Multi-Stage Real-Time Operation of a Multi-Energy Microgrid With Electrical and Thermal Energy Storage Assets: A Data-Driven MPC-ADP Approach

Zhengmao Li, *Member, IEEE*, Lei Wu, *Senior Member, IEEE*, Yan Xu, *Senior Member, IEEE*, Somayeh Moazeni, *Senior Member, IEEE*, and Zao Tang, *Member, IEEE*

***Abstract*—This paper studies the multi-stage real-time stochastic operation of grid-tied multi-energy microgrids (MEMGs) via the hybrid model predictive control (MPC) and approximate dynamic programming (ADP) approach. In the MEMG, practical power and thermal network constraints, heterogeneous energy storage devices, and distributed generations are involved. Given the relatively large thermal inertia and slow thermal energy fluctuation, only uncertainties of renewable energy sources and active/reactive power loads are considered. Then, historical data are adopted as training scenarios for the MPC-ADP method to acquire empirical knowledge for dealing with all the diverse uncertainties. Further, piecewise linear functions are used to approximate value functions with respect to the operation status of energy storage assets, which enables sequentially solving the Bellman's equation forward through time to minimize MEMG operation cost. Finally, numerical case studies are conducted to illustrate the effectiveness and superiority of the proposed MPC-ADP approach. Simulation results indicate that with sufficient information embedded, the MPC-ADP approach could obtain good-enough real-time operation solutions with the successively updated forecast. Further, it outperforms alternative real-time operation benchmarks in terms of optimality and convergence for various application scenarios.**

***Index Terms*—Hybrid model predictive control-approximate dynamic programming, multi-energy microgrid, stochastic operation, heterogeneous energy storage.**

NOMENCLATURE

Sets and Indices

i, j, n	Indices of electrical buses or thermal nodes
N_{SN}	Set of source nodes in the thermal network
p/b	Index of thermal pipes/electricity branches
$pair(p,n)$	Indicating pipe p is connected with node/bus n
$P^n_{s/r,(+)}$	Set of thermal pipes starting at node n in the supply/return network
$P^n_{s/r,(-)}$	Set of thermal pipes ending at node n in the supply/return network
t	Index of real-time operation intervals

Parameters

C_W	Water heat capacity
$E^{i,min}_{BS/TS}/E^{i,max}_{BS/TS}$	Min/max energy of the battery storage (BS)/ thermal storage (TS) on bus i
$H^{i,max}_{TC}/H^{i,max}_{TD}$	Max absorbing/releasing thermal energy of the TS on bus i
$H^{p,t}_{TL}$	Thermal demand for pipe p at interval t
N_I	Number of electricity buses
N_T	Number of real-time operation intervals
$P^{i,max}_{BC}/P^{i,max}_{BD}$	Max charging/discharging power of the BS on bus i
$P^{i,max}_{PtC}$	Maximum power input of the power-to-thermal conversion (PtC) unit on bus i
$P^{i,min}_{MT}/P^{i,max}_{MT}$	Min/max power output of the combined cooling, heat and power (CCHP) plant on bus i
$P^{t,i}_{EL}/Q^{t,i}_{RL}$	Active/reactive power demand of bus i at interval t
$P^{t,i}_{WT}/P^{t,i}_{PV}$	Active power output of the wind turbine (WT)/ photovoltaic (PV) cell on bus i at interval t
r_b/x_b	Resistance/reactance of branch b
$T^{min}_{s/r,(+)}/T^{max}_{s/r,(+)}$	Min/max temperature at the start of a supply/ return pipe
$T^{min}_{s/r,(-)}/T^{max}_{s/r,(-)}$	Min/max temperature at the end of a supply/ return pipe
$V^{i,min}_{BUS}/V^{i,max}_{BUS}$	Min/max nodal voltage magnitude of bus i
V_s	Voltage of the substation bus
Δt	Unit dispatch interval
η_{ME}	Power efficiency of a CCHP
η_{PtC}/η_{MT}	Power-to-thermal conversion efficiency of a PtC/ CCHP

Manuscript received April 23, 2021; revised August 1, 2021 and September 18, 2021; accepted October 9, 2021. Date of publication October 13, 2021; date of current version December 23, 2021. This work was supported in part by the U.S. National Science Foundation under Grant CNS-1915756. Paper no. TSG-00640-2021. *(Corresponding author: Lei Wu.)*

Zhengmao Li, Lei Wu, and Zao Tang are with the ECE Department, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: zli161@stevens.edu; lei.wu@stevens.edu; ztang17@stevens.edu).

Yan Xu is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

Somayeh Moazeni is with the School of Business, Stevens Institute of Technology, Hoboken, NJ 07030 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2021.3119972.

Digital Object Identifier 10.1109/TSG.2021.3119972

η_{BC}/η_{BD}	Charging/discharging efficiency of a BS
η_{TC}/η_{TD}	Absorbing/releasing efficiency of a TS
$\xi^{M}_{PtC}/\xi^{M}_{TS}/\xi^{M}_{MT}$	Unit maintenance cost of a PtC/TS/CCHP
$\xi^{M}_{WT}/\xi^{M}_{PV}/\xi^{M}_{BS}$	Unit maintenance cost of a WT/PV/BS
ξ_{NG}	Natural gas fuel price
τ_{BS}/τ_{TS}	The decay rate of BS/TS
τ_b/τ_s	Electrical energy purchasing/selling price

Variables

$E^{t,i}_{BS}/E^{t,i}_{TS}$	Energy stored in BS/TS on bus/node i at interval t
F^{MADP}	Optimal objective value obtained by the proposed MPC-ADP approach (from offline training or real-time operation layer)
F^{t}_{FU}/F^{t}_{OM}	Fuel/maintenance cost at interval t
F^{t}	Total MEMG operation cost at interval t
F^{t}_{EX}	Power transaction cost of MEMG at interval t
$H^{p,t}_{s/r,(+)}$	Thermal energy at the start of supply/return pipe p at interval t
$H^{p,t}_{s/r,(-)}$	Thermal energy at the end of supply/return pipe p at interval t
$H^{t,i}_{PtC}/H^{t,i}_{MT}$	Thermal output of the PtC/CCHP on bus i at interval t
$H^{t,i}_{TC}/H^{t,i}_{TD}$	Absorbing/releasing thermal energy of the TS on bus i at interval t
$m^{p,t}_{s/r}/T^{n,t}_{s/r}$	Mass flow rate/temperature of pipe p at interval t
P^{t}_{BUY}/P^{t}_{SELL}	Purchasing/selling power at interval t
$P^{t,0,b}_{AC}/Q^{t,0,b}_{RE}$	Active/reactive power on a lateral branch of b at interval t
$P^{t,b}_{AC}/Q^{t,b}_{RE}$	Active/reactive power flow of branch b at interval t
$P^{t,i}_{BC}/P^{t,i}_{BD}$	Charging/discharging power of the BS on bus i at interval t
$P^{t,i}_{PtC}$	PtC power consumption on bus i at interval t
$P^{t,i}_{MT}$	Active power output of the CCHP on bus i at interval t
$T^{p,t}_{s/r,(+)}$	Temperature at the start of supply/return pipe p at interval t
$T^{p,t}_{s/r,(-)}$	Temperature at the end of supply/return pipe p at interval t
$V^{t,i}_{BUS}$	Voltage magnitude of bus i at interval t

I. Introduction

MULTI-ENERGY microgrids (MEMGs) are the emerging paradigms around the world, whereby the heterogeneous energy carriers such as electrical and thermal (e.g., heat and cooling) energy are generated, transmitted, and consumed on the distribution network level. MEMGs have proven to be an effective way in providing cost-effective and reliable multi-energy supply via enhanced renewable energy source (RES) utilization as well as multi-energy coordination. Under this circumstance, one primary focus of the industry and academia is to explore the practical real-time MEMG operation methods against diverse uncertainties [1].

In a MEMG, as consumers typically demand a large amount of electrical and thermal energy simultaneously [2], extensive research works have been conducted on the optimal electrical-thermal energy coordination. Reference [3] presents an optimal coordinated operation model for a MEMG with RESs to minimize electrical and thermal energy supply cost in both the grid-tied and islanded mode. A multi-objective electrical-thermal energy coordinated dispatch model is proposed in [4] for a grid-tied MEMG while considering the demand response schemes. Reference [5] presents a detailed MEMG model to minimize the multi-energy supplying cost, subject to specified physical operation constraints. Although the effective multi-energy coordination is achieved in [3]–[5], the heterogeneous uncertainties are not effectively tackled, posing great threats to the reliable and economic MEMG operation.

To tackle various uncertainties gradually revealed in the MEMG operation, the robust optimization (RO) and stochastic programming (SP) methods [6] have been applied. A RO-based approach is proposed in [7] to minimize the MEMG operation cost in multiple timescales. In [8], to derive the day-ahead economic MEMG operation decisions while attenuating adverse effects from all various uncertainties, a RO-based dispatch method is discussed. A two-stage RO method is utilized in [9] to optimally coordinate multiple MEMG assets, including distributed generators, energy storage devices, and demand response assets. The RO methods in [7]–[9] contribute to immunizing the MEMG operation against all various uncertainties, however, their results can be too conservative as the occurrence possibility of the worst case is extremely low, hindering its economic efficiency in practice [6].

To mitigate solution conservativeness, the SP method can be applied by incorporating a set of scenarios sampled from historical data. In [10], diverse uncertainties from electricity transaction prices, electricity loads, and RES outputs are depicted by scenario trees and managed by a two-stage SP scheme to minimize the expected MEMG operation cost. The studies in [11] and [12] introduce an SP operation method for MEMGs with diverse uncertainties, aiming to coordinate the multiple energy in different timescales and reduce the holistic energy supply costs. Reference [13] proposes a two-timescale coordinated SP operation approach for the RES-integrated MEMG, immunizing against the diverse uncertainties with minimum operation cost. The SP methods employed in [10]–[13] can alleviate the adverse effects of diverse uncertainties on MEMG operations but they require the specific probability distributions of individual uncertainties which could be hard to obtain. Furthermore, the fact that only a limited number of scenarios is preserved after the scenario reduction in [10]–[13] could result in a coarse coverage of the uncertainty space. On the other hand, a large number of scenarios is required to ensure the solution quality, which however could be too computationally demanding [6].

The coordinated electrical and thermal energy operation researches in [3]–[13] assume that the thermal energy is consumed locally, deviating from the practical condition that thermal loads are geographically distributed. Indeed, the practical thermal flow constraints need to be enforced, while

considering that the thermal inertia could provide additional dispatch flexibilities [14]. In this light, reference [15] proposes a two-layer (i.e., day-ahead and real-time) risk-averse SP approach for the optimal coordination of multi-energy carries in the MEMG against diverse uncertainties. The practical thermal flow modeling is incorporated to enhance the overall dispatch flexibility. However, the SP method could be overly time-consuming. In [16], a two-stage RO-based dispatch model for a MEMG is presented to make optimal operation plans against power load uncertainties. The thermal dynamic characteristics of the district thermal network are utilized to provide operational flexibility and improve economic performance. To handle multiple uncertainties, reference [17] proposes a two-stage RO economic dispatch model with a robust thermal comfort management strategy and the practical thermal flow constraints. Though computationally effective, the solutions derived from [16] and [17] could be too conservative as [7]–[9].

