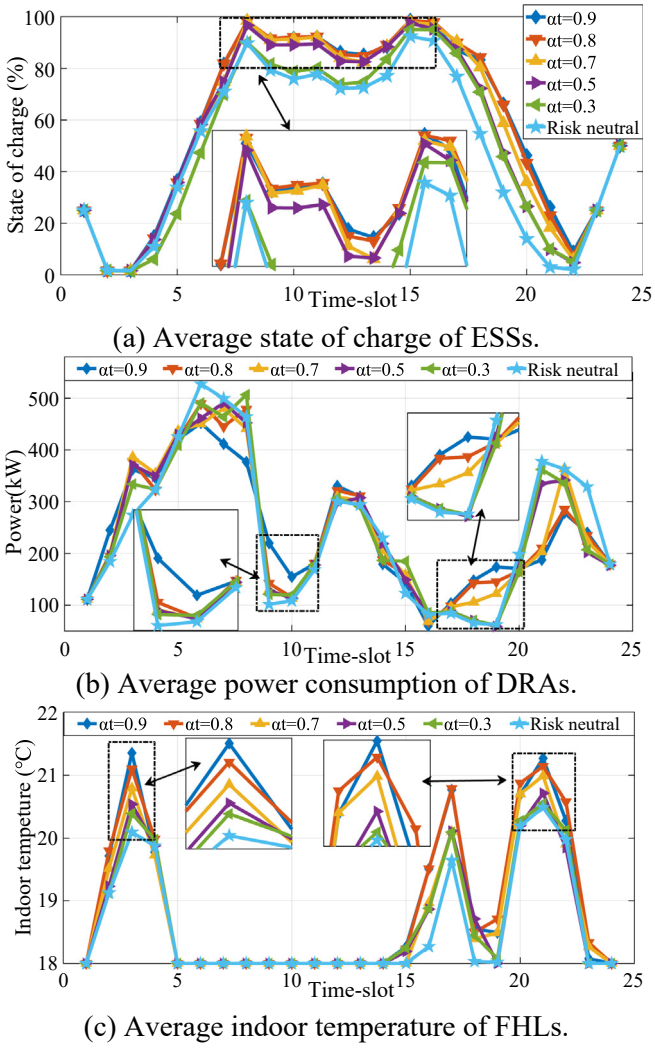


Fig. 4. Impacts of different risk-aversion preferences on RADP.

accommodate the possible high RES output in critical scenarios, as the penalty of RES curtailment becomes a major risk with the increasing β_{RES}. Table 2 compares the average curtailment of RESs and demand shedding of DRAs under different β_{RES} and α_t ($\lambda_t = 0.4$). When RES curtailment is not a major risk, a more risk-averse RADP incurs more RES curtailment because less capacities of ESSs and DRAs are reserved. As β_{RES} grows larger, RES curtailment reduces significantly while demand shedding of DRAs slightly increases, which is because demand shedding is more likely to occurs when DRAs have to reserve enough capacity to absorb extra RES power but RES output could suddenly decrease in critical scenarios.

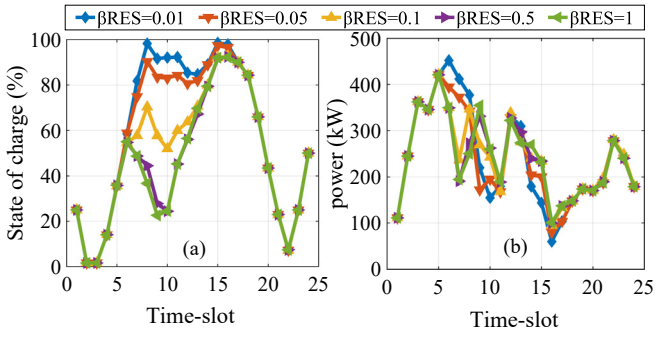


Fig. 5. Average state of charge of ESSs and average power consumption of DRAs under different β_{RES} ($\alpha_t = 0.8$, $\lambda_t = 0.4$).

