Double-Mode Energy Management for Multi-Energy System via Distributed Dynamic Event-Triggered Newton-Raphson Algorithm

Yushuai Li, *Member, IEEE,* David Wenzhong Gao, *Senior Member, IEEE,* Wei Gao *Student Member, IEEE,* Huaguang Zhang, *Fellow, IEEE,* and Jianguo Zhou, *Member, IEEE*

***Abstract*—The islanded and network-connected modes are expected to be modeled into a unified form as well as in a distributed fashion for multi-energy system. In this way, the adaptability and flexibility of multi-energy system can be enhanced. To this aim, this paper establishes a double-mode energy management model for the multi-energy system. It is formed by many energy bodies. With such a model, each participant is able to adaptively respond to the change of mode switching. Furthermore, a novel distributed dynamic event-triggered Newton-Raphson algorithm is proposed to solve the double-mode energy management problem in a fully distributed fashion. In this method, the idea of Newton descent along with the dynamic event-triggered communication strategy are introduced and embedded in the execution of the proposed algorithm. With this effort, each participant can adequately utilize the second-order information to speed up convergence. The optimality is not affected. Meanwhile, the proposed algorithm can be implemented with asynchronous communication and without needing special initialization conditions. It exhibits better flexibility and adaptability especially when the system modes are changed. In addition, the continuous-time algorithm is executed with discrete-time communication driven by the proposed dynamic event-triggered mechanism.It results in reduced communication interaction and avoids needing continuous-time information transmission. It is also proved that each participant can asymptotically converge to the global optimal point. Finally, simulation results show the effectiveness of the proposed model and illustrate the faster convergence feature of the proposed algorithm.**

***Index Terms*—Newton-Raphson, Energy Internet, distributed algorithm, multi-energy systems.**

I. Introduction

With the deep integration of advanced energy and information technologies, the concept of Energy Internet and multi-energy system (MES) are proposed recently. The key objective is to tackle the challenges of integrating multi-energy networks, accommodating high penetration of renewable energy resources and developing hybrid energy utilization model, etc [1-6]. As a fundamental element of MES, the energy management problem (EMP) studies how to cooperatively allocate energy generation resources and schedulable energy loads. In this way, the total social welfare can be maximized under multiple global supply-demand balance and local operation constraints.

A lot of centralized algorithms have been presented to solve the EMP. They mainly include analytical algorithms [7]–[9] and heuristic algorithms [10], [11]. However, the centralized algorithms rely on a powerful centralized controller, which easily undergoes single-point failures, modeling errors [12] and weakened privacy [13], etc. To overcome the drawbacks of traditional centralized algorithms, the distributed algorithm has become a popular and effective methodology to deal with the EMP. A lot of researchers have committed themselves to study the EMP by designing distributed algorithms in smart grid or microgrid. These algorithms mainly include consensus-based and alternating direction method of multipliers (ADMM) based methods. For instance, inspired by consensus protocol, Yand et al [14] first proposed a lambda-iteration approach to solve the economic dispatch problem in a fully distributed manner with quadratic cost function. On this basis, a distributed bisection method and a projected gradient method by making use of average consensus protocol to estimate critical variables were respectively proposed in [15] and [16], which are suitable for non-quadratic cost function. To improve convergence rate, some Newton-based distributed algorithms have been proposed and used in distribution network operation problems [17]–[19]. For instance, a distributed quasi-Newton was proposed to solve multi-area economic dispatch problem in [17]. This method is designed by using the concept of parallel primal-dual interior-point (PDIP) algorithm along with the quasi-Newton technology. Similarly, by decomposing interior-point method and embedding Newton descent procedure, a fully decentralized optimal power flow (OPF) algorithm and an incremental-oriented ADMM were proposed in [18] and [19], respectively. Note that literature [17]–[19] have done outstanding contributions on solving the problem of economic dispatch or OPF for multi-area interconnected power systems. Meanwhile, the Newton concept is utilized to improve convergence speed. However, the designed algorithms are built upon synchronous and periodic communication [18], [19] (or synchronous communication in initial step but needing multi-step communication during each updating [17]). Meanwhile, these algorithms require special initial conditions. It will lead

This work was supported by the U.S. National Science Foundation under Grant 1711951.

Y. Li, D. W. Gao and W. Gao are with the Department of Electrical and Computer Engineering, University of Denver, Denver, CO 80208 USA (e-mail: Wenzhong.Gao@du.edu)

H. Zhang is with State Key Laboratory of Synthetical Automation for Process Industries (Northeastern University), Shenyang, Liaoning, 110004, China. H. Zhang is also with the School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning, 110004, China.

J. Zhou is with the Tsinghua-Berkeley Shenzhen Institute (TBSI), Tsinghua Shenzhen International Graduate School (Tsinghua SIGS), Tsinghua University, 518055 Shenzhen, Guangdong, China.

* Corresponding Author: David Wenzhong Gao.

to two challenges. On one hand, the synchronous communication strategies rely on a global synchronization clock and require all areas to exchange information at the same time. It results in weakened flexibility. Paper [17] only requires synchronous information in initial step. However, each area needs to exchange multiple variables and execute multi-step communication interaction during each updating. Thus, the flexibility is also decreased. On the other hand, the algorithms with special initial procedure may not always ensure the convergence and optimality when the system operation mode or structure is changed. The most reliable method is to reset the initial value for all areas synchronously and re-run system algorithm. It results in weakened adaptability. Most recently, the algorithms' effectiveness under time delay [20], [21], uncertainty of power generation [22], [23], network attacks [24], [25], frequency control [26]–[28], and event-triggered communication [29], [30] etc., have gained wide interests in the field of distributed energy management for power systems.

The above research has made outstanding contributions for the paradigm shift from centralized operations to distributed operations. However, they mainly concentrate on solving the EMP for power system. Therein, only electrical power is seen as the major energy medium. Nevertheless, in MES, multiple energy resources are integrated to achieve the cooperative energy management of different energy networks rather than electrical power only. There are various types of energy mediums (i.e., power, heat and gas) with strong coupling relationship. Thus, the EMP of MES is more complex and difficult than smart grid EMP in both distributed model development and theoretical analysis. To address this issue, the energy body (EB) was recently presented within the context of MES in our previous work [31]. It features multi-coupling of different energy forms, diversified energy roles and peer-to-peer energy supply/demand. Meanwhile, a consensus-ADMM algorithm was proposed to find the optimal energy generations/demands and market clearing prices. In [32] and [33], the concept of two-stage modeling and optimization approaches were proposed to better integrate renewable energy into MES. However, they belong to centralized method. Our paper concentrates on distributed modeling and solutions.

On the basis of [31], an event-triggered based distributed gradient algorithm was further presented in [34]. This algorithm can solve the day-ahead and real-time EMP with less communication cost. Later, in [35], a neurodynamics-based algorithm was proposed to solve the co-planning of coupled power-heat-gas system, which is suitable for large-scale MES. A distributed double-consensus based distributed algorithm was proposed in [36] to achieve the cooperated energy management for multiple We-Energy. The aforementioned research [31], [34]–[36] have obtained some satisfactory results on dealing with the distributed EMP within the context of MES, which operate in either islanded mode or network-connected mode. To improve the adaptability of future MES, it is expected to establish a unified energy management model and strategy. In this way, the islanded and network-connected modes can be simultaneously encompassed together. Additionally, the buying and selling energy prices of the main networks may be different [37], which is also needed to be considered in the distributed system model development. More importantly, the above distributed ADMM or consensus-based algorithms mainly employ the gradient or sub-gradient descent direction. These methods only use first-order information, resulting in slower convergence rate. Note that the convergence rate is a very important index for real-time EMP to ensure effective operation scheduling. Although literature [17]–[19] have introduced the Newton-descent concept to improve convergence rate, they require synchronous and periodic communication as well as special initial conditions. In this paper, we hope that each participant within any EB can flexibly and adaptively operate in islanded mode, network-connected mode and during mode switching. To meet the expectation, further investigation on the design of new Newton-based distributed algorithm is necessary. Herein, the new method should be executed in asynchronous communication fashion, avoid special initialization procedure, and maintain faster convergence speed as well as global optimality. Additionally, although literature [29], [30], [34] have introduced the idea of event-triggered communication strategy into the distributed EMP, the employed triggering mechanisms are static. Further investigation on dynamic triggering mechanism is needed to increase inter-event time.

To overcome these challenges, this paper presents a double-mode energy management model for MES under EB framework [31]. Herein, a distributed dynamic event-triggered Newton-Raphson algorithm is proposed to solve the double-mode EMP in a distributed fashion. The contributions of this paper are summarized as follows:

1) By splitting the aggregated objective function, we regroup the items related to energy price of all participants and equivalently assigned them to the main network (MN) agent. Meanwhile, the corresponding decision variables and objective function are designed for the MN agent based on the global supply-demand constraints. To model the profit-driven characteristic, a correction item is further designed and added into the objective function of MN agent. In this way, it can effectively take the different buying and selling energy prices for the main networks into consideration. Subsequently, a double-mode energy management model is built for the MES with many EBs. It is a fully distributed model, which supports that each participant adaptively runs in islanded mode, network-connected mode and during mode switching.