From the above literature review, it can be seen that references [7]–[13] and [15]–[17] adopt a two-step framework, including a RO or SP-based day-ahead model and a real-time model, to approach the real-time MEMG operation with uncertainties. That is, the RO and SP-based two-stage day-ahead models are used to first obtain operation plans for the entire day, and the real-time model determines flexible adjustments over the day-ahead solution for each time slot in the real-time scale. As a matter of fact, the real-time MEMG operation is a multi-stage optimization problem, in which uncertainties are gradually revealed over time and the sequence of real-time operation decisions is made adaptively at each time slot throughout the day [18]. That is, non-anticipativity constraints are necessary for the real-time MEMG operation to ensure that the decisions taken at interval t depend only on the information revealed up to interval t, but not on the data that will be realized in the future. In this sense, the two-stage RO and SP methods fail to satisfy the requirements of the real-time non-anticipative operation, which will severely under-estimate the effect of uncertainties and lead to suboptimal or even infeasible operations.

As for the current research work on handling the non-anticipativity of uncertainty factors, reference [18] proposes two mixed-integer linear programming (MILP) methods, namely explicit and implicit decision methods, for solving the scheduling problem of thermal units and energy storage assets in a transmission network while ensuring solution robustness and non-anticipativity. However, the empirical knowledge from historical data is not fully utilized in those methods, and their computational performance is sensitive to the problem dimension. In [19], a stochastic dual dynamic programming approach while considering multi-stage uncertainties is developed to get the optimal operation decisions of power systems, and the non-anticipative requirements are satisfied through the multi-dimensional Benders cuts. However, the approach in [19] uses a sample approximation, which may produce certain errors; in addition, only a fixed set of samples is generated and used for all iterations, while the new information or samples are hard to be involved [20]. Finally, same as [18], reference [19] focuses on the electricity dispatch and does not address the additional flexibility from the multi-energy coordination. Reference [21] proposes a two-stage multi-period distributionally robust energy management model for a MEMG. Although the non-anticipativity of uncertainties is considered in the dispatch process, the non-anticipativity constraints are formulated in the form of uncertainty budgets, which fail to make full use of historical data to get practical operation solutions.

Indeed, the approximate dynamic programming (ADP) method, a modeling and algorithmic framework to solve the multi-stage SP problem with non-anticipativity constraints as a Markov decision process (MDP) in a real-time manner [22], [23], becomes promising to overcome all the shortcomings of existing studies. Specifically, following the basic concepts of the dynamic programming (DP) method, the Bellman's equation is applied to decompose all temporal dependencies so that a multi-stage SP problem can be divided into several small sub-problems and solved iteratively [6]. In addition, it does not require the probability distribution or uncertainty sets. Instead, it engages active learning through the successive interpretations of an increasing volume of historical data (i.e., data-driven) [24].

The ADP method has drawn broad research interest recently and demonstrates the effectiveness to solve SP problems in many areas such as energy storage operation [6], [24], energy policy and investment planning [18], [25], electricity market [23], etc. However, all these studies confine into the power system operation with one electrical energy storage asset and simple approximations, which may not be naturally applicable to the MEMG operation with multiple storage assets of different energy carriers. Further, the static ADP approach applied above cannot fully utilize all the real-time information that is constantly updated and successively revealed, which otherwise could further enhance solution quality and convergence performance. In this regard, the model predictive control (MPC), designed for online system operation [26] with constantly updated information, can be incorporated to make more practical decisions forward through time. It is especially suitable for handling energy storage assets and thermal energy systems whose operations are tightly time-coupled due to the state-of-charge evolution equations and large inertia. However, the MPC method alone, such as the research work [12]–[13], cannot guarantee globally optimal solutions [6].

To fill in the existing research gaps identified above, this paper discusses a data-driven MPC-ADP-based stochastic real-time operation method for the practical electricity and thermal network-based MEMG with multiple electrical and thermal energy storage assets. The proposed method makes full use of the available data to achieve uncertainty mitigation, multi-energy coordination, and cost-saving for the real-time MEMG operation.

The main contributions are summarized as follows.

1) The optimal coordinated MEMG operation with the comprehensive and practical power and thermal energy flows is formulated as a real-time multi-stage SP model.

2) A data-driven MPC-ADP-based approach for the MEMG operation is presented. It embeds the empirical knowledge obtained from the historical data for offline training so that

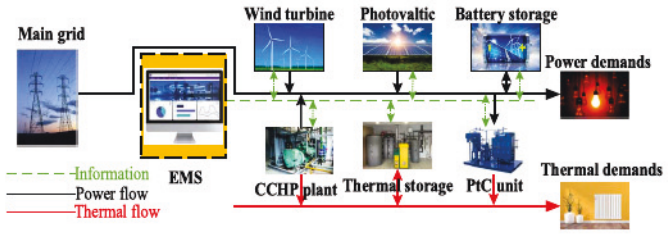

Fig. 1. An illustrative example of an MEMG.

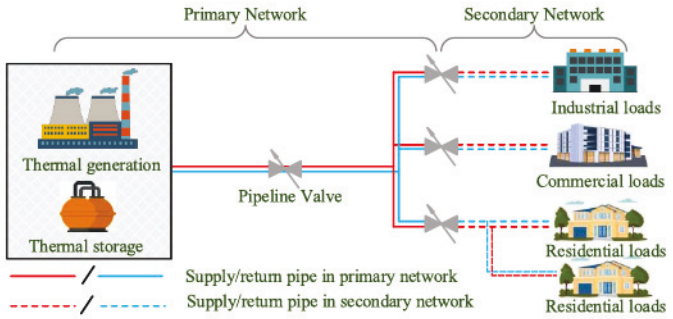

Fig. 2. A general structure of the thermal network.

the needs on the probability distribution of uncertainties by the traditional SP can be naturally waived. Furthermore, it fully utilizes constantly updated forecasts to derive good-enough real-time operation decisions and speeds up the convergence process.

3) The piecewise linear function (PLF) approximation approach for the MEMG with multiple energy storage assets is put forward, by converting the traditional convex solution space into multiple sets of PLFs. The simulation results illustrate its efficacy for the MEMG with multiple energy storage assets of different energy carriers.

The rest of this paper is organized as follows: The practical MEMG model is given in Section II; Section III presents the multi-stage real-time coordinated operation model of MEMGs; Section IV describes the data-driven MPC-ADP based approach; Case studies are conducted in Section V; The paper is concluded in Section V.

II. Practical Multi-Energy Microgrid Model

The typical structure of a grid-tied MEMG with energy storage assets and distributed generators is shown in Fig. 1. As for the distributed generators, wind turbines (WTs) and photovoltaic cells (PVs) are non-dispatchable generators with the high intermittency and fluctuation [12]; The combined cooling, heat, and power (CCHP) plants and power-to-thermal conversion (PtC) units as dispatchable assets deal with the electrical and thermal energy simultaneously and thus tighten their interactions. Specifically, the CCHP plants consume the natural gas fuel to generate electrical and thermal energy simultaneously [9], while the PtC units convert electricity into heat/cooling energy [3]. Heterogeneous energy storage assets, including battery storage (BS) and thermal storage (TS), can shift the electricity/thermal loads and enhance the MEMG operational flexibility [12]. The MEMG is also connected to the main grid for electrical energy exchange as appropriate.

A practical MEMG is built upon an integrated electricity-thermal distribution network. The electricity network models have been extensively studied in the literature, while practical thermal network models are rarely considered [9], [11]. The general structure of the thermal network is demonstrated in Fig. 2.

As shown in Fig. 2, a thermal network usually contains one source node and multiple intermediate/load nodes connected via supply and return pipelines. The generation-side network is called the primary network, and the consumption-side network is the secondary network [14]. The thermal energy is delivered via water through supply pipelines in the primary network; after entering the thermal exchanger in the secondary network, thermal energy is transmitted to all load nodes and water temperature dramatically changes [16], [17]. At last, the water flows back through return pipelines to finish the thermal circulation. Compared to the primary network, the pipelines in the secondary network are usually not modelled in detail [15] because they are short with only a few adjustable valves.

- *Nodal thermal flow balance:* The thermal flow follows the *Kirchhoff* law for the inflow-outflow energy balance (1).

$$\sum_{p\in \boldsymbol{P}^{n}_{s/r,(-)}} m^{p,t}_{s/r} = \sum_{p\in \boldsymbol{P}^{n}_{s/r,(+)}} m^{p,t}_{s/r} \tag{1}$$

- *Nodal temperature mixing:* According to *the first law of thermodynamics*, the nodal temperature depends on water mass and temperature as in (2). Besides, the temperature of a node is equal to its corresponding starting node of the same pipe (3).

$$\sum_{p\in \boldsymbol{P}^{n}_{s/r,(-)}} T^{p,t}_{s/r,(-)} \cdot m^{p,t}_{s/r} = T^{n,t}_{s/r} \cdot \left(\sum_{p\in \boldsymbol{P}^{n}_{s/r,(+)}} m^{p,t}_{s/r} \right) \tag{2}$$

$$T^{n,t}_{s/r} = T^{p,t}_{s/r,(+)}, pair(p,n) \tag{3}$$

- *Thermal energy calculation:* The thermal energy of each pipe can be quantified via (4).

$$H^{p,t}_{s/r,(-)/(+)} = C_W \cdot m^{p,t}_{s/r} \cdot T^{p,t}_{s/r,(-)/(+)} \tag{4}$$

- *Thermal transmission delay:* Different from the electric energy, the thermal energy transmission shows a non-negligible time delay [14], [15]. That is, due to the relatively slow flowing speed of water in the pipeline, the water temperature at the start and end of a pipe may not be the same. To quantify this transmission delay, detailed formulations are given in the Appendix-A.
- *Internal temperature range:* The inflow and outflow water temperatures of a pipe are bounded as in (5).

$$T^{\min}_{s/r,(-)/(+)} \le T^{p,t}_{s/r,(-)/(+)} \le T^{\max}_{s/r,(-)/(+)} \tag{5}$$

- *Thermal energy balance:* Thermal energy supply and demand shall be balanced at each time. The total thermal energy demand is calculated as the energy difference between the start of supply pipes and the end of return pipes (6).

$$H^{p,t}_{TL} = C_W \cdot \left[m^{p,t}_{s} \cdot T^{p,t}_{s/r,(+)} - m^{p,t}_{r} \cdot T^{p,t}_{s/r,(-)} \right], p \in \mathrm{N_{SN}} \tag{6}$$

III. Multi-Stage Real-Time Coordinated Operation of MEMGs

In this paper, the real-time MEMG operation is formulated as a multi-stage (i.e., N_T stages) SP problem (7)-(37), in which the sequence of real-time coordinated operation decisions is made at each interval t for $t = 1, \ldots, N_T$ with the gradually revealed uncertainties up to interval t. The multi-stage real-time MEMG operation is a non-anticipative process, i.e., the real-time operation decision at each interval t only depends on the information available up to interval t, but not on the future observations.