Table 2
Average demand shedding (kWh) and RES curtailment (kWh) under different penalty coefficients.

	$\beta_{\text{RES}} = 0.01$	$\beta_{\text{RES}} = 0.05$	$\beta_{\text{RES}} = 0.1$	$\beta_{\text{RES}} = 0.5$	$\beta_{\text{RES}} = 1$
$\alpha_t = 0.3$	9.20 (26.72)	9.93 (25.32)	10.86 (18.11)	11.34 (17.72)	11.84 (11.14)
$\alpha_t = 0.5$	8.36 (30.80)	9.91 (25.48)	10.36 (18.93)	11.26 (15.42)	12.17 (7.23)
$\alpha_t = 0.8$	8.22 (33.81)	9.84 (27.17)	10.12 (21.40)	11.88 (13.56)	13.38 (4.76)

*values in brackets are RES curtailment (kWh).

5.5. Comparison results with different policies

The results of four comparative policies, namely myopic policy, off-line policy, deterministic MPC (DMPC), and risk-averse MPC (RMPC) [22], are tested for comparison. Specifically, myopic policy ignores the impact of the current decisions on future. Off-line policy obtains the optimal dispatch at the beginning of time horizon using the expected exogenous information. The formulations of DMPC and RMPC are shown in Appendix C. RADP and RMPC have the same risk-aversion preferences, i.e., $\alpha_t = 0.8$ and $\lambda_t = 0.4$.

Table 3 lists the results of different policies in 1000 random samples. It can be seen that RADP and RMPC obtains similar results. RADP provides great robustness to multivariate uncertainties, it facilitates economy of IEHS as well as reducing high quantile cost in critical scenarios. RADP achieves 17.2%, 2.7%, and 0.7% improvements in mean cost when compared with myopic policy, off-line policy, and DMPC. It also leads to 21.6%, 15.1%, and 11.7% reductions in 95%-VaR of costs. Myopic policy ignores the impacts of current decision on future, which results in high cost as well as high risk. The results of off-line policy are also unsatisfactory since the decisions are not adapted to the realized system state. DMPC policy incurs high cost variability because although it uses forecasting information to achieve real-time dispatch, risks caused by uncertainties are not considered.

Table 1
Average cost and 95%-VaR of total cost under different risk-aversion preferences ($).

	Risk neutral	$\alpha_t = 0.3$	$\alpha_t = 0.5$	$\alpha_t = 0.7$	$\alpha_t = 0.8$	$\alpha_t = 0.9$
$\lambda_t = 0.4$	2715.16 (3412.41)	2725.81 (3308.42)	2782.18 (3258.84)	2824.45 (3198.45)	2896.16 (3095.54)	2944.87 (3061.17)
$\lambda_t = 0.6$	2715.16 (3412.41)	2738.99 (3305.36)	2798.63 (3237.85)	2842.58 (3197.61)	2916.78 (3061.14)	2986.47 (3051.64)
$\lambda_t = 0.8$	2715.16 (3412.41)	2798.46 (3299.94)	2836.64 (3232.44)	2874.47 (3208.34)	2948.69 (3081.15)	3024.94 (3099.77)

*values in brackets are 95%-VaR of total costs ($).

Table 3
Performance of different policies.

	RADP	Myopic policy	Off-line policy	DMPC	RMPC
Average cost($)	2896.16	3498.24	2975.58	2917.45	2812.32
95%-VaR cost($)	3095.58	3946.18	3645.12	3504.91	3185.06
RES curtailment(kWh)	33.81	16.84	59.68	45.17	41.58
Demand shedding (kWh)	8.22	82.16	45.06	18.39	11.62

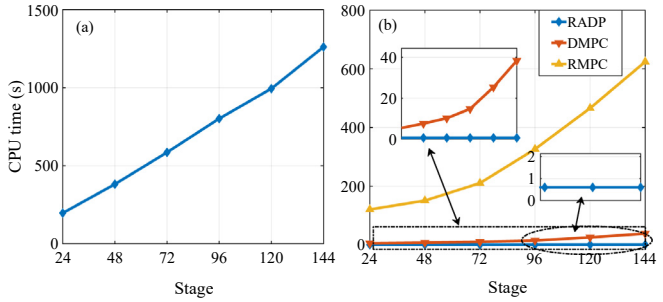


Fig. 6. Computation efficiency of RADP and other policies. (a). CPU time per iteration of RADP in off-line learning phase. (b) CPU time of different policies in on-line application phase.