2) A novel distributed dynamic event-triggered Newton-Raphson algorithm is presented. It is designed by employing primal-dual analysis, Taylor expansion and the differentiated projection operation, while embedding dynamic event-triggered communication strategy. Distinguished from existing gradient-based and Newton-based distributed methods, the proposed algorithm simultaneously possesses the advantages of faster convergence speed, no special initialization condition, and asynchronous communication, etc. With those efforts, each participant can fast obtain its optimal operation and adaptively respond to the model switching without needing to reset global initial value. Meanwhile, the asynchronous communication can avoid the requirement of global clock synchronization, which is more flexible and easier to be implemented. Thus, the proposed algorithm is more suitable for solving the double-

TABLE I
FREQUENTLY USED NOTATION

Notations	Description
p, h, g	Electricity power, heat, gas outputs
l^p, l^h, l^g	Electricity power, heat, gas loads
fbg	Index for the fuel-based generation device referring to the FG or FHD used in variable
rbg	Index for the renewable-based generation device referring to the RG or RHD used in variable
st	Index for the storage device referring to the ES or HS used in variable
chp	Index for CHP used in variable
gs	Index for GS used in variable
mr	Index for must-run energy load used in variable
sl	Index for schedulable energy load used in variable
$C(\bullet)$, $U(\bullet)$	Indices for the cost function and utility function
γ_t^p, γ_t^h, γ_t^g	Indices for the electricity power, heat gas market clearing prices

mode EMP for MES considered in this paper.

3) By designing dynamic triggering mechanism, each participant can exchange information at discrete instants if needed. The direct and major benefits of the proposed event triggered mechanism is to reduce the communication interaction and significantly relax the dependency on precise continuous-time information transmission. Meanwhile, it can also bring some indirect benefits such as the reduced communication expenditure and bandwidth, etc. Moreover, the proposed triggering mechanism involves extra dynamic variable, which holds longer inter-event time and explicitly excludes Zeno behavior.

4) By using Lyapunov technique, it is proved that the proposed algorithm holds asymptotic convergence. Meanwhile, the equilibrium point satisfies the Karush-Kuhn-Tucker (KKT) conditions. Thus, the optimality and convergency of the proposed algorithm is ensured in theory.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this paper, the MES is expected to adaptively work in not only islanded mode but also network-connected mode. In islanded mode, all EBs are cooperative with each other to maximize the total social welfare without being affected by the main energy networks. The main energy networks include electricity grid, heat network and gas network. In network-connected mode, the energy (i.e., electric power, heat and gas) prices of the main energy networks are further taken into consideration to regulate the energy allocation among all EBs. In the considered MES, there is an agent directly connected to the main energy networks, which is named as MN agent. The MN agent can monitor the energy prices of the main energy networks and be required to share information with less neighboring EBs. The proposed distributed algorithm will be discussed in section III. By implementing this algorithm, each participant can find its optimal operation under both islanded and network-connected modes via local communication.

A. Energy Body Model

For each EB, it integrates various kinds of energy resources, including renewable generators (RGs), renewable heating devices (RHDs), fuel-based generators (FGs), fuel-based heating devices (FHDs), combined heat and power (CHP) generators, electricity storages (ESs), heat storages (HSs), gas providers (GSs). Meanwhile, the energy loads in each EB contain electricity loads, heat loads and gas loads. Each of them contains an equivalent must-run energy load and a schedulable energy load (SEL). For simplified definitions of variables and symbols, please refer to Table I for quick reference. The imbalance (deficit or overabundance) power, heat and gas of the ith EB at time step t are respectively given by

$$\Delta p_{i,t} = p_{i,t}^{rbg} + p_{i,t}^{fbg} + p_{i,t}^{chp} + p_{i,t}^{st} - l_{i,t}^{p,mr} - l_{i,t}^{p,sl}, \quad (1)$$

$$\Delta h_{i,t} = h_{i,t}^{rbg} + h_{i,t}^{fbg} + h_{i,t}^{chp} + h_{i,t}^{st} - l_{i,t}^{h,mr} - l_{i,t}^{h,sl}, \quad (2)$$

$$\Delta g_{i,t} = g_{i,t}^{gs} - l_{i,t}^{g,mr} - l_{i,t}^{g,sl}, \quad (3)$$

where $p_{i,t}^{st}$ or $h_{i,t}^{st}$ is positive for discharging and negative for charging.

Under both islanded and network-connected modes, each participant within an EB makes decisions subject to a set of local constraints. In power system, we consider the power capability constraint and ramp rate limits for FG shown in (4) when $\phi = p$ and (8), respectively. The tradeoff constraint between optimality and possibility for RG is shown in (6) when $\phi = p$ (see Remark 1). The constraints for electricity storage are shown in (10-12) when $\phi = p$. The limits for schedulable power load and the corresponding ratios are shown in (13), (16) and (18), respectively. In heat system, we consider the capability constraint of FHD shown in (4) when $\phi = h$. The tradeoff constraint between optimality and possibility for RHD is shown in (6) when $\phi = h$ (see Remark 1). The constraints for HS are shown in (10-12) when $\phi = h$. The limits for heat load and the corresponding ratios are shown in (14), (17) and (18), respectively. In gas system, we consider the capability constraint for the gas provider shown in (5). The limits for gas load and the corresponding ratios are shown in (15), (16) and (17). The CHP is in both power system and heat system, whose local operation constraints are shown in (7) and (9). Moreover, some energy loads can be more flexible to be fed by different energy supplies. The concept is considered and modeled in (16-18). This model of (16) is first proposed for integrated electricity-gas system in [38]. It is further expanded to (16-18) for multiple-energy system in [31]. The mathematical expression of the above constraints are as follows [31]

$$\phi_i^{fbg,min} \le \phi_{i,t}^{fbg} \le \phi_i^{fbg,max}, \quad \phi \in p, h \quad (4)$$

$$0 \le g_i^{gs,min} \le g_{i,t}^{gs} \le g_i^{gs,max}, \quad (5)$$

$$\phi_{i,t}^{rbg,min} \le \phi_{i,t}^{rbg} \le \phi_{i,t}^{rbg,max}, \quad \phi \in p, h \quad (6)$$

$$0 \le \varpi_{i,\kappa}^1 p_{i,t}^{chp} + \varpi_{i,\kappa}^2 h_{i,t}^{chp} + \varpi_{i,\kappa}^3, \quad \kappa = 1, 2, 3, 4 \quad (7)$$

$$-p_i^{fbg,rp} \le p_{i,t}^{fbg} - p_{i,t-1}^{fbg} \le p_i^{fbg,rp}, \quad (8)$$

$$-p_i^{chp,rp} \le p_{i,t}^{chp} - p_{i,t-1}^{chp} \le p_i^{chp,rp}, \quad (9)$$

$$-\phi_i^{ch,max} \le \phi_{i,t}^{st} \le \phi_i^{ds,max}, \quad \phi \in p, h \quad (10)$$

$$SOC_{i,t}^{\phi} = SOC_{i,t-1}^{\phi} - \Im_i^{\phi} \phi_{i,t-1}^{st} \Delta t, \quad \phi \in p, h \quad (11)$$

$$SOC_i^{\phi,min} \le SOC_{i,t}^{\phi} \le SOC_i^{\phi,max}, \quad \phi \in p, h \quad (12)$$

$$0 \le l_{i,t}^{p,sl} \le l_{i,t}^{p,max} - l_{i,t}^{p,mr}, \tag{13}$$
$$0 \le l_{i,t}^{h,sl} \le l_{i,t}^{h,max} - l_{i,t}^{h,mr}, \tag{14}$$
$$0 \le l_{i,t}^{g,sl} \le l_{i,t}^{g,max} - l_{i,t}^{g,mr}, \tag{15}$$
$$\Upsilon_{i,g\to p}^{min} \le l_{i,t}^{p,sl}/(l_{i,t}^{p,sl} + \Psi l_{i,t}^{g,sl}) \le \Upsilon_{i,g\to p}^{max}, \tag{16}$$
$$\Upsilon_{i,g\to h}^{min} \le l_{i,t}^{h,sl}/(l_{i,t}^{h,sl} + \Psi l_{i,t}^{g,sl}) \le \Upsilon_{i,g\to h}^{max}, \tag{17}$$
$$\Upsilon_{i,h\to p}^{min} \le l_{i,t}^{p,sl}/(l_{i,t}^{p,sl} + l_{i,t}^{h,sl}) \le \Upsilon_{i,h\to p}^{min}, \tag{18}$$

where superscript "min" and "max" refer to the lower and upper bounds, respectively; $\varpi_{i,\kappa}^1$, $\varpi_{i,\kappa}^2$ and $\varpi_{i,\kappa}^3$ are the coefficients of κth linear inequality constraint determined by the feasible operation region of CHP i; $p_i^{fbg,rp}$ and $p_i^{chp,rp}$ are the ramp rates of $p_{i,t}^{fbg}$ and $p_{i,t}^{chp}$ between two consecutive time steps; $\phi_i^{ch,max}$ and $\phi_i^{ds,max}$ are the maximum charging and discharging rates; $\Im_i^{\phi}$ represents the charge and discharge efficiency; $SOC_{i,t}^{\phi}$ represents the state of charge (energy stored in the corresponding storage device); $\Upsilon_{i,g\to p}$, $\Upsilon_{i,g\to h}$ and $\Upsilon_{i,h\to p}$ represent the ratio of electric power with reference to combined power and gas load, the ratio of heat with reference to combined heat and gas load, the ratio of electric power with reference to combined power and gas load, respectively; Ψ represents the transformation ratio from SCM/h to MW, which is often set as $1/84$ [31].