The multi-stage real-time operation model is described as in (7), which minimizes the expected supply cost throughout the day subjecting to all the prevailing operation constraints and diverse uncertainties. Note that in (7):

- $E[*]$denotes the expectation operator;
- Uncertainty parameters at interval t W^t include RES outputs and active/reactive power demands, i.e., $W^t = \{P^{t,i}_{WT}, P^{t,i}_{PV}, P^{t,i}_{EL}, Q^{t,i}_{RL}\}$. Given that thermal energy fluctuation is relatively slower than electrical energy due to large thermal inertia [12], we focus on uncertainties of RESs and active/reactive power demands;
- y^t includes all real-time decision variables at interval t, which only depends on information of uncertainties up to interval t, i.e., $y^t = y^t\ (W^{[1,t]})$ where $W^{[1,t]} = \{W^1, \ldots, W^t\}$;
- The feasible region of y^t is denoted as set Ψ^t, which is described by prevailing MEMG operation constraints.

The MEMG operation cost at interval t $F^t\ (y^t, W^t)$ is calculated as in (8), including the fuel cost (9), maintenance cost (10), and electrical energy transaction cost with the main grid (11).

Set $\Psi^t\ (y^t, W^t)$ describes prevailing operation constraints of the MEMG operation at interval t, which can be formulated as in (12)-(37). The practical thermal flow constraints are described as in (12). Constraint (13) presents the thermal energy balance. Constraints (14) and (15) describe the coupling relationship between electrical and thermal energy for CCHP plants and PtC units [14].

The *Linear Dist-flow* model (16)-(19) is used to formulate the electricity distribution network [11], where the active and reactive power balance equations are described in (16) and (17), and nodal voltage is calculated in (18) and limited via (19). The voltage security is guaranteed via the safe voltage range in (19) [14], [15], [27]. Constraint (20) denotes the active power output limits of the CCHP plant, and (21) is the active power input limits of the PtC unit [15], [20]. Constraint (22) shows the relationship of the power flow and the power transactions. Constraints (23)-(24) limit the power transaction between the MEMG and the main grid, and (25) is the exclusive electrical energy purchasing and selling constraint [9]. Constraints (26)-(31) are the operation limits of BS, where (26)-(27) are the safe range of its charging and discharging power; (28) denotes its energy balance; (29) is the safe operation range of energy stored in the BS [8]; (30) indicates that the BS cannot charge and discharge at the same time [6], [14]; and (31) restricts the same initial and terminal energy levels for achieving consistent dispatch flexibility in the daily operation cycle. Similarly, operation constraints of the TS are presented in (32)-(37), where (32)-(33) are the absorbing and releasing power limits; (34) is the energy balance in the TS; (35) describes the energy limits; (36) restricts that the absorbing and releasing cannot happen at the same time; and (37) ensures the same dispatch flexibility of the TS [15], [27].

$$F^t(y^t, W^t) = F^t_{FU}(y^t, W^t) + F^t_{OM}(y^t, W^t) + F^t_{EX}(y^t, W^t) \tag{8}$$

$$F^t_{FU}(y^t, W^t) = \left(\sum_{i \in N_I} \xi_{NG} \cdot P^{t,i}_{MT}/\eta_{ME}\right) \cdot \Delta t \tag{9}$$

$$F^t_{OM}(y^t, W^t) = \sum_{i \in N_I} \begin{bmatrix} \xi^M_{MT} \cdot P^{t,i}_{MT} + \xi^M_{BS} \cdot \left(P^{t,i}_{BC} + P^{t,i}_{BD}\right) + \\ \xi^M_{PV} \cdot P^{t,i}_{PV} + \xi^M_{TS} \cdot \left(H^{t,i}_{TC} + H^{t,i}_{TD}\right) + \\ \xi^M_{PtC} \cdot P^{t,i}_{PtC} + \xi^M_{WT} \cdot P^{t,i}_{WT} \end{bmatrix} \cdot \Delta t \tag{10}$$

$$F^t_{EX}(y^t, W^t) = \left(\tau_B \cdot P^t_{BUY} - \tau_S \cdot P^t_{SELL}\right) \cdot \Delta t \tag{11}$$

$$\text{Constraints (1)} - \text{(6)},\ \text{(A1)} - \text{(A5)} \tag{12}$$

$$H^{p,t}_{TL} = H^{t,i}_{MT} + H^{t,i}_{PtC} - H^{t,i}_{TC} + H^{t,i}_{TD}, \quad pair(p, i) \tag{13}$$

$$H^{t,i}_{MT} = \eta_{MT} \cdot P^{t,i}_{MT} \tag{14}$$

$$H^{t,i}_{PtC} = \eta_{PtC} \cdot P^{t,i}_{PtC} \tag{15}$$

$$P^{t,b+1}_{AC} = \begin{bmatrix} P^{t,b}_{AC} - P^{t,0,b+1}_{AC} - P^{t,i}_{EL} + P^{t,i}_{WT} + \\ P^{t,i}_{PV} + P^{t,i}_{BD} - P^{t,i}_{BC} + P^{t,i}_{MT} - P^{t,i}_{PtC} \end{bmatrix} \tag{16}$$

$$Q^{t,b+1}_{RE} = Q^{t,b}_{RE} - Q^{t,0,b+1}_{RE} - Q^{t,i}_{RL} \tag{17}$$

$$V^{t,i+1}_{BUS} = V^{t,i}_{BUS} - \left(r^b \cdot P^{t,b}_{AC} + x^b \cdot Q^{t,b}_{RE}\right)/V_S \tag{18}$$

$$V^{i,\min}_{BUS} \le V^{t,i}_{BUS} \le V^{i,\max}_{BUS} \tag{19}$$

$$P^{i,\min}_{MT} \le P^{t,i}_{MT} \le P^{i,\max}_{MT} \tag{20}$$

$$0 \le P^{t,i}_{PtC} \le P^{i,\max}_{PtC} \tag{21}$$

$$P^t_{BUY} - P^t_{SELL} = P^{t,1}_{AC} \tag{22}$$

$$0 \le P^t_{SELL} \le P^{max}_{SELL} \tag{23}$$

$$0 \le P^t_{BUY} \le P^{max}_{BUY} \tag{24}$$

$$P^t_{BUY} \cdot P^t_{SELL} = 0 \tag{25}$$

$$0 \le P^{t,i}_{BC} \le P^{i,\max}_{BC} \tag{26}$$

$$0 \le P^{t,i}_{BD} \le P^{i,\max}_{BD} \tag{27}$$

$$E^{t,i}_{\mathrm{BS}} = (1 - \tau_{BS}) \cdot E^{t-1,i}_{\mathrm{BS}} + \left(P^{t,i}_{\mathrm{BC}} \cdot \eta_{BC} - P^{t,i}_{\mathrm{BD}}/\eta_{BD}\right) \cdot \Delta t \tag{28}$$

$$E^{i,\min}_{BS} \le E^{t,i}_{BS} \le E^{i,\max}_{BS} \tag{29}$$

$$P^{t,i}_{BC} \cdot P^{t,i}_{BD} = 0 \tag{30}$$

$$E^{0,i}_{BS} = E^{N_T,i}_{BS} \tag{31}$$

$$0 \le H^{t,i}_{TC} \le H^{i,\max}_{TC} \tag{32}$$

$$0 \le H^{t,i}_{TD} \le H^{i,\max}_{TD} \tag{33}$$

$$F^{MADP} = \min_{y^1 \in \Psi^1} F^1\left(y^1, W^1\right) + \mathbf{E}\left[\min_{y^2 \in \Psi^2\left(y^1, W^2\right)} F^2\left(y^2, W^2\right) + \mathbf{E}\left[\cdots + \mathbf{E}\left[\min_{y^{N_T} \in \Psi^{N_T}\left(y^{N_T-1}, W^{N_T}\right)} F^{N_T}\left(y^{N_T}, W^{N_T}\right)\right]\right]\right] \tag{7}$$

$$E_{\mathrm{TS}}^{t,i} = (1-\tau_{TS})\cdot E_{\mathrm{TS}}^{t-1,i} + \left(H_{\mathrm{TC}}^{t,i}\cdot\eta_{TC} - H_{\mathrm{TD}}^{t,i}/\eta_{TD}\right)\cdot\Delta t \quad (34)$$

$$E_{TS}^{i,\min} \le E_{TS}^{t,i} \le E_{TS}^{i,\max} \quad (35)$$

$$H_{TC}^{t,i}\cdot H_{TD}^{t,i} = 0 \quad (36)$$

$$E_{TS}^{0,i} = E_{TS}^{N_T,i} \quad (37)$$

The proposed multi-stage stochastic real-time MEMG operation model includes several nonlinear terms, such as the nonlinear functions in (2), (4), (6), and (A5) as well as bilinear terms in (25) and (30). The convexification and linearization approaches are adopted to convert them into linear forms, as detailed in the Appendix-B and C.

IV. Hybrid MPC-ADP-Based Solution Methodology

A. The MDP Reformulation

The multi-stage stochastic real-time operation model (7)-(37) can be reformulated as an MDP problem and solved by the traditional DP method. Specifically, DP decomposes the original high-dimensional problem into multiple sequential sub-problems that can be solved iteratively [23]. An MDP framework typically includes three categories of state, decision, and exogenous variables [23], [24], which are linked via the transition functions [22].

- **State variables** S^t is a set of variables that are used to describe the mathematical "state" of a dynamical system. This state provides sufficient information about the system to determine its future behavior in the absence of any external force affecting the system [19], [22]. In this sense, the energy levels of heterogeneous energy storage assets at the beginning of interval *t, i.e.*, the energy level at the end of interval *t-1* as shown in (38), are used as the state variables [20].

$$S^t = \left\{E_{TS}^{t-1,i}, E_{BS}^{t-1,i}\right\} \quad (38)$$

- **Decision variables** x^t describes the system actions when all the system states at interval t are observed [22]. They are given as follows:

$$x^t = \left\{\begin{array}{c} P_{BC}^{t,i}, P_{BD}^{t,i}, H_{TC}^{t,i}, H_{TD}^{t,i}, P_{BUY}^{t}, P_{SELL}^{t}, \\ P_{MT}^{t,i}, P_{PtC}^{t,i}, P_{AC}^{t,b}, Q_{RE}^{t,b}, V_{BUS}^{t,i}, H_{MT}^{t,i}, H_{PtC}^{t,i} \end{array}\right\} \quad (39)$$

- **Exogenous variables** $\hat{W}^t$ denote the forecasts of all the uncertainty sources at interval t. As this study focuses on uncertainties from the power system [15], the exogenous variables are denoted as follows:

$$\hat{W}^t = \left\{\hat{P}_{WT}^{t,i}, \hat{P}_{PV}^{t,i}, \hat{P}_{EL}^{t,b}, \hat{Q}_{RE}^{t,b}\right\} \quad (40)$$

- **Transition functions** map the current state to the next state according to the decision and exogenous variables, as below:

$$S^{t+1} = S^t + x^t + \hat{W}^t \quad (41)$$

Fig. 3. The relationship between *St x*, S^t, x^t, and $\hat{W}^t$.

Given the above MDP elements, the multi-stage stochastic real-time operation model can be reformulated as in (42).