5.6. Computation efficiency and scalability of RADP

This section investigates the computation efficiency and scalability of RADP. Fig. 6(a) demonstrates that the CPU time per iteration of RADP in off-line learning grows linearly as the number of dispatch stage grows. Fig. 6(b) shows that RADP trades off-line learning for superior on-line computation performance. Although RADP and RMPC obtains similar results (see Table 3). However, the computational efficiency of RADP is significantly higher than RMPC. In on-line application phase, RADP obtains the dispatch decision in only about 0.6s, which is dramatically lower than DMPC and RMPC. The computation time of DMPC and RMPC grows significantly as the problem size grows. When dispatch horizon contains 144 stages, it takes more than 35s for DMPC and 600s for RMPC to complete computation. In contrast, computation time of RADP is not affected by the number of stage and is only 2% and 0.1% of that of DMPC and RMPC, because RADP only solves a single stage sub problem at a time. These facts suggest that compared with MPC based approaches, RADP enjoys higher computation efficiency and better scalability when applied to more complex problems.

6. Conclusion

This paper investigates the real-time optimal stochastic dispatch of IEHS where the flexibilities of ESSs, DRAs, and FHLs are jointly utilized and multivariate uncertainties are considered. The original problem is formulated as a multistage risk-averse stochastic sequential optimization problem with dynamic risk measure. A data-driven ADP based framework is proposed to tackle such problem and a risk-averse ADP based policy which adapts to uncertainties is developed. Simulation results illustrate that RADP not only facilitates economy of IEHS but also reduces high quantile of total cost. It also helps in reducing RES curtailment and demand shedding. The risk-aversion preferences of RADP and penalty coefficients of should be properly chosen due to their obvious influences on the dispatch decisions. RADP trades off-line learning for remarkably high on-line computation efficiency and does not require forecasting information to achieve real-time dispatch, which promises an attractive application and scalability in other complex real-time stochastic dispatch problems in the future smart grids.

Future studies will investigate the optimal dispatch of IEHS with high penetration of RES [38] and more flexible resources, e.g., heat storages [39] and heat pumps [40], and a more practical HDN considering varying mass flow rates [41]. Besides, the optimal dispatch of integrated multi-energy system [42] with multiple energy hub will be further studied and a distributed ADP based on multi-agent learning and edge computing will be developed.

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.

CRediT authorship contribution statement

Zhenning Pan: Conceptualization, Methodology, Software, Writing - original draft. **Tao Yu:** Conceptualization, Resources, Supervision, Funding acquisition. **Jie Li:** Formal analysis. **Kaiping Qu:** Investigation. **Bo Yang:** Writing - review & editing.

Acknowledgments

The authors gratefully acknowledge the support of National Natural Science Foundation of China (51777078, 61963020), the Fundamental Research Funds for the Central Universities (D2172920), and the Key Projects of Basic Research and Applied Basic Research in Universities of Guangdong Province (2018KZDXM001).

Appendix A

The approximation method in Ref. [31] is adopted to approximate Eq. (2). Its main idea is to use a set of linear constraints to approximate a convex function. By choosing two sets of operational points of CHP, Eq. (2) is approximated by

$$C_n^{\mathrm{CHP}}(t) \geq \frac{\partial C_n^{\mathrm{CHP}}(t)}{\partial P_{n,i}^{\mathrm{CHP}}}\left(P_n^{\mathrm{CHP}}(t) - P_{n,i}^{\mathrm{CHP}}\right) + \frac{\partial C_n^{\mathrm{CHP}}(t)}{\partial H_{n,j}^{\mathrm{CHP}}}\left(H_n^{\mathrm{CHP}}(t) - H_{n,j}^{\mathrm{CHP}}\right) + C_{n,i,j}^{\mathrm{CHP}}, \forall i \in I_n, \forall j \in J_n, \tag{A1}$$