Driven by profit, each EB can simultaneously act as the energy supplier and consumer. For instance, the EB may purchase deficit power from other EBs or main grid if its power demand cannot be satisfied. Under this circumstance, it is viewed as a power consumer. Meanwhile, the EB may also sell the overabundance heat to other EBs or main heat network to make additional profit. Under this circumstance, it is regarded as a heat supplier. At time step t, the objective function of each EB composed of the profit part and the cost part is expressed as [31]

$$\begin{aligned} F_{i,t} = &- \sum_{\phi\in p,h} \big(C(\phi_{i,t}^{fbg}) + C(\phi_{i,t}^{rbg}) + C(\phi_{i,t}^{st})\big) - C(g_{i,t}^{gs}) \\ &- C(p_{i,t}^{chp}, h_{i,t}^{chp}) + U(l_{i,t}^{p,sl}, l_{i,t}^{h,sl}, l_{i,t}^{g,sl}) + \gamma_t^p \Delta p_{i,t} \\ &+ \gamma_t^h \Delta h_{i,t} + \gamma_t^g \Delta g_{i,t}, \qquad \phi \in p,h \end{aligned} \tag{19}$$

with

$$\begin{aligned} C(\phi_{i,t}^{fbg}) &= a_i^{\phi,fbg}(\phi_{i,t}^{fbg})^2 + b_i^{\phi,fbg}\phi_{i,t}^{fbg} + c_i^{\phi,fbg} \\ &\quad + \varepsilon_i^{\phi,fbg}\exp(\xi_i^{\phi,fbg}\phi_{i,t}^{fbg}), \end{aligned}$$

$$C(\phi_{i,t}^{rbg}) = a_i^{\phi,rbg}\phi_{i,t}^{rbg} + \epsilon_i^{\phi,rbg}\exp\big(\xi_i^{\phi,rbg}\frac{\phi_{i,t}^{rbg,max} - \phi_{i,t}^{rbg}}{\phi_{i,t}^{rbg,max} - \phi_{i,t}^{rbg,min}}\big),$$

$$C(\phi_{i,t}^{st}) = a_i^{\phi,st}(\phi_{i,t}^{st} + b_i^{\phi,st})^2,$$

$$C(g_{i,t}^{gs}) = a_i^{gs}(g_{i,t}^{gs})^3 + b_i^{gs}(g_{i,t}^{gs})^2 + d_i^{gs}g_{i,t}^{gs} + c_i^{gs},$$

$$\begin{aligned} C(p_{i,t}^{chp}, h_{i,t}^{chp}) &= a_i^{chp}(p_{i,t}^{chp})^2 + b_i^{chp}p_{i,t}^{chp} + d_i^{chp}p_{i,t}^{chp}h_{i,t}^{chp} \\ &\quad + e_i^{chp}(h_{i,t}^{chp})^2 + f_i^{chp}h_{i,t}^{chp} + c_i^{chp}, \end{aligned}$$

$$\begin{aligned} U(l_{i,t}^{p,sl}, l_{i,t}^{h,sl}, l_{i,t}^{g,sl}) &= -\alpha_i^p(l_{i,t}^{p,mr} + l_{i,t}^{p,sl})^2 + \beta_i^p(l_{i,t}^{p,mr} + l_{i,t}^{p,sl}) \\ &\quad - \alpha_i^h(l_{i,t}^{h,mr} + l_{i,t}^{h,sl})^2 + \beta_i^h(l_{i,t}^{h,mr} + l_{i,t}^{h,sl}) \\ &\quad - \alpha_i^g(l_{i,t}^{g,mr} + l_{i,t}^{g,sl})^2 + \beta_i^g(l_{i,t}^{g,mr} + l_{i,t}^{g,sl}), \end{aligned}$$

where $a_i^{\phi,fbg}$, $b_i^{\phi,fbg}$, $c_i^{\phi,fbg}$, $\varepsilon_i^{\phi,fbg}$, $\xi_i^{\phi,fbg}$, $a_i^{\phi,rbg}$, $\epsilon_i^{\phi,rbg}$, $a_i^{\phi,st}$, $b_i^{\phi,st}$, a_i^{gs}, b_i^{gs}, c_i^{gs}, d_i^{gs}, a_i^{chp}, b_i^{chp}, c_i^{chp}, d_i^{chp}, e_i^{chp} and f_i^{chp} are cost coefficients; $\xi_i^{\phi,rbg} < 0$ is penalty coefficient; α_i^p, β_i^p, α_i^h, β_i^h, α_i^g and β_i^g are utility coefficients. Note that $C(g_{i,t}^{gs})$ is convex in the region determined by (5).

Remark 1: The models of RG and RHD with the consideration of uncertainties come from [31]. To be specific, it is assumed that the forecast error of the power generation of RG (or heat generation of RHD) caused by uncertainties obeys Gaussian distribution. On this basis, the confidence intervals of the forecasted power or heat generation can be further obtained based on the preset confidence level of the forecast error. Then, we get constraint (6) by letting the forecasting value plus the corresponding forecast error to obtain $\phi_{i,t}^{rbg,min}$ and $\phi_{i,t}^{rbg,max}$, which takes the uncertainty into account. Moreover, the exponential penalty of renewable power or heat generation is considered in the cost function $C(\phi_{i,t}^{rbg})$, i.e., the second item of $C(\phi_{i,t}^{rbg})$. The purpose is to make the optimality and possibility be traded off suitably. By using this model, it can be derived that the Hessian matrix of $C(\phi_{i,t}^{rbg})$, i.e., $\epsilon_i^{\phi,rbg}\big(\xi_i^{\phi,rbg}/(\phi_{i,t}^{rbg,max} - \phi_{i,t}^{rbg,min})\big)^2\exp\big(\xi_i^{\phi,rbg}(\phi_{i,t}^{rbg,max} - \phi_{i,t}^{rbg})/(\phi_{i,t}^{rbg,max} - \phi_{i,t}^{rbg,min})\big)$ is positive definite within local constraint (6). Thus, $C(\phi_{i,t}^{rbg})$ is convex within constraint (6).

B. Double-Mode Energy Management of MES

Considering a MES with n EBs, the objective of the optimal energy management is to maximize the total social welfare under both islanded and network-connected modes. The mathematical expression is given by

$$\max \quad Obj = \sum_{i=1}^{n} F_{i,t}. \tag{20}$$

Although the islanded and network-connected modes hold the same objective function, they have different forms. This is caused by the different energy supply-demand balance constraints and energy trading mechanism. To show it clearly, $F_{i,t}$ is divided into two terms as follows

$$F_{i,t} = F_{i,t}^{one} + F_{i,t}^{two}, \tag{21}$$

where

$$\begin{aligned} F_{i,t}^{one} = &- \sum_{\phi\in p,h} \big(C(\phi_{i,t}^{fbg}) + C(\phi_{i,t}^{rbg}) + C(\phi_{i,t}^{st})\big) - C(g_{i,t}^{gs}) \\ &- C(p_{i,t}^{chp}, h_{i,t}^{chp}) + U(l_{i,t}^{p,sl}, l_{i,t}^{h,sl}, l_{i,t}^{g,sl}), \end{aligned} \tag{22}$$

$$F_{i,t}^{two} = \gamma_t^p \Delta p_{i,t} + \gamma_t^h \Delta h_{i,t} + \gamma_t^g \Delta g_{i,t}. \tag{23}$$

The first term, i.e., $F_{i,t}^{one}$, is the same for both islanded mode and network-connected mode. The difference is mainly reflected in the second term, i.e., $F_{i,t}^{two}$. To be specific, the global energy supply-demand constraints for islanded mode (24) and network-connected mode (25) at each time step t are respectively given by

$$\sum_{i=1}^{n} \Delta p_{i,t} = 0, \quad \sum_{i=1}^{n} \Delta h_{i,t} = 0, \quad \sum_{i=1}^{n} \Delta g_{i,t} = 0, \tag{24}$$

$$\sum_{i=1}^{n} \Delta p_{i,t} = p_t^M, \ \sum_{i=1}^{n} \Delta h_{i,t} = h_t^M, \ \sum_{i=1}^{n} \Delta g_{i,t} = g_t^M, \tag{25}$$

where p_t^M, h_t^M and g_t^M are the exchanged power, heat and gas between the main networks and EBs. Therein, if the main networks inject power, heat or gas into the EBs, then we let $p_t^M < 0$, $h_t^M < 0$ or $g_t^M < 0$. Otherwise, $p_t^M \geq 0$, $h_t^M \geq 0$ or $g_t^M \geq 0$. In addition, we let

$$|p_t^M| \leq p_t^{M,max},\ |h_t^M| \leq h_t^{M,max},\ |g_t^M| \leq g_t^{M,max}, \quad (26)$$

where $p_t^{M,max}$, $h_t^{M,max}$ and $g_t^{M,max}$ represent the maximum values of p_t^M, h_t^M or g_t^M, respectively.