Following Bellman's optimality principle, the tail problem of (42) starting at S^t of interval t can be described as in (43). $V^t(S^t)$ denotes the value function in state S^t, i.e., the optimal cost starting from state S^t, which includes the function value of the current state plus the cost-to-go function value. The cost-to-go function means the operation cost of the MEMG from the immediate future dispatch period $(t+1)$ to the last period N_T [23], [24]. γ is the discount factor that leverages the importance of the immediate reward and future rewards in the MDP. This value, by setting between 0 and 1, helps avoid an infinite reward in continuous tasks [28].

$$V^t(S^t) = \min_{x^t\ \hat{W}^t}\left\{\underbrace{C^t(S^t,x^t)}_{Current\ cost} + \gamma\cdot\underbrace{\mathbf{E}\left[V^{t+1}\left(S^{t+1}\right)|S^t\right]}_{Cost-to-gofunction}\right\} \quad (43)$$

B. The Proposed Hybrid MPC-ADP Method

The classical DP method solves Bellman's equations (43) backward through time to recursively obtain value functions corresponding to all individual possible states and then derives optimal solutions by solving Bellman's equations forward through time with obtained value functions. However, the extremely large state and action spaces would make it super hard to calculate the expected future cost in (43). This is the so-called "curse of dimensionality" issue in the traditional DP method, which makes the DP ineffective for problem-solving [6]. Hence, the ADP approach is developed as in [22], [23], which replaces the expectation in (43) by the value function $V_x^t(S_x^t)$ approximated around the post-decision state S_x^t, where $S_x^t = S^t + x^t$, to quantify the impact of current decisions on future costs [22]. The post-decision state *St x* is the state of MEMG soon after the decisions at interval t are made, but before any uncertainty information at $(t+1)$ is released (i.e., $\hat{W}^{t+1}$). The relationship between S^t, x^t, $\hat{W}^t$, and S_t^x is illustrated in Fig. 3.

The near-optimal solutions to (43) can be obtained by solving (44) given information of the current state.

$$V_x^{t-1}\left(S_x^{t-1}\right) = \mathbf{E}_{\hat{W}^t}\left\{\min_{x^t}\left[C^t(S^t,x^t) + \gamma\cdot V_x^t(S_x^t)|S_x^{t-1}\right]\right\} \quad (44)$$

However, ADP methods do not fully utilize the near-future forecasting information which is constantly updated in real time. The MPC method hence can be incorporated to leverage

$$\min_{x^1\ \hat{W}^1} C^1\left(S^1,x^1\right) + \mathbf{E}_{\hat{W}^2}\left[\min_{x^2} C^2\left(S^2,x^2\right) + \mathbf{E}_{\hat{W}^3}\left[\cdots + \mathbf{E}_{\hat{W}^{N_T}}\left[\min_{x^{N_T}} C^{N_T}\left(S^{N_T},x^{N_T}\right)\right]\right]\right] \quad (42)$$

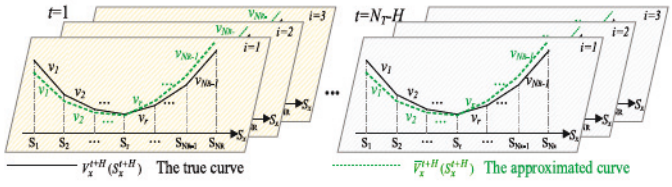

Fig. 4. The construction of the value function via PLFs.

such near-future forecasts, process more information, and achieve faster convergence [26]. As a result, the hybrid MPC-ADP method by taking advantage of both ADP and MPC is adopted in this study, which replaces (44) by (45) with an H-step look-ahead time horizon.

$$V_x^{t-1}\left(S_x^{t-1}\right) = \mathbf{E}_{\hat{W}^t,\ldots,\hat{W}^{t+H}}\left[\min_{x^t,\ldots,x^{t+H}}\left\{\begin{array}{l}\sum_{\tau=t}^{t+H} C^\tau(S^\tau, x^\tau)+\\ \gamma \cdot V_x^{t+H}\left(S_x^{t+H}\right)|S_x^{t-1}\end{array}\right\}\right] \tag{45}$$

Next, we discuss how to solve (34).

- **Successive Projective Approximation Routine (SPAR):** Calculating optimal solutions to (45) requires evaluating V_x^{t+H} (*). As stated before, it is computationally intractable via the traditional DP method when the state and action spaces are in extremely high dimensions [6]. In this regard, a set of PLFs can be used to approximate $V_x^{t+H}(S_x^{t+H})$ around S_x^{t+H} with the favorable computation burden [24] (see Fig. 4). Actually, other function forms besides the PLF can also be employed to approximate V_x^{t+H} (S_x^{t+H}); However, as the PLF only introduces LP problems that can be effectively solved by the industry-proven solvers like Gurobi and Cplex [15], it is used in most of the research works for the approximation [22], [24].

The mathematical formulations of PLFs are presented as in (46), where r and N_R are index and the total number of PLF blocks (i.e., the whole energy level is divided into N_R levels); $v_r^{t,i}$ is the slope of the PLF block r for energy storage of bus i at interval t; $y_r^{t,i}$ captures the energy quantity allocated to block r for energy storage of bus i at interval t.

$$V_x^{t+H}\left(S_x^{t+H}\right) \approx \bar{V}_x^{t+H}\left(S_x^{t+H}\right) = \sum_{i\in N_I, r\in N_R} y_r^{t,i}\cdot v_r^{t,i} \tag{46}$$

Note that as the original value function $V_x^{t+H}(S_x^{t+H})$ is convex, to ensure the same convexity of the constructed value function by PLFs, slopes of PLFs shall be monotone increasing as in Fig. 4. With the convexity property, no value can be assigned to later blocks until former ones are fully filled up [24].

Accordingly, (45) is converted into a linear programming problem as in (47)-(49), where m denotes the index of ADP iterations; S_{max}^i and S_{min}^i denote the minimal and maximal energy levels of the energy storage on bus i. In (47), the hybrid MPC-ADP method recursively minimizes the MEMG operation cost (*i.e.*, the summation of the H-step look-ahead cost and the cost-to-go value); As the energy storage capacity is evenly divided into N_R blocks in PLFs, $y_r^{t,i}$ is non-negative and no larger than the average value y_ρ^i in (48); The post-decision state or the final energy level in the storage at time t can be calculated by accumulating quantities of all blocks as in (49).

Algorithm 1 Data-Driven Offline Training of the Hybrid MPC-ADP Method

Step 1: Set the number of iterations N_M, the number of PLF blocks N_R, the initial energy level S_0^i in the storage, and all initial slopes $v_{r,0}^{t,i}$ (monotone increasing for $t = 1,\ldots,N_T-H+1$ and equal to 0 for $t = N_T - H$).
For $m = 1,\ldots,N_M$:
  For $t = 1,\ldots,N_T-H$:
    For $i = 1,\ldots,N_I$:
      Step 2: Observe uncertainty realizations for $[t, t+H]$ from historical data.
      Step 3: Obtain the decisions $x_m^t,\ldots,x_m^{t+H}$ and the post-decision state $S_{x,m}^{t+H}$ by solving (47)-(49);
      Step 4: Update slopes if $1 \le t \le N_T - H$
        Step 4a: Observe the gradients at $S_{x,m}^{t+H}$ and $S_{x,m}^{t+H}+y_\rho^i$, i.e., $v_m^{t+1,}$ $(S_{x,m}^{t+H})$ and $v_m^{t+1,i}(S_{x,m}^{t+H}+y_\rho^i)$, via (50).
        Step 4b: Get the intermediate slopes $z_m^{t,i}$ for $S_{x,m}^{t+H}$ and $S_{x,m}^{t+H}+y_\rho^i$ as in (51).
        Step 4c: To update slopes of PLFs while retaining the convexity, perform the projection operation to get the final slopes: $v_m^{t,i} = \Pi_C(v_{m-1}^{t,i})$ as in (53).
    End for (i);
  End for (t);
End for (m);
Step 5: Output all the resulted PLFs.

$$\min_{x^t,\ldots,x^{t+H},y_r^{t,i}\hat{W}^t,\ldots,\hat{W}^{t+H}}\left[\sum_{\tau=t}^{t+H} C^\tau\left(S^\tau,x^\tau\right)+\gamma \cdot \sum_{i\in N_I, r\in N_R} y_r^{t,i}\cdot v_{r,(m-1)}^{t,i}|S_x^{t-1}\right] \tag{47}$$

$$s.t. 0 \le y_r^{t,i} \le \left(S_{max}^i - S_{min}^i\right)/N_R = y_\rho^i \tag{48}$$

$$\sum_{r\in N_R} y_r^{t,i}\cdot v_r^{t,i} = S_x^{t+H} \tag{49}$$

The problem in (47)-(49) can be solved by an exploitation algorithm named SPAR [25]. The SPAR method can provably learn the optimal decisions to be taken at parts of the state space, which can be reached by an optimal policy. By learning the slopes of PLFs at important parts of state spaces and constructing V_x^t (*) that are convex with all the state points as in Fig. 4, the exact (but unknown) value function V_x^t (*) and its approximation $\bar{V}_x^t$ (*) will match only in the vicinity of the optimum. Given the convexity property of the original problem, after the intensive training by visiting different states infinitely often, the corresponding approximation slopes can converge to the slopes of the optimal value functions [24]. In addition, as the SPAR algorithm learns an optimal decision for all states that can be reached by an optimal policy, decisions obtained can be regarded as the optimal solution. The detailed theoretical proof for the convergence and optimality can be referred to Section V of [24].

With the SPAR method, the MPC-ADP approach can be trained offline with the historical data as the inputs (i.e., data-driven) to derive slopes $vt,i\ r$, as outlined in Algorithm 1 [24].

To update the slopes of PLFs, in the *mth* iteration, a sample observation of the marginal value $v_m^{t,i}(*)$ in the storage is calculated as in (50).

$$\mathrm{v}_m^{t+1,i}\left(S_{x,m}^{t+H}\right) = \partial\left[V^{t+1}\left(\mathrm{S}_{x,m}^{t+H}\right)\right]/\partial S_{x,m}^{t+H} \tag{50}$$

Then, in *Step 4b*, a set of ancillary slopes $z_m^{t,i}$ is produced by combining the current approximation and sample slopes via the step size or learning factor α_m^t as (51).

$$z_m^{t,i}(*) = \left(1-\alpha_m^t\right)\cdot v_{m-1}^{t,i}(*) + \alpha_m^t \cdot \underline{v}_m^{t+1,i}(*) \tag{51}$$

α_m^t is constrained by (52) with B being a finite positive constant and no less than 1 [24].

$$\left\|\alpha_m^t\right\|_1 = \infty, \quad \left\|\alpha_m^t\right\|_2 \le B \le \infty \tag{52}$$

The intermediate slopes $z_m^{t,i}$ from (51) are used to update the slopes of PLFs while retaining the convexity of PLFs. That is, the projection operation (53) in *Step 4c* is applied to force violated slopes equal to the newly updated ones, maintaining the slopes of the curves to be monotone increasing. In (53), S_r^i describes the energy level corresponding to the r-th breakpoint of the PLF block for the energy storage at bus i.

$$\Pi_C\left(v_{r,m-1}^{t,i}\right) = \begin{cases} z_m^{t,i}\left(S_{x,m}^{t+H}\right) \\ \text{if}\left(S_r^i \le S_{x,m}^{t+H}\right)\&\left(v_{r,m-1}^{t,i}\left(S_r^i\right) \ge z_m^{t,i}\left(S_{x,m}^{t+H}\right)\right) \\ z_m^{t,i}\left(S_{x,m}^{t+H}+y_\rho^i\right) \\ \text{if}\left(S_r^i \ge S_{x,m}^{t+H}+y_\rho^i\right)\&\left(v_{r,m-1}^{t,i}\left(S_r^i\right) \le z_m^{t,i}\left(S_{x,m}^{t+H}+y_\rho^i\right)\right) \\ v_{r,m-1}^{t,i}\left(S_r^i\right) \quad \text{otherwise} \end{cases} \tag{53}$$

Equation (53) implicates the following rules to retain the convexity (monotone increasing) of the PLF curves:

a) If the energy level S_r^i corresponding to the r-th breakpoint of the PLF block is no larger than $S_{x,m}^{t+H}$ but the corresponding slope from the last iteration $v_{r,(m-1)}^{t,i}$ is no less than $z_m^{t,i}(S_{x,m}^{t+H})$, the slopes of blocks with energy level no larger than $S_{x,m}^{t+H}$ are updated to be $z_m^{t,i}(S_{x,m}^{t+H})$;

b) If S_r^i is no less than $S_{x,m}^{t+H}+y_\rho^i$ but $v_{r,(m-1)}^{t,i}$ is no larger than $z_m^{t,i}(S_{x,m}^{t+H}+y_\rho^i)$, the slopes of blocks with energy level no less than $S_{x,m}^{t+H}+y_\rho^i$ are updated to be $z_m^{t,i}(S_{x,m}^{t+H}+y_\rho^i)$;

c) For all other cases, slopes of the PLF blocks remain the same with the $v_{r,m-1}^{t,i}$.