where $P_{n,i}^{\mathrm{CHP}}$, $H_{n,j}^{\mathrm{CHP}}$, and $C_{n,i,j}^{\mathrm{CHP}}$ are the electric power, heat power, and operation cost of a CHP at specify operational point, respectively; I_n and J_n represent the sets of operational points, gives

$$\frac{\partial C_n^{\mathrm{CHP}}(t)}{\partial P_{n,i}^{\mathrm{CHP}}} = 2a_{n,5}P_{n,i}^{\mathrm{CHP}} + a_{n,3}H_{n,j}^{\mathrm{CHP}} + a_{n,2}, \tag{A2}$$

$$\frac{\partial C_n^{\mathrm{CHP}}(t)}{\partial H_{n,j}^{\mathrm{CHP}}} = 2a_{n,4}H_{n,j}^{\mathrm{CHP}} + a_{n,3}P_{n,i}^{\mathrm{CHP}} + a_{n,1}, \tag{A3}$$

$$C_{n,i,j}^{\mathrm{CHP}} = a_{n,5}\left(P_{n,i}^{\mathrm{CHP}}\right)^2 + a_{n,4}\left(H_{n,j}^{\mathrm{CHP}}\right)^2 + a_{n,3}P_{n,i}^{\mathrm{CHP}}H_{n,j}^{\mathrm{CHP}} + a_{n,2}P_{n,i}^{\mathrm{CHP}} + a_{n,1}H_{n,j}^{\mathrm{CHP}} + a_{n,0}. \tag{A4}$$

Appendix B

The parameters of HDN, CHP, EDN, ambient temperature, and STATCOM are from Ref. [8]. The expectation of loads, RES output, and real-time price are listed in Table B1. The parameters of ESSs and FHLs are listed in Table B2 to Table B3. The data of DRAs are from Ref. [25]. Finally, the forecasting errors of inflexible electrical and heat loads, real-time price, and RES output are assumed to follow the normal distribution $N(0,0.03^2)$, $N(0,0.1^2)$, and $N(0,0.2^2)$.

Table B1
Expectation of inflexible electrical load, heat load, RES output, and real-time price.

Time-slot	1	2	3	4	5	6
$P_i^{\mathrm{EL}}(t)$(kW)	480.2	420.1	450.2	510.2	570.2	600.2
$Q_i^{\mathrm{EL}}(t)$(kVAR)	240.1	210.0	225.1	215.1	285.1	300.1
$H_n^{\mathrm{IHL}}(t)$(kW)	138.6	166.3	184.8	194.1	184.9	172.9
$\overline{P}_n^{\mathrm{RES}}(t)$(kW)	0	0	0	0	0	12
$\varepsilon(t)$($)	0.124	0.107	0.094	0.084	0.083	0.093
Time-slot	7	8	9	10	11	12
$P_i^{\mathrm{EL}}(t)$(kW)	660.3	690.3	720.3	780.2	840.3	870.4
$Q_i^{\mathrm{EL}}(t)$(kVAR)	330.2	345.2	360.2	390.1	420.2	435.2
$H_n^{\mathrm{IHL}}(t)$(kW)	162.7	152.5	152.5	151.6	148.8	147.9
$\overline{P}_n^{\mathrm{RES}}(t)$(kW)	30	78	108	150	180	192
$\varepsilon(t)$($)	0.084	0.089	0.113	0.140	0.159	0.162
Time-slot	13	14	15	16	17	18
$P_i^{\mathrm{EL}}(t)$(kW)	876.4	840.3	780.3	757.5	753.9	744.3
$Q_i^{\mathrm{EL}}(t)$(kVAR)	438.2	420.2	390.2	378.8	377.0	372.2
$H_n^{\mathrm{IHL}}(t)$(kW)	138.6	120.1	74.0	42.2	27.7	18.4
$\overline{P}_n^{\mathrm{RES}}(t)$(kW)	210	192	180	108	72	12
$\varepsilon(t)$($)	0.172	0.170	0.162	0.160	0.157	0.160
Time-slot	19	20	21	22	23	24
$P_i^{\mathrm{EL}}(t)$(kW)	739.5	737.1	732.3	720.3	660.3	540.2
$Q_i^{\mathrm{EL}}(t)$(kVAR)	370.0	368.6	366.1	360.2	330.2	270.1
$H_n^{\mathrm{IHL}}(t)$(kW)	21.3	31.4	41.6	70.2	83.2	110.9
$\overline{P}_n^{\mathrm{RES}}(t)$(kW)	6	0	0	0	0	0
$\varepsilon(t)$($)	0.167	0.169	0.169	0.161	0.150	0.139