According to (24), we can get $F_{i,t}^{two} = 0$, which also means that $Obj = \sum_{i=1}^{n} F_{i,t}^{one}$ in islanded mode. Different from islanded mode, $F_{i,t}^{two}$ is not always equal to zero in network-connected mode. Meanwhile, γ_t^p, γ_t^h and γ_t^g are the energy prices set by the main networks. It can be observed from (1-3) and (23) that the energy prices are included in the objective function of individual participant. In distributed network, each participant may not have direct access to γ_t^p, γ_t^h and γ_t^g. This is because each participant only exchanges information with its neighbors. To achieve fully distributed computation, we make proper transformation for $F_{i,t}^{two}$. From (23) and (25), it can be derived that

$$F_T^{pr} = \sum_{i=1}^{n} F_{i,t}^{two} = \gamma_t^p p_t^M + \gamma_t^h h_t^M + \gamma_t^g g_t^M. \quad (27)$$

Thus, we can make p_t^M, h_t^M and g_t^M be seen as decision variables in the main networks side and assign a local objective function, i.e., F_T^{pr} for the MN agent. In this way, the energy prices are reflected only in F_T^{pr}. Thus, we can eliminate the requirement that each participant should access the main network energy prices, which is conducive to distributed implementation.

Moreover, the future MES shall operate in peer-to-peer energy supply/demand structure. The monopoly of the main networks will be weakened and one of the major profit modes is to charge service fees. In this situation, the buying energy prices and selling prices of the main networks may be different. To be specific, the main networks tend to either buy energy from the EBs if the clearing energy prices are cheaper than $b\gamma_t^p$, $b\gamma_t^h$ and $b\gamma_t^g$, or sell energy to the EBs if the clearing energy prices are more expensive than $b\gamma_t^p$, $b\gamma_t^h$ and $b\gamma_t^g$. Therein, $b\gamma_t^p$, $b\gamma_t^h$ and $b\gamma_t^g$ are named as benchmark energy prices. It is worth noting that we do not know the final trading status before calculation. To capture this scenario, a corrected objective function is designed based on the optimality conditions and the characteristics of p_t^M, h_t^M and g_t^M. The mathematical expression is given by

$$\begin{aligned} F_t^{Main} =& b\gamma_t^p p_t^M + b\gamma_t^h h_t^M + b\gamma_t^g g_t^M \\ & - a_p^M (p_t^M)^2 - a_h^M (h_t^M)^2 - a_g^M (g_t^M)^2, \end{aligned} \quad (28)$$

where a_p^M, a_h^M and a_g^M are positive correction coefficients (The detailed design of correction coefficients are illustrated in Appendix A.). Next, we take the electric power price as an example to show this concept. Based on (28), the final power market clearing price in optimal state (without consideration of (26) for example) is $b\gamma_t^p - 2a_p^M p_t^M$. Note that $p_t^M > 0$ if the main grid purchases power from the EBs. In this case, $b\gamma_t^p - 2a_p^M p_t^M < b\gamma_t^p$, which implies that the main grid will buy cheaper power from EBs. Otherwise, $p_t^M < 0$ and the main grid will sell more expensive power to the EBs. In addition, it is worth noting that the buying and selling prices are the same one when $a_p^M = a_h^M = a_g^M = 0$. In this scenario, F_t^{Main} is the same as F_t^{pr}. Thus, F_t^{pr} can be seen as a special case of F_t^{Main}.

Based on the aforementioned discussion, the double-mode EMP for MES is rewritten as

$$\max \quad Obj = \sum_{i=1}^{n} F_{i,t}^{one} + \pi F_t^{Main}, \quad (29)$$

subject to (4-18, 24) for islanded mode and (4-18, 25-26) for network-connected mode. Therein, $\pi = 0$ or 1 represents islanded mode or network-connected mode, respectively.

III. Transformation and Distributed Algorithm

We consider a MES with n EBs and one MN agent, in which each EB has $\wp_i$ participants (i.e., energy devices or schedulable energy loads). To simplify notations, at time t, we make use of $\{x_{ij} \in \mathrm{R}^3 | i = 1, \cdots, n; j = 1, \cdots, \wp_i\}$ to represent the decision variables of jth participant of ith energy body. $\{x_{ij} \in \mathrm{R}^3 | i = 0; j = 1\}$ is denoted as the decision variable of MN agent. $\{d_{ij} \in \mathrm{R}^3 | i = 1, \cdots, n; j = 1, \cdots, \wp_i\}$ is denoted as the must-run energy loads connected to jth participant of ith energy body. Note that x_{ij} is a three-dimensional decision variable composed of power, heat and gas, some of which may be zero(s). d_{ij} is also a three dimensional vector composed of power, heat and gas loads. If some participant does not have all the three energy variables, it can extend the dimension of its variable(s) to x_{ij} by setting the bounds of zero variable(s) as zeros. Meanwhile, the cost function of each zero variable is set to any kind of strongly convex function like $C(\phi_{i,t}^{fbg})$. Then, we use $W(x_{ij})$ to re-denote the corresponding cost function or negative utility function or $-F_t^{Main}$. Then, the double-mode EMP for MES, i.e., (29), can be further expressed as the following form,

$$\min \quad Obj = \sum_{i=0}^{n} \sum_{j=1}^{\wp_i} W(x_{ij}) \quad (30)$$

$$s.t. \quad \sum_{i=0}^{n} \sum_{j=1}^{\wp_i} B_{ij} x_{ij} = \sum_{i=0}^{n} \sum_{j=1}^{\wp_i} d_{ij}, \quad (31)$$

$$g(x_{ij}) \leq 0 \rightarrow x_{ij} \in \Omega_{ij}, \quad (32)$$

where $B_{ij} = -I_3$ if x_{ij} represents schedulable energy load, otherwise $B_{ij} = I_3$; $g(x_{ij})$ is the local inequality constraint related to x_{ij}; I_3 represents a three-dimensional identify matrix; Ω_{ij} is the local closed convex set determined by $g(x_{ij})$. In the aforementioned EMP, the MN agent can also be seen as a special EB with single participant, i.e., $\wp_0 = 1$. When the MN agent takes part in the EMP, the system is running in network-connected mode; otherwise, the system is running in islanded mode.

Some preliminary knowledge are presented in Appendix B, including strongly convex property of the aforementioned EMP, the differentiated projection operation and its relevant properties. These will be employed for the convergence analysis of the subsequent distributed algorithm.

A. Distributed Communication Network Topology

The communication network topology among EBs and MN agent can be modeled as a graph $G = (\mathcal{V}, \mathcal{E}, \mathcal{A})$. Therein, $\mathcal{V} = \{v_{ij} | i = \aleph, \cdots, n; j = 1, \cdots, \wp_i\}$ is a set of nodes representing participants or MN agent. $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is a set of edges representing the communication links. $\mathcal{A} = [a_{ij,\overline{ij}}]$ is the adjacency matrix. Therein, $\aleph = 0$ (or 1) refers to network-connected mode (or islanded mode indicating that the MN agent is not considered). Node $v_{\overline{ij}}$ is called the neighbor node of v_{ij} if v_{ij} can get the information from $v_{\overline{ij}}$. Then, we have $(v_{ij}, v_{\overline{ij}}) \in \mathcal{E}$ with $a_{ij,\overline{ij}} = 1$; otherwise $a_{ij,\overline{ij}} = 0$. Meanwhile, the neighbor set of v_{ij} is denoted as $\mathcal{N}_{ij}$ with degree $|\mathcal{N}_{ij}|$. Currently, there are two ways to design the communication links among nodes. One is to make communication network overlay the physical network. In this way, the communication links are aligned with the physical transmission lines. Thus, the neighbors of each node in communication network are the same as the ones in physical structure. The second one is not limited by the physical structure. In this way, each node can exchange information with any other one as needed. In this paper, we take multiple types of energy resources into consideration. The physical links may not be available among some of them, but any needed communication links can be achieved. Thus, we employ the second way to design communication structure in this paper. This implies that the definition of "neighbors" in communication network may be different from the one in physical structure. To avoid confusion, the "neighbors" used in this paper refers to the definition of "neighbors" in communication network.

When the possible expansion of the energy (power, heat and gas) transaction restrictions are taken into consideration, there are two feasible methods to design the corresponding communication network. For the first one, we can let the communication network mimic the physical network. This method possesses good compatibility when considering the energy (power, heat and gas) transaction restrictions directly. However, it may also reduce the flexibility. For the other one, the communication network can be designed to be independent of the physical topology. This kind of method has been widely employed in distributed networks, such as [14-15, 20-21, 34, 37]. It implies that a suitable communication network can be designed in light of cost, location, convenience and technology, etc. Note that if the physical transaction restrictions can be modeled and translated as the local inequality constraints for each agent, we can employ the second one. Since the transaction restrictions have been included in the local operation constraints, each participant can obtain the feasible solution. Otherwise, it is better to employ the first one. Note that no matter what kind of communication network is used, each participant needs to collect its local and neighboring information. More importantly, this paper only requires the communication graph to be connected. Thus, the proposed algorithm can be implemented in any kind of connected communication network architectures. This means that the communication network architecture can be chosen based on actual condition and different application scenarios.