- **McClain's Step-Size Rule:** In (53), the costant ($\alpha_m^t = \alpha_0, 0 < \alpha_0 < 1$) and harmonic ($0 < \alpha_0 < 1, \alpha_m^t = \alpha_{m-1}^t)/m$ step sizes are commonly used in the literature [22], [24]. However, the constant step size would make the learning process oscillating after several training iterations while the harmonic step size declines too fast to converge. Thus, McClain's step size in (54) which combines advantages of both approaches can be applied, where $\underline{\alpha}$ is a tunable parameter in the range of [0, α_0].

$$\alpha_m^t = \begin{cases} \alpha_0 & \text{if } m = 1 \\ \alpha_{m-1}^t/\left(1+\alpha_{m-1}^t-\underline{\alpha}\right) & \text{if } m \ge 2 \end{cases} \tag{54}$$

As a matter of fact, McClain's step size behaves like the harmonic rule for early iterations, and as the step size approaches $\underline{\alpha}$, it starts mimicking the constant step size rule. All the nonzero step sizes can capture changes that may occur in the later iterations.

C. Real-Time MEMG Operation

After converged slopes are obtained in Algorithm 1 after the intensive offline training, the MPC-ADP method has been embedded with empirical knowledge. Then, the pre-trained PLFs are used to calculate the near-optimal real-time MEMG operation solutions as Algorithm 2 [22].

Algorithm 2 Real-Time MEMG Operation via the Trained Slopes From MPC-ADP

for $t = 1, \ldots, N_T$
  Step 1: Input the real-time information of all uncertainty sources at interval t;
  Step 2: Calculate x^t and S_x^t by solving (47)-(49) with the trained slopes;
End for (t);

From Algorithms 1 and 2, it can be seen that effective offline training is the pre-requisite for an excellent real-time operation, while real-time operation is the direct application of the offline training results. In this sense, they are effectively coordinated with each other.

D. Performance Evaluation

To evaluate the convergence performance and solution quality of our proposed MPC-ADP method, it would be ideal to use the original multi-stage real-time SP model (7) as the comparison base. However, solving (7) would be overly computationally demanding, given the excessively high problem dimension. Alternatively, the anticipative SP operation model in (55) is used as the base, which assumes that for each scenario s, the information for the entire day is known prior to $t = 1$. The anticipative method (55) is computationally friendly with complete information prior to the operation [22]. However, with regard to the Jensen's inequality [29], the resulted solution $\sum_s F_s^{ASP}$ is a lower bound estimation to the original multi-stage non-anticipative real-time problem (7).

$$F_s^{ASP} = \min_{y^{t,s}\in\Psi^{t,s}} \sum_{t=1}^{N_T} F^t\left(y^{t,s}\right) \tag{55}$$

With this, for the offline training, the solution gap E_m^{OMADP} of the MPC-ADP approach with regard to the anticipative method after the m-th offline training iteration is formulated as in (56), which is evaluated via the average system operation costs of m offline training scenarios for the MPC-ADP method and the anticipative method. F_s^{MADP} and F_s^{ASP} are system operation costs of the MPC-ADP method and the anticipative method for the s-th offline training scenario.

$$E_m^{OMADP} = \frac{\left(\sum_{s=1}^m F_s^{MADP}\right)/m - \left(\sum_{s=1}^m F_s^{ASP}\right)/m}{\left(\sum_{s=1}^m F_s^{ASP}\right)/m} \tag{56}$$

For the real-time operation, the solution gap E_s^{RMADP} of the MPC-ADP approach in the s-th real-time operation application is denoted as in (57). The finalized average solution gap for real-time operation can be calculated in the same way as (56).

$$E_s^{RMADP} = \left(F_s^{MADP} - F_s^{ASP}\right)/F_s^{ASP}. \tag{57}$$

V. Case Studies

A grid-tied MEMG, combining an IEEE 15-bus distribution network and an 8-node thermal network in Tianjin, China, is used to evaluate the effectiveness of the proposed MPC-ADP

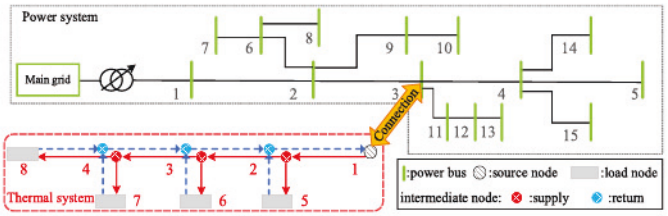

Fig. 5. The structure of the test MEMG.

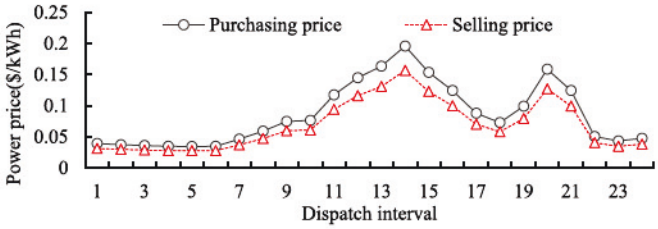

Fig. 6. Power transaction price between the MEMG and main grid.

TABLE I
PARAMETERS OF GENERATORS AND STORAGE [9], [11], [12]

Bus	WT (kW)	PV (kW)	BS (kW/kWh)	TS(kW/kWh)	CCHP (kW)	PtC (kW)
3	300	150	-	-	1,000	500
4	250	100	500/1,500	500/1,600	-	-
6	200	-	-	-	-	-
9	-	200	-	-	-	-

TABLE II
PARAMETERS OF THE THERMAL NETWORK [15]

Start node	End node	Length (m)	Type*	Mass flow rate (kg/h)
1	2	1,050	DN80	23,017.08
2	3	525	DN78	20,905.19
3	4	525	DN60	12,457.36
2	5	525	DN32	2,111.89
3	6	225	DN50	8,447.83
4	7	525	DN54	10,345.87
4	8	225	DN32	2,111.49

*DN is a technical term in thermal/hydraulic engineering and means "nominal diameter".

approach. As shown in Fig. 5, electrical and thermal networks are coupled via the CCHP plant and PtC unit on bus 3 of the electricity network and node 1 of the thermal network. For the electricity network, V_s is set as 1.0 p.u., and voltages limits of all other buses are sets as [0.95, 1.05] p.u. [9], [27].

The electricity transaction prices are given in Fig. 6. Parameters for generators and energy storage devices are given in Table I. The thermal network parameters are listed in Table II. The temperature of supply and return pipes is limited as [353.15, 373.15] K and [323.15, 343.15] K.

The winter case is studied, thus only heat and electricity demands are to be satisfied. Some of the key MEMG operation parameters are listed in Table III.

The real-time operation window is 24 hours with 1-hour granularity. One deterministic forecast (lines with marks) and multiple scenarios (generated via the historical data) of all uncertainty sources are presented in Fig. 7 [12]. The data in Fig. 7 represent the observed historical data of uncertainty sources in all dispatch intervals, which are used in the training process of the MPC-ADP method to calculate the finalized operation cost of the MEMG.

TABLE III
MEMG OPERATION PARAMETERS [12], [15], [27]

η_{ME}	η_{BC}/η_{BD}	ξ^{M}_{TS}	ξ^{M}_{PV}
0.29	0.98	0.005$/kWh	0.0035$/kWh
η_{PtC}	η_{TC}/η_{TD}	ξ^{M}_{MT}	ξ^{M}_{BS}
3	0.95	0.0032$/kWh	0.01$/kWh
η_{MT}	ξ_{NG}	ξ^{M}_{WT}	ξ^{M}_{PtC}
1.26	0.03578$/kWh	0.003$/kWh	0.002$/kWh

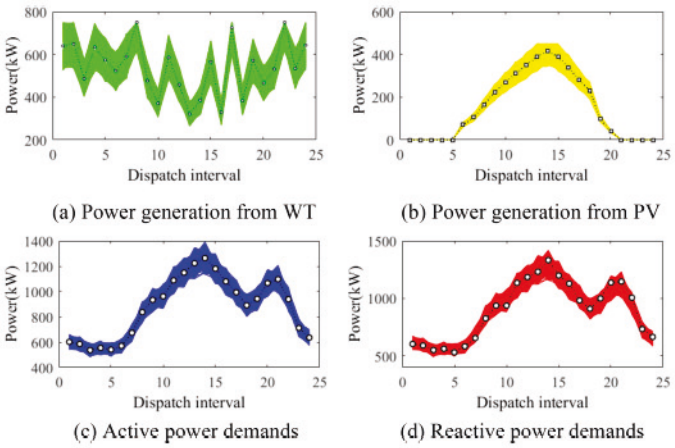

Fig. 7. Deterministic forecast and uncertainty scenarios generated via the historical data in winter.

For the MPC-ADP method, the look-ahead horizon H is set as 2 hours. As the original problem is bounded and the future reward is as important as the current one, the discount factor γ is set as 1 [22], [24]. The initial energy levels for the BS and TS are set as their minimum values. The PLF for each storage at each time interval has 20 equal blocks. The initial slope for the first block is set as 0, with the step increment of 0.1 for successive blocks. $\alpha_0 = 0.9$ and $\underline{\alpha} = 0.2$ are used in (42) as shown at the bottom of the p. 6, to set McClain's step size.

The simulations are implemented via Yalmip with MATLAB and executed on a 64-bit PC with a 3.70 GHz CPU and 16 GB RAM. The problems in (47)-(49) are solved by Gurobi.

A. Deterministic Benchmarking

The proposed MPC-ADP algorithm is first benchmarked on a deterministic version of the MEMG operation problem to analyze its computational performance and solution quality [30]. That is, a single deterministic scenario is utilized to iteratively train the PLFs and to calculate the real-time MEMG operation. With this, the global optimal real-time solution can be obtained via deterministic linear programming (LP) for exact comparison.