Table B2
Parameters of ESS.

$\overline{P}_{c,n}^{\mathrm{ESS}}$ 100 kW	$\overline{E}_n^{\mathrm{ESS}}$ 400 kWh	$\eta_{\mathrm{ESS},c}$ 1	β_{ESS} $0.001/kW
$\overline{P}_{d,n}^{\mathrm{ESS}}$ 100 kW	$\underline{E}_n^{\mathrm{ESS}}$ 10 kWh	$\eta_{\mathrm{ESS},d}$ 1	

Table B3
Parameters of FHL

C_{air} 1.85 kWh/°C	$\overline{\tau}_n^{\mathrm{FHL}}$ 22 °C
r_n^{FHL} 1.3 °C/kWh	$\underline{\tau}_n^{\mathrm{FHL}}$ 18 °C

Appendix C

By using the expected exogenous information, the optimal dispatch of DMPC at time-slot t can be obtained by solving (C1), as

$$X_t^{\mathrm{DMPC}} = \mathrm{argmin}\left(C_t(S_t,X_t) + \sum_{k=t+1}^{|T|} C_k(S_k,X_k)\right), \tag{C1}$$

subject to (1)–(15), (16a)-(16e), (17)–(21), and (31).

The optimal dispatch of RMPC at time-slot t can be obtained by solving (C2), as

$$X_t^{\mathrm{RMPC}} = \mathrm{argmin}\left(C_t(S_t,X_t) + \frac{1-\lambda_{t+1}}{|N_s|}\left(\sum_{s=1}^{|N_s|}\sum_{k=t+1}^{|T|} C_k^s(S_k^s,X_k^s)\right) + \lambda_{t+1}\left(u + \frac{1}{\alpha_{t+1}|N_s|}\sum_{s=1}^{|N_s|}\varpi(s)\right)\right), \tag{C2}$$

subject to (1)–(15), (16a)-(16e), (17)–(21), (31), and (C3)

$$\begin{cases} \varpi(s) \geq 0, \forall s, \\ u - \sum_{k=t+1}^{|T|} C_k^s(S_k^s,X_k^s) + \varpi(s) \geq 0, \forall s, \end{cases} \tag{C3}$$

where ϖ is the auxiliary variable; N_s denotes the set of forecasting scenarios which contains 100 scenarios; s is scenario subscript; and X_t^{RMPC} represents the decision variables, yields

$$X_t^{\mathrm{RMPC}} = \left[X_t, X_{t+1}^s, \ldots, X_{|T|}^s\right], \forall s. \tag{C4}$$

Note that although solution of MPC involves dispatch decisions from current stage to the end of dispatch horizon, only the current stage decision will be implemented.