The system Laplacian matrix of G is defined as $L = D - \mathcal{A}$, where $D = \text{diag}\{|\mathcal{N}_{ij}|\}$ is called the degree matrix. In this paper, it is assumed that G is a connected graph. Then, L has one simple zero eigenvalue and positive eigenvalues. The eigenvalues of L are ordered as $0 = \lambda_1 < \lambda_2 \leq, \cdots, \leq \lambda_N$, where N is the total number of nodes.

B. Main Algorithm

In this section, we would like to solve the problems (30-32) with faster convergence rate, reduced communication, asynchronous implementation and without special initialization condition. To this end, we focus on embedding the Newton descent concept along with the dynamic event-triggered communication strategy into the design of the distributed optimization algorithm. By primal-dual analysis, Taylor expansion and differentiated projection operation, the distributed dynamic event-triggered Newton-Raphson algorithm is proposed as follows

$$\begin{aligned}\dot{x}_{ij} =&\Gamma_{\Omega_{ij}}\Big(\nabla^2 W(x_{ij})^{-1}x_{ij}, \nabla^2 W(x_{ij})^{-1}\big(-\nabla W(x_{ij}) \\ &+ B_{ij}^T\hat{y}_{ij}\big)\Big),\end{aligned} \tag{33}$$

$$\begin{aligned}\dot{y}_{ij} =&-\sum_{\overline{ij}\in\mathcal{N}_{ij}} a_{ij,\overline{ij}}(\hat{y}_{ij} - \hat{y}_{\overline{ij}}) \\ &-\sum_{\overline{ij}\in\mathcal{N}_{ij}} a_{ij,\overline{ij}}(\hat{z}_{ij} - \hat{z}_{\overline{ij}}) + d_{ij} - B_{ij}x_{ij},\end{aligned} \tag{34}$$

$$\dot{z}_{ij} = \sum_{\overline{ij}\in\mathcal{N}_{ij}} a_{ij,\overline{ij}}(\hat{y}_{ij} - \hat{y}_{\overline{ij}}), \tag{35}$$

$$\begin{aligned}\dot{\mathcal{Q}}_{ij} =&-\vartheta_{1,ij}\mathcal{Q}_{ij} + \vartheta_{2,ij}\big(\frac{1}{4}\sum_{\overline{ij}\in\mathcal{N}_{ij}} a_{ij,\overline{ij}}||\hat{y}_{ij} - \hat{y}_{\overline{ij}}||^2 \\ &-\vartheta_{3,ij}||(\hat{y}_{ij} - y_{ij})||^2 - \vartheta_{4,ij}||(\hat{z}_{ij} - z_{ij})||^2\big),\end{aligned} \tag{36}$$

with triggering mechanism

$$\begin{aligned}t_{ij}^{k+1} =&\max\Big\{t > t_{ij}^k | \vartheta_{5,ij}\big(\vartheta_{3,ij}||(\hat{y}_{ij} - y_{ij})||^2 \\ &+\vartheta_{4,ij}||(\hat{z}_{ij} - z_{ij})||^2 \\ &-\frac{1}{4}\sum_{\overline{ij}\in\mathcal{N}_{ij}} a_{ij,\overline{ij}}||\hat{y}_{ij} - \hat{y}_{\overline{ij}}||^2\big) \leq \mathcal{Q}_{ij}\Big\},\end{aligned} \tag{37}$$

where $\vartheta_{1,ij}$, $\vartheta_{2,ij}$, $\vartheta_{3,ij}$, $\vartheta_{4,ij}$ and $\vartheta_{5,ij}$ are positive parameters; $\Gamma_{\Omega_{ij}}$ represents the differentiated projection operator, whose detailed definition can be seen in Appendix B; $\overline{ij} \in \mathcal{N}_{ij}$ is the neighbor of node ij; $\hat{y}_{ij} \equiv y_{ij}(t_{ij}^k)$, $\hat{z}_{ij} \equiv z_{ij}(t_{ij}^k)$ and $t \in [t_{ij}^k, t_{ij}^{k+1})$; $\{t_{ij}^k\}_{k=1}^{\infty}$ and $\{t_{ij}^{k+1} - t_{ij}^k\}_{k=1}^{\infty}$ refer to the triggering times and inter-event time intervals; $y_{ij} \in \mathrm{R}^3$ and $z_{ij} \in \mathrm{R}^3$ are designed auxiliary variables. Therein, y_{ij} is the Lagrangian dual multiplier with respect to (31), whose physical significance is the energy market clearing prices; $\mathcal{Q}_{ij}$ is the designed extra dynamic state to achieve dynamic event-triggering communication. In addition, there is no initialization requirement for the Lagrange dual multipliers and primal variables. Thus, choosing different initial values does not affect the convergence of the algorithm.

In the proposed algorithm, formulation (33) is used to update primal variable x_{ij}. $\nabla^2 W(x_{ij})^{-1}\big(-\nabla W(x_{ij})+B_{ij}^T\hat{y}_{ij}\big)$ is obtained by calculating the Newton descent direction of x_{ij} related to the dual problem of (30) with (31). To further deal with the inequality constraint (32), we employ differentiated projection operator (see Appendix B). In this way, the infeasible point can be projected into the corresponding feasible region. In addition, the second-order information is embedded into (33) to speed up the convergence, which will be discussed in Remark 2 for details. Formulations (34) and (35) are designed to estimate the global Lagrangian dual multiplier. The major purpose is to make all the local estimated Lagrangian dual multipliers y_{ij}'s converge to the same value, while meeting the global supply-demand balance, i.e., (31). In this process, the consensus protocol (i.e., $-\sum_{\overline{ij}\in\mathcal{N}_{ij}} a_{ij,\overline{ij}}(\hat{y}_{ij}-\hat{y}_{\overline{ij}})$ and $-\sum_{\overline{ij}\in\mathcal{N}_{ij}} a_{ij,\overline{ij}}(\hat{z}_{ij}-\hat{z}_{\overline{ij}})$) is employed to achieve distributed communication and calculation. The item $d_{ij}-B_{ij}x_{ij}$ is designed to ensure that the global supply-demand balance can be achieved when the algorithm converges. Moreover, to achieve asynchronous communication and reduce communication interaction among participants and MN agent, we further design the extra dynamics (36) and triggering mechanism (37). Dynamics (36) generates a positive variable $\mathcal{Q}_{ij}$. It is further used as a threshold value in judgement condition (37). Based on the information at current triggering time instant, the next triggering time t_{ij}^{k+1} is determined by the designed triggering mechanism. An event is triggered once the inequality in (37) is not satisfied. The major functionality of (36) and (37) is to make each participant or MN agent share the information of variables $\hat{y}_{ij}$ and $\hat{z}_{ij}$ with its neighbors only at some intermittent instants, i.e., triggering times. The proposed event-triggering strategy possesses two major advantages. On one hand, the continuous-time algorithm is implemented by using discrete-time communication in this paper. As a result, the communication interaction can be greatly reduced. On the other hand, each participant or MN agent can independently decide the next triggering time, which results in the ability of asynchronous communication. In other word, the proposed algorithm can be implemented in asynchronous communication fashion without a global clock synchronization. As a consequence, it is more feasible and convenient for applications. Based on the aforementioned discussion, for each participant or MN agent, the updating of x_{ij}, y_{ij}, z_{ij} and $\mathcal{Q}_{ij}$ in (33-36) only involves neighbors' information, i.e., $\hat{y}_{\overline{ij}}$ and $\hat{z}_{\overline{ij}}$, and its own information. Thus, the proposed algorithm can be implemented in a distributed fashion. In addition, each participant does not need to switch algorithm or strategy when the system operation mode is changed. In this way, the designed algorithm is adaptive and suitable for not only islanded mode but also network-connected mode.

Remark 2: It is worth noting that the proposed algorithm can effectively employ Newton descent direction by making use of both first- and second-order information, i.e., $\nabla W(x_{ij})$ and $\nabla^2 W(x_{ij})$. As a consequence, the convergence rate can be sped up, which is one of the most important advantages when compared with (sub)-gradient-based algorithms. Additionally, the computation complexity of $\nabla^2 W(x_{ij})$ is not very big, because x_{ij} is only three-dimensional. Here, we focus on the comparison of computational burden for Newton descent method and for gradient descent method. Let ϱ represent the dimensionality of x_{ij}. The computational complexity of the proposed Newton method and the gradient method in [34] are $O(\varrho^3)$ and $O(\varrho^2)$, respectively. The proposed algorithm undergoes heavy computational burden only when ϱ becomes very large. Note that $\varrho=3$ for each individual participant in our considered system model. Thus, the computational burden of the proposed distributed algorithm when used in this paper is not very big, although it is higher than the gradient method in [34]. But the overall computation time needed to achieve optimal solution is much shorter with our proposed algorithm due to the faster convergence rate. As we know, for each participant or MN agent, they are expected to calculate the global optimal solutions as fast as possible to guarantee effective energy allocation. Thus, the proposed algorithm is more suitable for solving the EMP considered in this paper.