- **Computational Performance:** The convergence curve in terms of solution gaps (43) is shown in Fig. 8. The MPC-ADP method starts from the gap similar to the MPC method and converges in less than $N_M = 100$ iterations via Algorithm 1 (i.e., approaching the true optimal condition with a 2.81% solution gap in 52 iterations).

The approximated value functions of iterations 1-10 and 31-40 are given in Fig. 9. It clearly shows that the differences among function values of iterations 1-10 are rather significant,

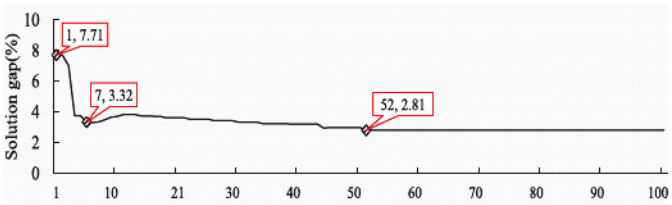

Fig. 8. The convergence curve of the MPC-ADP method.

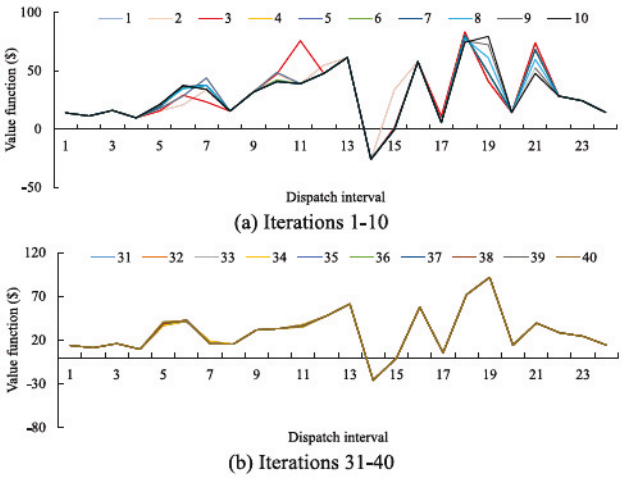

Fig. 9. Function values of iterations 1-10 and iterations 31-40.

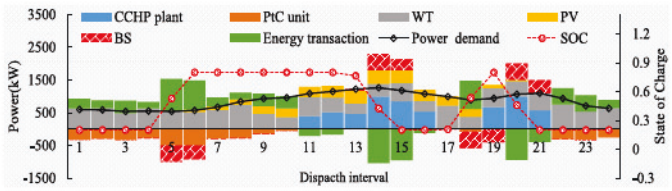

Fig. 10. The electrical energy balance condition in the MEMG.

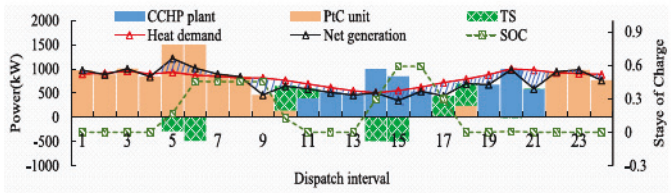

Fig. 11. The thermal energy balance condition in the MEMG.

but after further training, they gradually converge in iterations 31-40.

In addition, the solution time of the single deterministic LP model [22] is 2.23s, and the solution time of the MPC-ADP method is only 0.88s. This is because the proposed approach approximates the original 24-hour LP model via 24 smaller-scale single-hour LP models, by taking the advantage of the ADP method.

- **Solution Quality:** With the deterministic inputs, the resulted electrical and thermal energy balance conditions from the hybrid MPC-ADP method are presented in Figs. 10 and 11. They clearly show that solutions from the MPC-ADP method can effectively coordinate all sources to meet the multi-energy demands.

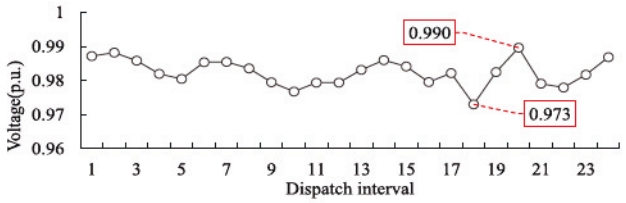

Fig. 12. The voltage profile of bus 8 for all the dispatch intervals.

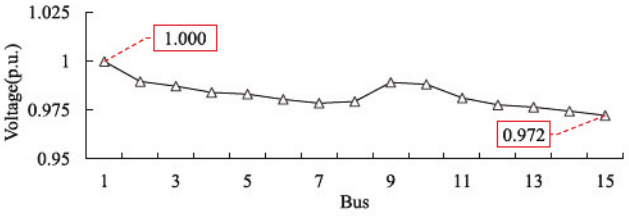

Fig. 13. The voltage profile of all buses at dispatch interval 12.

(i) As for electricity/thermal generations and electrical energy transactions, when electricity prices are low (*i.e.*, intervals 1-10 and 22-24 in Fig. 6), the MEMG purchases electrical energy from the main grid, and thermal energy is mostly supported by the PtC using electrical energy. With high electricity prices (*i.e.*, intervals 11-16 and 19-21 in Fig. 6), the MEMG sells electrical energy to the main grid, and thermal energy is dominantly supported by the CCHP. That is, the MEMG can leverage the temporal differences of electricity prices and multi-energy loads to optimally coordinate the multi-energy generation and transaction.

(ii) As for energy storage assets, TS stores thermal energy when thermal generation cost is low (*i.e.*, intervals 5-6 and 14-15), and releases it at high thermal generation cost intervals 10-11 and 17-18. In this way, TS can effectively reduce the thermal supply cost by coordinating with the CCHP and PtC. BS collaborates with all generators to shift power loads for peak-shaving and cost-saving purposes. That is, BS charges at intervals 5-6 and 18-19 when electricity prices and load levels are low in Fig. 6 while discharging at intervals 14-15 and 20-21 with high electricity prices in Fig. 6 and load levels. The real-time operation decisions of TS and BS illustrate that the MPC-ADP method can effectively capture dynamic behaviors via the step-by-step learning process.

(iii) As for thermal inertia, Fig. 11 shows that the thermal generation and loads at individual time intervals may not be equal. This is because different from instant electrical energy transmission, thermal inertia makes the whole thermal network a virtual storage that can coordinate multiple dispatch intervals. The shaded area between thermal generation and load curves in Fig. 10 indicates its storage capability.

To show the effectiveness of our proposed method in guaranteeing the secure operation, the voltage profile of bus 8 for all dispatch intervals in the deterministic case after 100 training iterations is shown in Fig. 12, and the voltage profile of the whole distribution system at interval 12 is given in Fig. 13. In Figs. 12 and 13, the two labeled values indicate the maximal and minimal voltage levels of the voltage profile.

The voltage profile in Fig. 12 shows that although the voltage values of the same bus at different dispatch intervals

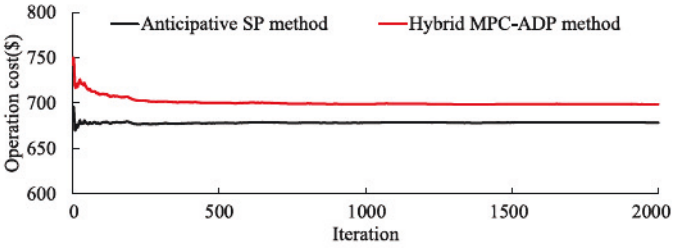

Fig. 14. Operation cost comparison of MPC-ADP and anticipative SP methods.

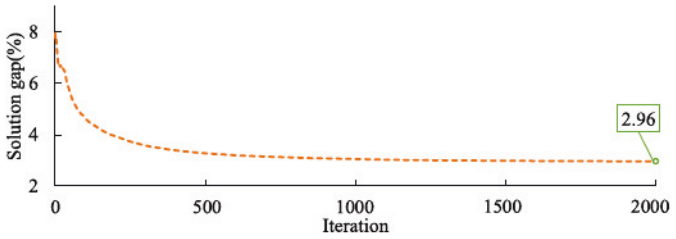

Fig. 15. Convergence curves of the MPC-ADP method in terms of solution gap.

vary, it still falls within the safe range, i.e., 0.95-1.05 p.u. Furthermore, Fig. 13 shows that voltage values of all buses are also in the safe range. The voltage results in Figs. 12 and 13 indicate that after training, the proposed MPC-ADP method could still guarantee voltage security of the MEMG operation.

In summary, the above results illustrate that the proposed MPC-ADP method can contribute to obtaining good-enough solutions, reducing the solution time significantly, and guaranteeing the secure operation of MEMGs.

B. Stochastic Simulations

The stochastic simulation involves 3,000 scenarios sampled from the historical data in Fig. 7, with 2,000 scenarios (i.e., N_M in *Algorithm 1* is 2,000) used for the offline training and 1,000 scenarios for the real-time application.

Fig. 14 shows the average operation costs of the 2,000 scenarios from the proposed MPC-ADP offline training and the anticipative SP method, and the corresponding solution gaps are shown in Fig. 15. As the anticipative SP method assumes the full availability of the 24-hour forecast information, the corresponding operation cost is used as a lower bound to evaluate the solution quality of the MPC-ADP approach. The figures show that similar to the deterministic case, the cost obtained from the MPC-ADP method starts from a high value close to the MPC method and gradually approaches the SP solution within the 3% gap (note that as the anticipative SP method only provides a lower bound to the true optimal solution of the original multi-stage problem (7), the actual gap of the MPC-ADP solution to the true optimal solution shall be smaller than 3%). This is rational because in the first iteration, given the prespecified positive slopes of PLFs, BS and TS do not work as they shall. That is, the system operation solution from the first iteration is myopic, which incurs additional costs according to (47). After informative training and learning, parts of the slopes in the PLFs become negative, which drives BS and TS to gradually operate towards the optimal operation conditions. Thus, the solution gaps gradually reduce. However, because the MPC-ADP method is an approximation approach in nature, the solution gap may not be able to reach 0 exactly [30]. The results verify the effectiveness of the proposed MPC-ADP method in obtaining good-enough solutions for the multi-stage real-time coordinated MEMG operation.

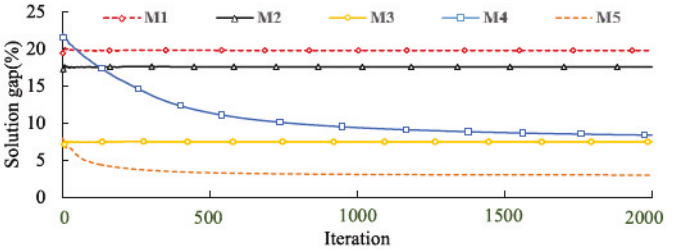

Fig. 16. The convergence curves of all the comparison cases in winter.

After the slopes of PLFs are sufficiently trained, the real-time operation for 1,000 scenarios is conducted. The average solution gap of the real-time operation cost is 2.90%, which shows the effectiveness of the MPC-ADP method in learning the optimal system operations. Furthermore, as the MPC-ADP method embeds the historical knowledge and solves the operation forward through time, it can obtain good-enough solutions for the current interval t even without knowing system information for $t+1$ to N_T. On the contrary, the SP methods heavily rely on the full availability of the 24-hour forecasted or simulated information. This is another advantage of the MPC-ADP method.

C. Comparison With Other Real-Time Operation Schemes

The proposed MPC-ADP method is further compared with other real-time MEMG operation schemes to illustrate its advantages.