References

[1] Shankar K, Manohar K, Michael DM, Hossein S. Efficiency improvements through combined heat and power for on-site distributed generation technologies. Cogener. Distrib. J. 2007;22(3):19–34.
[2] East M, Africa N. World energy outlook 2014. International Energy Agency; 2014. p. 2010.
[3] Yang B, Wang JB, Zhang XS, Yu T, Yao W, Shu HC, Zeng F, Sun LM. Comprehensive overview of meta-heuristic algorithm applications on PV cell parameter identification. Energy Convers Manag 2020;208:112595.
[4] Li Z, Wu W, Shahidehpour M, Wang J, Zhang B. Combined heat and power dispatch considering pipeline energy storage of district heating network. IEEE Trans. Sustain. Energy 2016;7(1):12–22.
[5] Jiang XS, Jing ZX, Li YZ, Wu QH, Tang WH. Modelling and operation optimization of an integrated energy based direct district water-heating system. Energy 2014;64:375–88.
[6] Liu X, Mancarella P. Modelling, assessment and Sankey diagrams of integrated electricity-heat-gas networks in multi-vector district energy systems. Appl Energy 2016;167:336–52.
[7] Qin X, Sun H, Shen X, Guo Y, Guo Q, Xia T. A generalized quasi-dynamic model for electric-heat coupling integrated energy system with distributed energy resources. Appl Energy 2019;251:113270.
[8] Liu B, Meng K, Dong ZY, Wei W. Optimal dispatch of coupled electricity and heat system with independent thermal energy storage. IEEE Trans Power Syst 2019;34(4):3250–63.
[9] Li R, Wei W, Mei S, Hu Q, Wu Q. Participation of an energy hub in electricity and heat distribution markets: an MPEC approach. IEEE Trans. Smart Grid 2019;10(4):3641–53.
[10] Wang D, Hu Q, Jia H, Hou K, Du W, Chen N, Wang X, Fan M. Integrated demand response in district electricity-heating network considering double auction retail energy market based on demand-side energy stations. Appl Energy 2019;248:656–78.
[11] Li X, Li W, Zhang R, Jiang T, Chen H, Li G. Collaborative scheduling and flexibility assessment of integrated electricity and district heating systems utilizing thermal inertia of district heating network and aggregated buildings.