Remark 3: The event-triggering based energy management strategies have been studied in [29], [30], [34]. Different from those work, this paper aims at designing a dynamic triggering mechanism, i.e., (37), which takes advantages of the extra dynamic state $\mathcal{Q}_{ij}$ in (36) to increase inter-event time intervals. If $\mathcal{Q}_{ij}$ is not involved, then the dynamic triggering mechanism will reduce to the static one. It can be viewed as a limit case of (37). Since $\mathcal{Q}_{ij}>0$ (seen from (62) in Appendix C), the dynamic triggering mechanism processes larger inter-event time intervals than the static ones. In addition, we focus on designing Newton descent method to improve the convergence rate, while literature [29], [30], [34] are based on gradient descent information.

Remark 4: Note that (19) is the objective function of each EB. Therein, the exponential items exist only in the cost functions $C(\phi_{i,t}^{fbg})$ and $C(\phi_{i,t}^{rbg})$. The local constrains for $\phi_{i,t}^{fbg}$ and $\phi_{i,t}^{rbg}$, i.e., (4) and (6), are linear constraints. Unlike the strong conic constraint in [39], (26) is a special conic constraint. It is equivalent to $-p_t^{M,max}\le p_t^M\le p_t^{M,max}$, $-h_t^{M,max}\le h_t^M\le h_t^{M,max}$ and $-g_t^{M,max}\le g_t^M\le g_t^{M,max}$. Thus, in our considered system model shown in Section-II, it can avoid involving conic constraints. Moreover, it should be pointed out that the proposed algorithm can be expanded and used to solve general optimization problems with the common form shown in (30-32). Therein, $g(x_{ij})$ can be chosen as conic constraint or any other kinds of inequality constraints as long as Ω_{ij} is a closed convex set. Note that the physical limits of power, heat or gas transactions can be modeled and accounted for with some of the inequality constraints. In this case, we need to model the corresponding cone when calculating the differentiated projection.

Next, the following Theorem 1 is developed to show that the proposed algorithm (33-36) with triggering mechanism (37) can asymptotically converge to the Karush-Kuhn-Tucker (KKT) point of problem (30-32). Namely, each participant and MN agent can locally obtain its optimal operation by implementing the proposed algorithm. Thus, the studied double-mode EMP can be solved.

Theorem 1: Suppose that the communication graph G is

connected. We choose $\vartheta_{1,ij} > 0$, $0 < \vartheta_{2,ij} < 1$, $\vartheta_{3,ij} = \frac{1}{2\emptyset_{ij}} + 5|\mathcal{N}_{ij}| + 4\pounds_4|\mathcal{N}_{ij}|$, $\vartheta_{4,ij} = 5\pounds_4|\mathcal{N}_{ij}| + 4|\mathcal{N}_{ij}|$, $\vartheta_{5,ij} > \frac{1-\vartheta_{2,ij}}{\vartheta_{1,ij}}$, where $0 < \frac{\lambda_2}{3\lambda_2-4\pounds_3} < \pounds_4 < \frac{3\lambda_2}{\lambda_2+4\pounds_3}$, $0 < \pounds_3 < \min\{\frac{1}{2\lambda_2}, \frac{\sqrt{\lambda_2^2\emptyset_{ij}^2+48\emptyset_{ij}\lambda_2}-\emptyset_{ij}\lambda_2}{8\emptyset_{ij}}\}$ and $\emptyset_{ij}$ is the strongly convex coefficient obtained from (38) in Appendix B. Then, given the first triggering time $t_{ij}^1 = t0$ and $\mathcal{Q}_{ij}(t0) > 0$, the implementation of algorithm (33-36) with triggering mechanism (37) is asymptotically convergent. Meanwhile, the equilibrium point of dynamics (33-35) satisfies the optimality (KKT) conditions for problem (30-32).

The proof of Theorem 1 is presented in Appendix C. On the basis of Theorem 1, we also provide the theoretical analysis for the lower bound between two adjacent triggering time, i.e., $t_{ij}^{k+1} - t_{ij}^k$. The theoretical analysis is built upon the condition of guaranteed global optimality. The theoretical result is presented in Appendix D.

The proposed algorithm possesses better robustness property against the change of the system model parameters and algorithm parameters. On one hand, we only let $\mathcal{Q}_{ij}(t0) > 0$ in the proposed algorithm. It is also proved that $\mathcal{Q}_{ij} > 0$ after $t0$ (seen from (62) in Appendix C). More importantly, x_{ij}, y_{ij} and z_{ij} are free of initialization. Thus, when some system model parameters change (e.g., the change of $W(x_{ij})$ and d_{ij}), each participant only needs to independently adjust local data and projects its local generation or consumption onto its feasible region, and then to reach a new equilibrium. In this process, the execution of the algorithm is continuous, and we do not need to stop the dynamics, reset the initial value, and then re-run the dynamics. This implies that the proposed algorithm responds automatically to the changes in system model parameters and converges to new equilibrium with better robustness as well as adaptation properties. On the other hand, it can be seen from the results in Theorem 1 that the proposed algorithm can converge to the global optimal solutions if parameters $\vartheta_{1,ij}$, $\vartheta_{2,ij}$, $\vartheta_{3,ij}$, $\vartheta_{4,ij}$ and $\vartheta_{5,ij}$ meet the designed requirements. This implies that the proposed algorithm can be robust against the bounded perturbations in regard to $\vartheta_{1,ij}$, $\vartheta_{2,ij}$, $\vartheta_{3,ij}$, $\vartheta_{4,ij}$ and $\vartheta_{5,ij}$ if they are within the limits as shown in Theorem 1.

C. Benefits of Distributed Computation Implementation

The major differences in computational implementation between Newton-based centralized methods (e.g, [8], [9]) and the proposed Newton-based distributed method are summarized as follows. For centralized method, all the computation tasks are implemented in a centralized controller. The centralized controller needs to collect all the information (including the objective functions and constraints) of all participants and formulate the global optimization problem. In general, the individual participant is unwilling to reveal its private information to external centralized controller. Thus, the centralized computation suffers from weakened privacy. Next, the centralized controller computes the (approximated) Hessian matrix of the whole system. Then, the centralized controller determines the feasible search direction and step size for all primal and dual variables. It is worth noting that the centralized controller suffers from a huge computing burden as the system

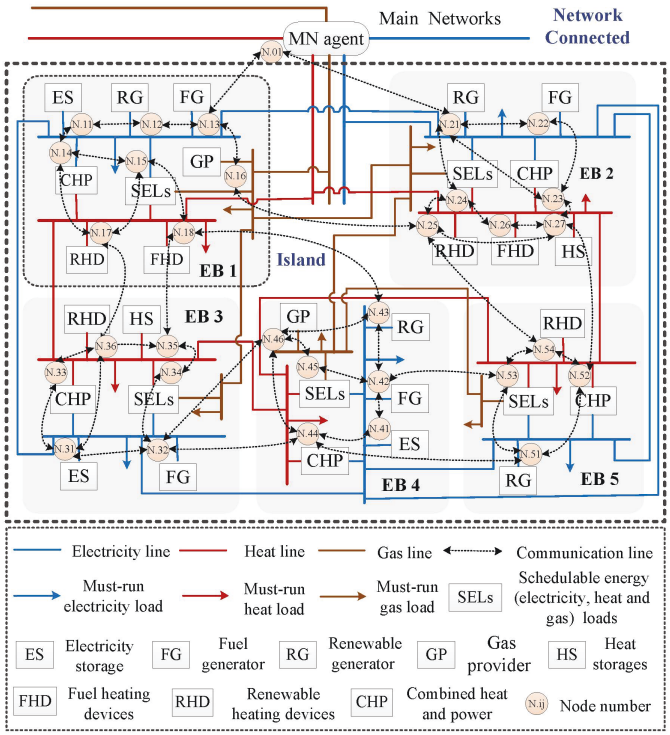

Fig. 1. Test system with five EBs and one MN agent.

scale expanses. In addition, if the centralized controller is in fault, the optimal calculation results cannot be obtained. Different from the centralized computation scheme, we aim at disaggregating the global computation problem and letting the individual participant and the MN agent solve it locally. Each of them only needs to be equipped with a small programmable processor. As discussed earlier, each participant or MN agent only needs to exchange the information of $\hat{y}_{ij}$ and $\hat{z}_{ij}$ with its neighbors. And the updating of x_{ij}, y_{ij}, z_{ij} and $\mathcal{Q}_{ij}$ are implemented via local calculation. It results in better privacy. More importantly, we can avoid making all heavy computational tasks be executed in the powerful centralized controller. The computational process can be implemented in individual participant or MN agent, each of which performs a local optimization. Meanwhile, a very small amount of computation is required for the distributed and small processor, resulting in reduced computational burden. Moreover, the failure of the individual processor only affects itself. The rest of participants can still calculate the new optimal operation as long as the graph connectivity is maintained. Consequently, the system possesses better robustness against single-node failure by using the distributed calculation. Meanwhile, it also provides better flexibility and scalability for integration of distributed energy resources.