- **M1**: Myopic policy, i.e., the real-time operation decision at hour t is determined by solving a stochastic operation problem involving information and variables for hour t only.
- **M2**: Myopic H-hour look-ahead policy, i.e., operation decisions for future H hours are determined by solving a stochastic problem with the information and variables for every H hours.
- **M3**: Traditional MPC-based H-hour look-ahead policy.
- **M4**: Traditional ADP method, i.e., H in (34) is set as 0.

The convergence curves of **M1**-**M4** and the proposed MPC-ADP method (**M5**) are reported in Fig. 16. From Fig. 16, the below can be observed:

(i) The myopic and MPC methods (i.e., **M1**, **M2**, and **M3**) could only obtain local optimal solutions with high solution gaps (against the optimal solution from the anticipative SP method). Indeed, their gaps do not show a convergence trend. In comparison, final solution gaps of the ADP-based methods (i.e., **M4** and **M5**) converge to the optima gradually. It indicates that the ADP-based methods contribute to learning from the historical data and converging to the global optimal solutions.

TABLE IV
REAL-TIME OPERATION RESULTS OF ALL THE METHODS IN WINTER

#Method	M1	M2	M3	M4	M5
Average solution gap	19.70%	14.47%	7.40%	7.08%	2.90%
Average computation time	0.72s	0.42s	0.84s	0.76s	0.87s

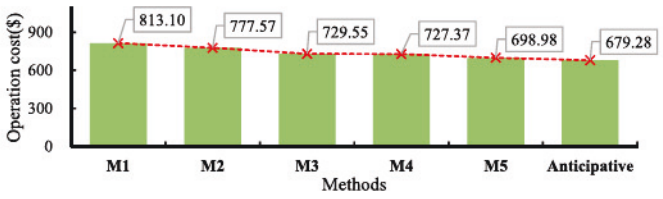

Fig. 17. The final operation costs of all the methods in the winter case.

(ii) Incorporating more information could lead to solutions of higher quality. This is supported by the fact that the solution gaps of **M3** are lower than those of **M2** and **M1**. This also explains that **M5** has a much smaller solution gap in the first iteration and converges faster to the optima than **M4**. This demonstrates the benefits of integrating MPC and ADP for the multi-stage real-time operation.

The average gaps of the real-time operation solutions from the five methods in 1,000 scenarios, together with their average computational time, are further compared in Table IV. The final operation costs of all the methods in the winter case are compared in Fig. 17.

With the real-time operation results in Table IV and Fig. 17, it can be inferred that the proposed MPC-ADP method has the smallest real-time solution gap and the lowest operation cost with regard to the anticipative SP method. Furthermore, the computation time of our method is acceptable for real-time applications. Thus, all the comparison results indicate the effectiveness of the hybrid MPC-ADP method in obtaining the near-optimal solutions with a faster convergence rate.

D. A Summer Case

To further illustrate the effectiveness of our proposed MPC-ADP approach for the MEMG under various application occasions, an additional case study for summer is further conducted. For the summer case, system parameters including capacities of generators and energy storage assets as well as power transaction prices between the MEMG and main grid, etc., remain the same as the winter case. On the other hand, in summer, space cooling demands are needed with the TS referring to cooling energy storage, and temperature ranges of supply and return pipes are respectively set as [308.15, 333.15] K and [338.15, 363.15] K [12].

The deterministic forecast (lines with the mark) and 3,000 scenarios (sampled from historical data) of all uncertainty sources are presented in Fig. 18 [15].

The parameter setting for the MPC-ADP algorithm is the same as the winter case. Among the 3,000 scenarios, 2,000 scenarios are used for the offline training and the remaining 1,000 are used for the real-time application. Afterward, the final simulation results of our proposed method **M5** are compared with the traditional methods **M1**-**M4** in Fig. 19.

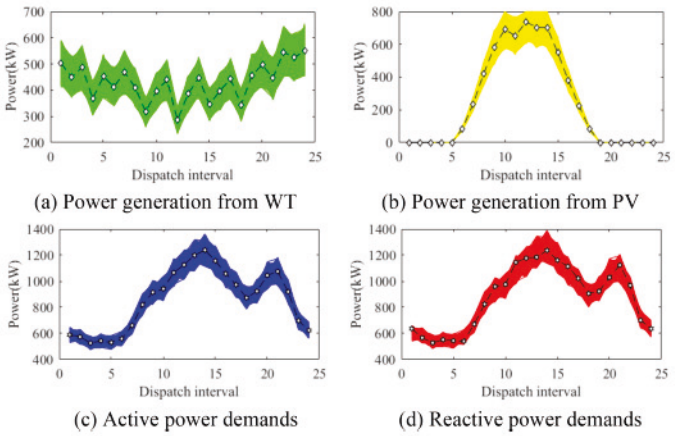

Fig. 18. Deterministic forecast and uncertainty scenarios generated via the historical data in summer.

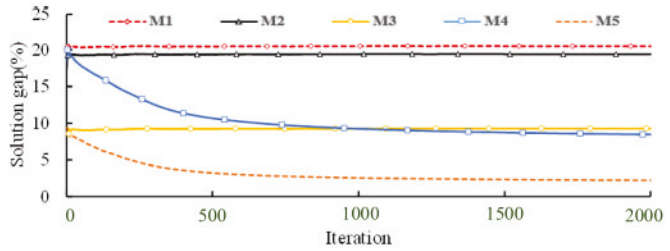

Fig. 19. The convergence curves of all the comparison cases in summer.

TABLE V
REAL-TIME OPERATION RESULTS OF ALL THE METHODS IN SUMMER

#Method	M1	M2	M3	M4	M5
Average solution gap	20.38%	19.29%	9.23%	6.94%	1.95%
Average computation time	0.71s	0.40s	0.83s	0.74s	0.86s

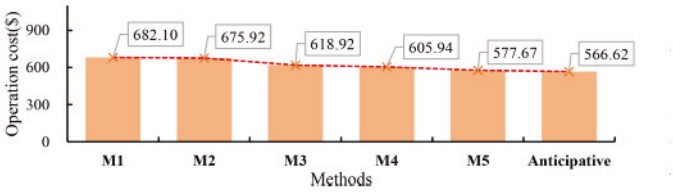

Fig. 20. The convergence curves of all the comparison cases in summer.

After the training, average solution gaps of the real-time operation from all the five methods in 1,000 scenarios are compared in Table V. The final operation costs of all methods in the summer case are shown in Fig. 20.

From the simulation results in Figs. 19–20 and Table V, it can be seen that our MPC-ADP method remains valid in the summer case, with good-enough solutions and fast convergence. To this end, the superiority of our proposed hybrid MPC-ADP method is comprehensively validated via various cases.

VI. CONCLUSION

This paper proposes a data-driven MPC-ADP method for the multi-stage stochastic real-time operation of an MEMG considering practical thermal flow constraints as well as heterogeneous energy storage assets. The historical data are

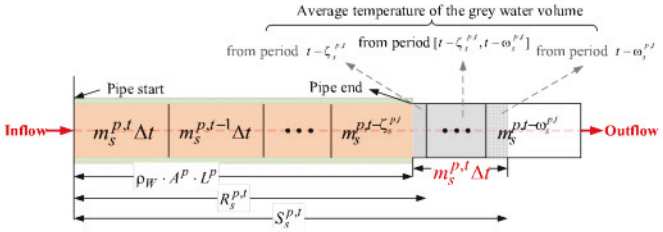

Fig. AI. The vertical section of the supply pipe p.

used for the offline training of the MPC-ADP method, which approximates the original solution space via sets of PLFs for energy storage assets. The simulation results indicate that by making full use of the historical data, the MPC-ADP method could mimic the original multi-stage SP approach and derive good-enough solutions, while is computational more efficient than the non-anticipative SP approach. The comparative studies with other real-time operation benchmarks indicate that the MPC-ADP method could derive better solutions to guarantee the secure and economic operations of MEMGs with a much faster convergence process. Furthermore, the sensitivity analysis is conducted with regard to various application scenarios and all the corresponding results justify the advantage of the hybrid MPC-ADP method.

For future work, the policy-based ADP approach could be explored to study the real-time optimal operations of MEMGs while rigorously considering their non-linear and/or non-convex physical and operational characteristics [6]. In addition, without the powerful support from the main grid, the real-time operation of the islanded MEMG via the hybrid MPC-ADP method can be further investigated.

APPENDIX

A. Inertia-Based Thermal Transmission Delay

To quantify the transmission delay indicated in Section II, the vertical section of a supply pipe as shown in Fig. AI is utilized [15].

In Fig. A1, $m_s^{p,t}$ is the mass flow rate of the supply pipe at time t; the orange and grey volumes denote the water flowing in and out of a supply pipe; L^p is the length of pipe p; A^p denotes the area of the cross-section. With this, the water volume in this pipe can be calculated as $\rho_W A^p L^p$, where ρ_W is the water density. The grey volume at the right side means the water mass that has flowed out by the end of time interval t. Specifically, after the interval $\zeta_s^{p,t}$, the water starts to flow out; while after $\omega_s^{p,t}$, the entire water mass completely flows out. Regarding this, $\zeta_s^{p,t}$ and $\omega_s^{p,t}$ are calculated as in (A1) and (A2), where t and t_p are indices of operation time intervals, t_m and t_n are the objectives to be minimized, Δt is the unit dispatch interval, and N_T is the total number of dispatch intervals.

$$\zeta_s^{p,t} = \min_{t_m \in N_T} \left\{ t_m : \sum_{t_p=t-t_m}^{t} m_s^{p,t_p} \cdot \Delta t \geq \rho_W \cdot A^p \cdot L^p \right\} \tag{A1}$$

$$\omega_s^{p,t} = \min_{t_n \in N_T} \left\{ t_n : \sum_{t_p=t-t_n}^{t} m_s^{p,t_p} \cdot \Delta t \geq \rho_W \cdot A^p \cdot L^p + m_s^{p,t} \cdot \Delta t \right\} \tag{A2}$$

Accordingly, $R_s^{p,t}$ and $S_s^{p,t}$ respectively represent the total water mass flowing into the pipeline during periods $[t-\zeta_s^{p,t}, t]$ and $[t-\omega_s^{p,t}, t]$, which can be calculated as in (A3) and (A4).

$$R_s^{p,t} = \sum_{t_p=t-t_m}^{t} \left(m_s^{p,t_p} \cdot \Delta t \right) \tag{A3}$$

$$S_s^{p,t} = \begin{cases} \sum_{t_p=t-t_n+1}^{t} \left(m_s^{p,t_p} \cdot \Delta t \right), & t_n \geq t_m + 1 \\ R_s^{p,t}, & t_n < t_m + 1 \end{cases} \tag{A4}$$

Finally, the equivalent temperature $T_{s,(+)}^{p,t}$ at the start of the p-th supply pipe in interval t can be estimated as the average temperature of water in the grey volume of Fig. A2, as shown in (A5) [16].

$$T_{s,(+)}^{p,t} = \left[\begin{array}{l} \left(m_s^{p,t} \cdot \Delta t + \rho_W \cdot A^p \cdot L^p - R_s^{p,t} \right) \cdot T_{s,(+)}^{p,t-\zeta_s^{p,t}} \\ + \sum_{t_p=t-\zeta_s^{p,t}+1}^{t-\omega_s^{p,t}-1} \left(m_s^{p,t_p} \cdot \Delta t \cdot T_{s,(+)}^{p,t_p-\zeta_s^{p,t}} \right) \\ + \left(S_s^{p,t,} - \rho_W \cdot A^p \cdot L^p \right) \cdot T_{s,(+)}^{p,t-\omega_s^{p,t}} \end{array} \right] / \left(m_s^{p,t} \Delta t \right) \tag{A5}$$

B. Linearizing the Thermal Flow Model in Constraints (2), (4), (6), and (A5)

Though mass flow rate and temperature would vary in the thermal network operation, it remains a practical assumption that the mass flow rate is constant [15]. Thus, the thermal flow model with the constant mass flow rate presents a linear formulation.