Appl Energy 2020;258:114021.
[12] Xue Y, Li Z, Lin C, Guo Q, Sun H. Coordinated dispatch of integrated electric and district heating systems using heterogeneous decomposition. IEEE Trans. Sustain. Energy 2020. https://doi.org/10.1109/TSTE.2019.2929183. In press.
[13] Huang J, Li Z, Wu QH. Coordinated dispatch of electric power and district heating networks: a decentralized solution using optimality condition decomposition. Appl Energy 2017;206:1508–22.
[14] Pan Z, Guo Q, Sun H. Feasible region method based integrated heat and electricity dispatch considering building thermal inertia. Appl Energy 2017;192:395–407.
[15] Yang B, Yu T, Shu HC, Dong J, Jiang L. Robust sliding-mode control of wind energy conversion systems for optimal power extraction via nonlinear perturbation observers. Appl Energy 2018;210:711–23.
[16] Xu D, Zhou B, Chan KW, Li C, Wu Q, Chen B, Xia S. Distributed multi-energy coordination of multi-microgrids with biogas-solar-wind renewables. IEEE Trans. Ind. Informat 2019;15(6):3254–66.
[17] 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.
[18] Tan J, Wu Q, Hu Q, Wei W, Liu F. Adaptive robust energy and reserve co-optimization of integrated electricity and heating system considering wind uncertainty. Appl Energy 2020;260:114230.
[19] Moazeni S, Miragha AH, Defourny B. A risk-averse stochastic dynamic programming approach to energy hub optimal dispatch. IEEE Trans Power Syst 2019;34(3):2169–78.
[20] Bellman RE. Dynamic programming. Princeton, NJ, USA: Princeton Univ. Press; 1957.
[21] Gu W, Wang Z, Wu Z, Luo Z, Tang Y, Wang J. An online optimal dispatch schedule for CCHP microgrids based model predictive control. IEEE Trans. Smart Grid 2017;8(5):2332–42.
[22] Ma L, Liu N, Zhang J, Wang L. Real-time rolling horizon energy management for the energy-hub-coordinated prosumer community from a coorperative perspective. IEEE Trans. Power Syst 2019;34:1227–42.
[23] Nascimento J, Powell WB. An optimal approximate dynamic programming algorithm for concave, scalar storage problems with vector-valued controls. IEEE Trans Automat Contr 2013;58(12):2995–3010.
[24] Shuai H, Fang J, Ai X, Tang Y, Wen J, He H. Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming. IEEE Trans. Smart Grid 2019;10(3):2440–52.
[25] Pan Z, Yu T, Chen L, Yang B, Wang B, Guo W. Real-time stochastic optimal scheduling of large-scale electric vehicles: a multidimensional approximate dynamic programming approach. Int J Electr Power Energy Syst 2020;216:105542. https://doi.org/10.1016/j.ijepes.2019.105542.
[26] Philpott A, de matos V. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. Eur J Operation Res 2012;218:470–83.
[27] Shapiro A, Tekaya W, da Costa JP, Soares MP. Risk neutral and risk averse stochastic dual dynamic programming method. Eur J Operation Res 2013;224:375–91.
[28] Ruszczyński A. Risk-averse dynamic programming for Markov decision processes. Math Program 2010;125:235–61.
[29] Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk. J Risk 2000;2(3):21–41.
[30] Yang B, Zhong LE, Yu T, Li HF, Zhang XS, Shu HC, Sang YY, Jiang L. Novel bio-inspired memetic salp swarm algorithm and application to MPPT for PV systems considering partial shading condition. J Clean Prod 2019;215:1203–22.
[31] Frangioni A, Gentile C, Lacalandra F. Tighter approximated milp formulations of unit commitment problems. IEEE Trans Power Syst 2019;24(1):105–13.
[32] Xu Z, Hu Z, Song Y, Wang J. Risk-averse optimal bidding strategy for demand-side resource aggregators in day-ahead electricity markets under uncertainty. IEEE Trans. Smart Grid 2017;8(1):96–105.
[33] Yang B, Jiang L, Wang L, Yao W, Wu QH. Nonlinear maximum power point tracking control and modal analysis of DFIG based wind turbine. Int J Electr Power Energy Syst 2016;(74):429–36.
[34] Yang B, Yu T, Zhang XS, Li HF, Shu HC, Sang YY, Jiang L. Dynamic leader based collective intelligence for maximum power point tracking of PV systems affected by partial shading condition. Energy Convers Manag 2019;179:286–303.
[35] Yang P, Nehorai A. Joint optimization of hybrid energy storage and generation capacity with renewable energy. IEEE Trans. Smart Grid 2014;5(4):1566–74.
[36] Yuan H, Li F, Wei Y, Zhu J. Novel linearized power flow and linearized OPF model for active distribution networks with application in distribution LMP. IEEE Trans. Smart Grid 2018;9(1):438–48.
[37] Powell WB, Meisel S. Tutorial on stochastic optimization in energy—part I: modeling and policies. IEEE Trans Power Syst 2016;31(2):1459–67.
[38] Yang B, Zhang XS, Yu T, Shu HC, Fang ZH. Grouped grey wolf optimizer for maximum power point tracking of doubly-fed induction generator based wind turbine. Energy Convers Manag 2017;133:427–43.
[39] Hast A, Rinne S, Syri S, Kiviluoma J. The role of heat storages in facilitating the adaptation of district heating systems to large amount of variable renewable electricity. Energy 2017;137:775–88.
[40] Urbanucci L, Bruno JC, Testi D. Thermodynamic and economic analysis of the integration of high-temperature heat pumps in trigeneration systems. Appl Energy 2019;238:516–33.
[41] Yang B, Yu T, Shu HC, Zhang YM, Chen J, Sang YY, Jiang L. Passivity-based sliding-mode control design for optimal power extraction of a PMSG based variable speed wind turbine. Renew Energy 2018;119:577–89.
[42] Huang S, Tang W, Wu Q, Li C. Network constrained economic dispatch of integrated heat and electricity systems through mixed integer conic programming. Energy 2019;179:464–74.