IV. Simulation Results

The performance of the proposed distributed Newton-Raphson algorithm is tested on a MES system with five EBs and one MN agent, whose configuration and communication structures are shown in Fig. 1. Each EB is equipped with its own energy (i.e., electricity, heat, and gas) generation devices and energy loads (including schedulable energy loads

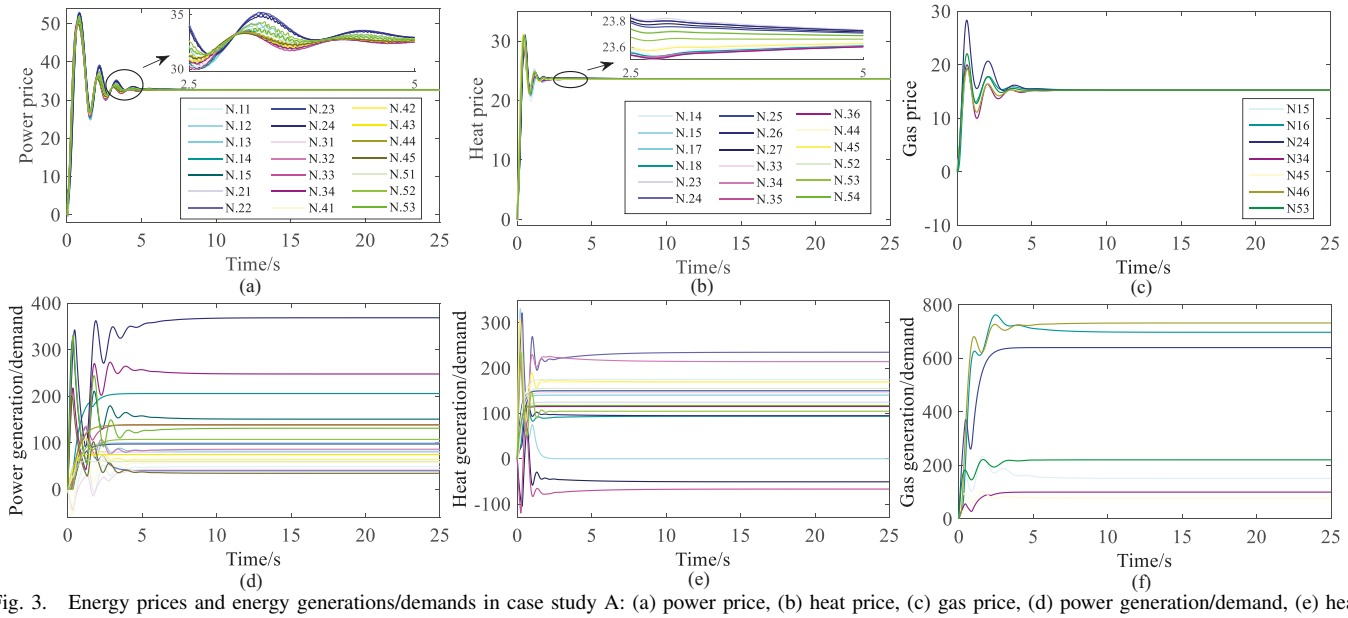

Fig. 3. Energy prices and energy generations/demands in case study A: (a) power price, (b) heat price, (c) gas price, (d) power generation/demand, (e) heat generation/demand, (f) gas generation/demand.

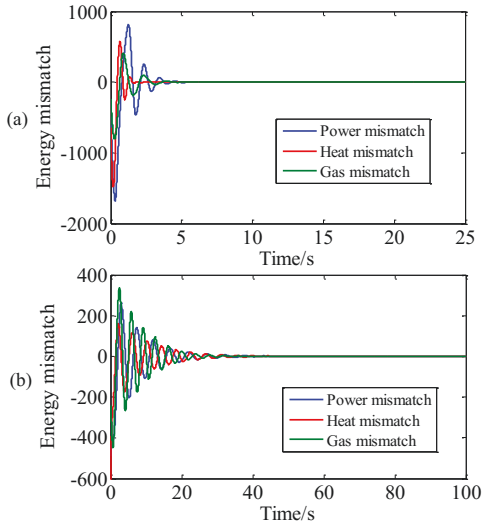

Fig. 2. Energy mismatch in case study A: (a) distributed dynamic event-triggered Newton-Raphson algorithm, (b) gradient-based method.

and must-run energy loads). The MN agent serves as the interface between the EBs and the main networks. The system is operated in network-connected mode if the MN agent takes part in the EMP; otherwise, it is operated in islanded mode. EBs are connected with each other via the interconnected lines, where the green, red and yellow solid lines represent the electricity, heat and gas lines, respectively. Moreover, the point-to-point communication network topology is adopted, which is also exhibited in Fig. 1 via dashed lines. Each participant only exchanges information with its neighbors. It is not very difficult to verify that the communication graph is connected, which meets the assumption in Theorem 1. The point-to-point communication is also a typical communication structure, which is much simpler than mesh network. We can employ either the wired or wireless communication technologies to achieve point-to-point communication in the real-world. Meanwhile, multiple communication protocols are also available, such as the Ethernet TCP/IP, ZigBee over IEEE 802.15.4 and Bluetooth over IEEE 802.15.1, etc. To simplify notation, we use "N.ij" as the index for the participant or MN agent, each of which is seen as a node. Therein, "N." is the abbreviation of number, $\{ij|i = 1, \cdots, n; j = 1, \cdots, \wp_i\}$ represents jth participant of ith EB, and $\{ij|i = 0; j = 1\}$ represents the MN agent. For example, the third participant, i.e., the FG in EB1 is numbered as N.13. The simulation test platform is built in the MATLAB environment. We coded the proposed algorithm for each participant and MN agent by programing individual M-function with ode45 solver. The algorithm is implemented on a computer with Intel Xeon E-2186M (Six Core Xeon 2.90GHz, 4.80GHz Turbo, 12MB 45W) and 64 GB memory. The parameters of cost and utility functions and constraints of each participant within each EB are obtained from [31]. The bases for the per-unit calculation are as follows: 1 p.u.=1MW for power or heat, 1 p.u.=84SCM/h for gas, and 1 p.u.=1$/MWh for price.

A. Convergence Analysis and Comparison

In the case study, the system is running in islanded mode. We focus on showing convergence of the proposed algorithm toward optimality. We also show its better performance in terms of faster convergence speed and less communication times by comparing with the most recent gradient-based method proposed in [34]. Without loss of generality, the must-run power, heat and gas loads for EB1 to EB5 are set as [150(p.u.), 124(p.u.), 50(p.u.)], [105(p.u.), 150(p.u.), 60(p.u.)], [85(p.u.), 135(p.u.), 80(p.u.)], [100(p.u.), 90(p.u.), 50(p.u.)], [50(p.u.), 140(p.u.), 0(p.u.)], respectively. By implementing the proposed algorithm, the trajectories of global energy supply-demand mismatches, dual variables (i.e., energy market clearing prices), and energy generation/demand of each participant are depicted in Fig. 2. (a), Figs. 3. (a)-(c), and Figs. 3. (d)-(f), respectively. It can be

TABLE II
COMMUNICATION TIMES

CE	-1	-2	-3	-4	-5
Proposed algorithm	2710	3405	4105	4830	5870
Gradient-based method	8286	10696	12904	15212	17530

seen that the global energy mismatches gradually converge to zeros, which implies that the global energy supply-demand balances have been satisfied. As shown in Fig. 3(a), the estimated power prices for the corresponding participants converge to a common value which is the final power market clearing price. Since there are multiple curves in Fig. 3(a), it is hard to identify the trajectories of different participants. To clearly exhibit the convergence process, we plot the zoomed-in trajectories from 2.5s to 5s. Furthermore, the legend for the Fig. 3(d) is the same as that Fig. 3(a). It can be observed from Fig. 3(d) that the corresponding power generation/demand of each participant can finally converge to steady value satisfying the local operation constraints. Similarly, the estimated heat and gas prices for the corresponding participants converge to the final heat and gas market clearing prices as shown in Fig. 3(b) and Fig. 3(c), respectively. Meanwhile, the heat and gas generation/demand of each participant also converges to the feasible steady value. Therein, Fig. 3(b) and Fig. 3(c) use the same legends with Fig. 3(e) and Fig. 3(f), respectively. Besides, as a demonstration, Fig. 4 shows the event-triggered instants of the participants in EB 1 (i.e., the participants from number N.11 to N.18). Therein, the marker "+" is used to represent the instant that an event is triggered by the corresponding participant. It can be observed from Fig. 4 that the communication pattern of each participant follows asynchronous and discrete-time fashion. The aforementioned results illustrate the behavior ensured by Theorem 1. Thus, the correctness and effectiveness of the proposed algorithm are verified.

Next, with the same topology and load conditions, the gradient-based method is employed to solve the problems (30-32). The results of the energy mismatches are depicted in Fig. 2(b). Comparing Fig. 2(a) with Fig. 2(b), it can be observed that the proposed Newton-Raphson algorithm converges within 11s, whereas it takes 50s for gradient-based method to converge. This implies that the proposed Newton-Raphson algorithm can greatly accelerate the convergence rate. The reason is that the Newton method can take advantages of both second-order and first-order information rather than first-order information (or gradient information) only. Furthermore, to clearly show the better performance in terms of the less communication times, we define the computation error as $CE = \frac{1}{N}\ln(\sum_{i=1}^{n}\sum_{j=1}^{\wp_i}|x_{ij} - x_{ij}^*|)$. The total numbers of communication times of all participants for different computation accuracy by using the proposed algorithm and the method in [34] are listed in Table II. It can be seen that the proposed algorithm needs less communication times than the method in [34] to reach the same CE. This is mainly because that we employ dynamic triggering mechanism which significantly improves the inter-event time.