C. Relaxing Bilinear Terms in (25) and (30)

The bilinear terms can be relaxed without compromising the solution optimality.

Proof: Take (25) for instance. As the objective (7) is monotone increasing with $P_{BC}^{t,i}$ and $P_{BD}^{t,i}$, if both were non-zero with $P_{BC}^{t,i} \geq P_{BD}^{t,i}$, there must exist another pair $\underline{P}_{BC}^{t,i}$ and $\underline{P}_{BD}^{t,i}$, with $\underline{P}_{BC}^{t,i} - \underline{P}_{BD}^{t,i} = P_{BC}^{t,i} - P_{BD}^{t,i}$ and $\underline{P}_{BD}^{t,i} = 0$. This would produce the same net BS power injection at a cheaper cost. This completes the proof.

REFERENCES

[1] E. Guelpa, A. Bischi, V. Verda, M. Chertkov, and H. Lund, "Towards future infrastructures for sustainable multi-energy systems: A review," *Energy*, vol. 184, pp. 2–21, Oct. 2019.

[2] J. Wang, S. You, Y. Zong, C. Træholt, Y. Zhou, and S. Mu, "Optimal dispatch of combined heat and power plant in integrated energy system: A state of the art review and case study of Copenhagen," *Energy Procedia*, vol. 158, pp. 2794–2799, Feb. 2019.

[3] Z. Li and Y. Xu, "Optimal coordinated energy dispatch of a multi-energy microgrid in grid-connected and islanded modes," *Appl. Energy*, vol. 210, pp. 974–986, Jan. 2018.

[4] H. Pashaei-Didani, S. Nojavan, R. Nourollahi, and K. Zare, "Optimal economic-emission performance of fuel cell/CHP/storage based microgrid," *Int. J. Hydrogen Energy*, vol. 44, no. 13, pp. 6896–6908, 2019.

[5] C. Mu, T. Ding, Z. Zeng, P. Liu, Y. He, and T. Chen, "Optimal operation model of integrated energy system for industrial plants considering cascade utilisation of heat energy," *IET Renew. Power Gener.*, vol. 14, no. 3, pp. 352–363, 2019.

[6] W. B. Powell and S. Meisel, "Tutorial on stochastic optimization in energy—Part I: Modeling and policies," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1459–1467, Mar. 2015.

[7] C. Zhang, Y. Xu, and Z. Y. Dong, "Robustly coordinated operation of a multi-energy micro-grid in grid-connected and islanded modes under uncertainties," *IEEE Trans. Sustain. Energy*, vol. 11, no. 2, pp. 640–651, Apr. 2020.
[8] L. Wang, Q. Li, M. Sun, and G. Wang, "Robust optimisation scheduling of CCHP systems with multi-energy based on minimax regret criterion," *IET Gener. Transm. Distrib.*, vol. 10, no. 9, pp. 2194–2201, 2016.
[9] C. Zhang, Y. Xu, Z. Li, and Z. Y. Dong, "Robustly coordinated operation of a multi-energy microgrid with flexible electric and thermal loads," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2765–2775, May 2019.
[10] T. Zhao, X. Pan, S. Yao, C. Ju, and L. Li, "Strategic bidding of hybrid AC/DC microgrid embedded energy hubs: A two-stage chance constrained stochastic programming approach," *IEEE Trans. Sustain. Energy*, vol. 11, no. 1, pp. 116–125, Jan. 2020.
[11] Y. Chen, X. Feng, Z. Li, Y. Xu, and A. Miragha, "Multi-stage coordinated operation of a multi-energy microgrid with residential demand response under diverse uncertainties," *Energy Convers. Econ.*, vol. 1, no. 1, pp. 20–33, 2020.
[12] Z. Li and Y. Xu, "Temporally-coordinated optimal operation of a multi-energy microgrid under diverse uncertainties," *Appl. Energy*, vol. 240, pp. 719–729, Apr. 2019.
[13] S. Sharma, Y. Xu, A. Verma, and B. K. Panigrahi, "Time-coordinated multienergy management of smart buildings under uncertainties," *IEEE Trans. Ind. Informat.*, vol. 15, no. 18, pp. 4788–4798, Aug. 2019.
[14] Y. Chen, Y. Xu, Z. Li, and X. Feng, "Optimally coordinated dispatch of combined-heat-and-electrical network with demand response," *IET Gener. Transm. Distrib.*, vol. 13, no. 11, pp. 2216–2225, 2019.
[15] Z. Li, L. Wu, and Y. Xu, "Risk-averse coordinated operation of a multi-energy microgrid considering voltage/var control and thermal flow: An adaptive stochastic approach," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 3914–3927, Sep. 2021, doi: 10.1109/TSG.2021.3080312.
[16] S. Lu, W. Gu, S. Zhou, S. Yao, and G. Pan, "Adaptive robust dispatch of integrated energy system considering uncertainties of electricity and outdoor temperature," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4691–4702, Jul. 2020.
[17] S. Lu, W. Gu, K. Meng, and Z. Dong, "Economic dispatch of integrated energy systems with robust thermal comfort management," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 222–233, Jan. 2021.
[18] Y. Zhou, Q. Zhai, and L. Wu, "Multistage transmission-constrained unit commitment with renewable energy and energy storage: implicit and explicit decision methods," *IEEE Trans. Sustain. Energy*, vol. 12, no. 2, pp. 1032–1043, Apr. 2021.
[19] P. Fatouros, I. Konstantelos, D. Papadaskalopoulos, and G. Strbac, "Stochastic dual dynamic programming for operation of der aggregators under multi-dimensional uncertainty," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 459–469, Jan. 2019.
[20] T. Asamov, D. F. Salas, and W. B. Powell, "SDDP vs. ADP: The effect of dimensionality in multistage stochastic optimization for grid-level energy storage," 2016, *arXiv:1605.01521*.
[21] Z. Shi, H. Liang, and V. Dinavahi, "Distributionally robust multi-period energy management for CCHP-based microgrids," *IET Gener. Transm. Distrib.*, vol. 14, no. 19, pp. 4097–4107, 2020.
[22] H. Shuai, J. Fang, X. Ai, Y. Tang, J. Wen, and H. He, "Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2440–2452, May 2019.
[23] D. R. Jiang and W. B. Powell, "An approximate dynamic programming algorithm for monotone value functions," *Oper. Res.*, vol. 63, no. 6, pp. 1489–1511, 2015.
[24] J. Nascimento and W. B. Powell, "An optimal approximate dynamic programming algorithm for concave, scalar storage problems with vector-valued controls," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 2995–3010, Dec. 2013.
[25] Z. Chen, L. Wu, and M. Shahidehpour, "Effective load carrying capability evaluation of renewable energy via stochastic long-term hourly based SCUC," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 188–197, Jan. 2015.
[26] P. Li *et al.*, "MPC-based local voltage control strategy of DGs in active distribution networks," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2911–2921, Oct. 2020.
[27] Z. Li, Y. Xu, X. Feng, and Q. Wu, "Optimal stochastic deployment of heterogeneous energy storage in a residential multi-energy microgrid with demand-side management," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 991–1004, Feb. 2021.
[28] R. Lam, K. Willcox, and D. H. Wolpert, "Bayesian optimization with a finite budget: An approximate dynamic programming approach," *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 883–891.
[29] "Jensens Inequality." Oct. 2021. [Online]. Available: https://en.wikipedia.org/wiki/Jensen%27s_inequality (Accessed: Oct. 2021).
[30] D. F. Salas, "Approximate dynamic programming algorithms for the control of grid-level storage in the presence of renewable generation," Ph.D. dissertation, Dept. Oper. Res. Financial Eng., Princeton Univ., Princeton, NJ, USA, 2014.

Zhengmao Li (Member, IEEE) received the B.E. degree in information engineering and the M.E. degree in electrical engineering from Shandong University, Jinan, China, in 2013 and 2016, respectively, and the Ph.D. degree in electrical engineering from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2020. He was a Research Fellow with the Stevens Institute of Technology, Hoboken, NJ, USA, from 2020 to 2021. His research interests include renewable energy integration, multienergy systems, such as multienergy microgrids, multienergy ships, optimization, and reinforcement learning techniques, such as approximate dynamic programming, robust and stochastic optimization method, and the resilience of multienergy systems.

Lei Wu (Senior Member, IEEE) received the B.S. degree in electrical engineering and the M.S. degree in systems engineering from Xi'an Jiaotong University, Xi'an, China, in 2001 and 2004, respectively, and the Ph.D. degree in electrical engineering from the Illinois Institute of Technology (IIT), Chicago, IL, USA, in 2008.

From 2008 to 2010, he was a Senior Research Associate with the Robert W. Galvin Center for Electricity Innovation, IIT. He was a summer Visiting Faculty with NYISO in 2012. He was a Professor with the Electrical and Computer Engineering Department, Clarkson University, Potsdam, NY, USA, until 2018. He is currently a Professor with the Electrical and Computer Engineering Department, Stevens Institute of Technology, Hoboken, NJ, USA. His research interests include power systems operation and planning, energy economics, and community resilience microgrid.

Yan Xu (Senior Member, IEEE) received the B.E. and M.E. degrees in electrical engineering from the South China University of Technology, Guangzhou, China, in 2008 and 2011, respectively, and the Ph.D. degree in electrical engineering from the University of Newcastle, Callaghan, NSW, Australia, in 2013.

He is currently the Nanyang Associate Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is a Cluster Director of Energy Research Institute. He held the University of Sydney Postdoctoral Fellowship in Australia. His research interests include power system stability and control, microgrid, and data analytics for smart grid applications. He is currently an Editor for the IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE TRANSACTIONS ON SMART GRID, IEEE POWER ENGINEERING LETTERS, *CSEE Journal of Power and Energy Systems*, and an Associate Editor for the *IET Generation, Transmission and Distribution*.

Somayeh Moazeni (Senior Member, IEEE) received the Ph.D. degree in computer science and the Math degree in combinatorics and optimization from the University of Waterloo, Waterloo, ON, Canada. She is currently an Assistant Professor with the School of Business, Stevens Institute of Technology, Hoboken, NJ, USA. Her research interests include continuous optimization, convex optimization, stochastic optimization, machine learning, and their applications in energy, finance, and business analytics.

Zao Tang (Member, IEEE) received the B.S. degree in electrical engineering from Sichuan University, Chengdu, China, in 2016, and Ph.D. degree from the College of Electrical Engineering, Sichuan University in 2021. She was a Visiting Scholar with the Electrical and Computer Engineering Department, Stevens Institute of Technology, USA, from 2018 to 2020. Her current research interests are the optimal planning and operation of energy storage in power system.