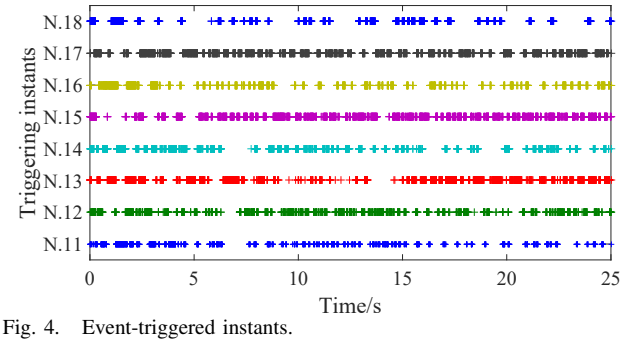

Fig. 4. Event-triggered instants.

B. Effectiveness Analysis in Mesh Communication Network

In this case study, we aim at showing the effectiveness of the proposed algorithm in mesh communication architecture instead of the point-to-point communication architecture in Fig. 1. To simplify communication, only one data concentrator is employed in each EB. The data concentrator collects the information of all the participants within the corresponding EB and achieves the information interaction with other data concentrators. The five data concentrators are interconnected with each other to form a typical mesh network. In this structure, the privacy of each participant is compromised. The system parameters and initial values are the same as the first case study. The trajectories of the energy generations/demands are shown in Fig. 5. It can be seen that the proposed algorithm is convergent. In addition, the final energy generations/demands of each participant under point-to-point communication structure and the mesh communication architecture are both listed in Table III. It can be observed that the final convergence values are almost the same. The results show that the proposed algorithm can work well in mesh communication architecture. In fact, as shown in Theorem 1, this paper only requires the communication graph to be connected. Thus, the proposed algorithm can be implemented in any kind of communication network architectures such as the typical mesh network, tree network as well as point-to-point network.

C. Mode Switching Test

In this case study, the system will be switched to different modes to verify the effectiveness of the proposed energy management model. All the topology, load and initial conditions are the same as those in the first case study. Then, at $t1 = 25s$, the system is switched from islanded mode to network-connected mode. We set $p_t^{M,max}$, $h_t^{M,max}$ and $g_t^{M,max}$ as the 50% of the total capabilities of power, heat and gas generations of all participants, respectively. For the MN agent, it buys power, heat and gas at prices 31.5(p.u.) $\sim$ 35(p.u.), 22.5(p.u.) $\sim$ 25(p.u.) and 12.6(p.u.) $\sim$ 14(p.u.), respectively. Meanwhile, it sells power, heat and gas at prices 35(p.u.) $\sim$ 38.5(p.u.), 25(p.u.) $\sim$ 27.5(p.u.) and 14(p.u.) $\sim$ 15.4(p.u.), respectively. According to Appendix A, we have $a_p^M = 0.0063$, $a_h^M = 0.0047$ and $a_g^M = 0.0019$. The trajectories of global energy supply-demand mismatches, energy market clearing prices, and energy generations/demands are shown in Fig. 6, Figs. 7(a)-(c), and Figs. 7(d)-(f), respectively. It can be seen

TABLE III
RESULTS OF ENERGY GENERATIONS AND CONSUMPTIONS

	Node	Point-to-point architecture			Mesh architecture		
		Power	Heat	Gas	Power	Heat	Gas
EB1	N.11	48.7267	0.0000	0.0000	48.7267	0.0000	0.0000
	N.12	99.9955	0.0000	0.0000	99.9955	0.0000	0.0000
	N.13	81.9282	0.0000	0.0000	81.9282	0.0000	0.0000
	N.14	205.9113	175.0625	0.0000	205.9113	175.0625	0.0000
	N.15	150.9604	0.0000	151.1226	150.9603	0.0000	151.1224
	N.16	0.0000	0.0000	697.1799	0.0000	0.0000	697.1795
	N.17	0.0000	140.4004	0.0000	0.0000	140.4004	0.0000
	N.18	0.0000	93.8493	0.0000	0.0000	93.8494	0.0000
EB2	N.21	79.8126	0.0000	0.0000	79.8126	0.0000	0.0000
	N.22	40.0000	0.0000	0.0000	40.0000	0.0000	0.0000
	N.23	97.5582	124.7394	0.0000	97.5582	124.7394	0.0000
	N.24	368.7100	234.9923	640.0000	368.7102	234.9922	640.0000
	N.25	0.0000	149.9000	0.0000	0.0000	149.9000	0.0000
	N.26	0.0000	94.9339	0.0000	0.0000	94.9338	0.0000
	N.27	0.0000	-50.9574	0.0000	0.0000	-50.9574	0.0000
EB3	N.31	43.0000	0.0000	0.0000	43.0000	0.0000	0.0000
	N.32	86.3168	0.0000	0.0000	86.3168	0.0000	0.0000
	N.33	139.1637	145.6757	0.0000	139.1637	145.6757	0.0000
	N.34	247.9712	214.2603	100.0000	247.9711	214.2604	100.0000
	N.35	0.0000	-66.9451	0.0000	0.0000	-66.9451	0.0000
	N.36	0.0000	115.3434	0.0000	0.0000	115.3434	0.0000
EB4	N.41	57.3022	0.0000	0.0000	57.3022	0.0000	0.0000
	N.42	63.8648	0.0000	0.0000	63.8648	0.0000	0.0000
	N.43	74.8884	0.0000	0.0000	74.8884	0.0000	0.0000
	N.44	137.9981	168.0150	0.0000	137.9981	168.0150	0.0000
	N.45	34.6088	169.6300	77.3928	34.6087	169.6301	77.3927
	N.46	0.0000	0.0000	731.8677	0.0000	0.0000	731.8679
EB5	N.51	59.7181	0.0000	0.0000	59.7181	0.0000	0.0000
	N.52	107.4697	155.1540	0.0000	107.4697	155.1540	0.0000
	N.53	131.4039	104.8088	220.5323	131.4040	104.8087	220.5323
	N.54	0.0000	117.5203	0.0000	0.0000	117.5203	0.0000

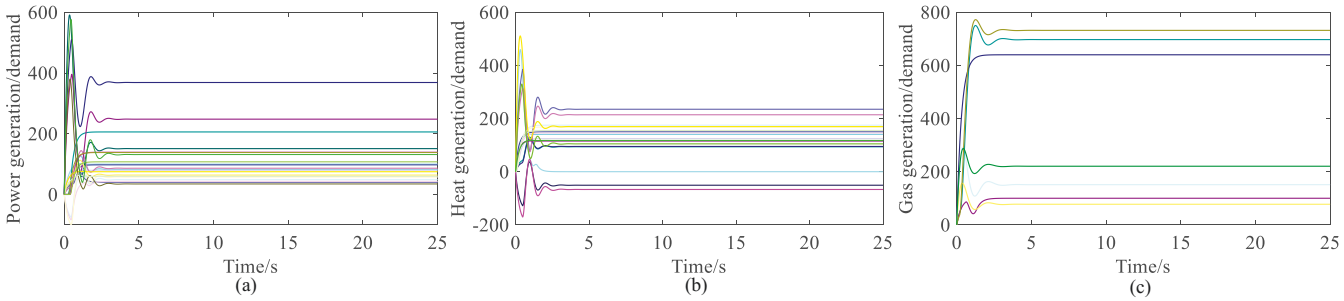

Fig. 5. Effectiveness test in mesh communication architecture: (a) power generation/demand, (b) heat generation/demand, (c) gas generation/demand.

that each participant can automatically respond to the changes of mode switching and converge to new optimal states. The final convergence values of energy market clearing prices for power, heat and gas are 33.4060(p.u.), 24.0344(p.u.) and 14.5382(p.u.), respectively. The final values of p_t^M, h_t^M and g_t^M are 143.3578(p.u.), 109.1287(p.u.), and -122.6236(p.u.), respectively. This implies that the main networks will buy power and heat from the EBs and sell gas to the EBs. In addition, it can be seen that the power (or heat) market clearing price is cheaper than the benchmark power (or heat) price. Meanwhile, the gas market clearing price is more expensive than the benchmark gas price. The results show the effectiveness of the designed model. Further, at $t2 = 50s$, the system switches to islanded mode. The simulation results are also shown in Figs. 6 and 7. It can be observed that each participant again converges to the new solutions responding to the new mode switching. Moreover, the final solutions are the same as those prior to the islanded mode. The results again exhibit the effectiveness of the proposed model and better adaptive performance of the proposed algorithm.

D. Fluctuations of Renewable Energy Sources

In this case study, we analyze the effectiveness of the proposed algorithm considering large output fluctuations of RGs and RHDs in islanded mode. The system structure and parameters are the same as those in the first case study. We let the forecasting output of each RG and RHD randomly change to the 50%-150% of the current value at $t1 = 25s$, and then change back to the original one at $t2 = 50s$. By implementing the proposed algorithm, the trajectory of the

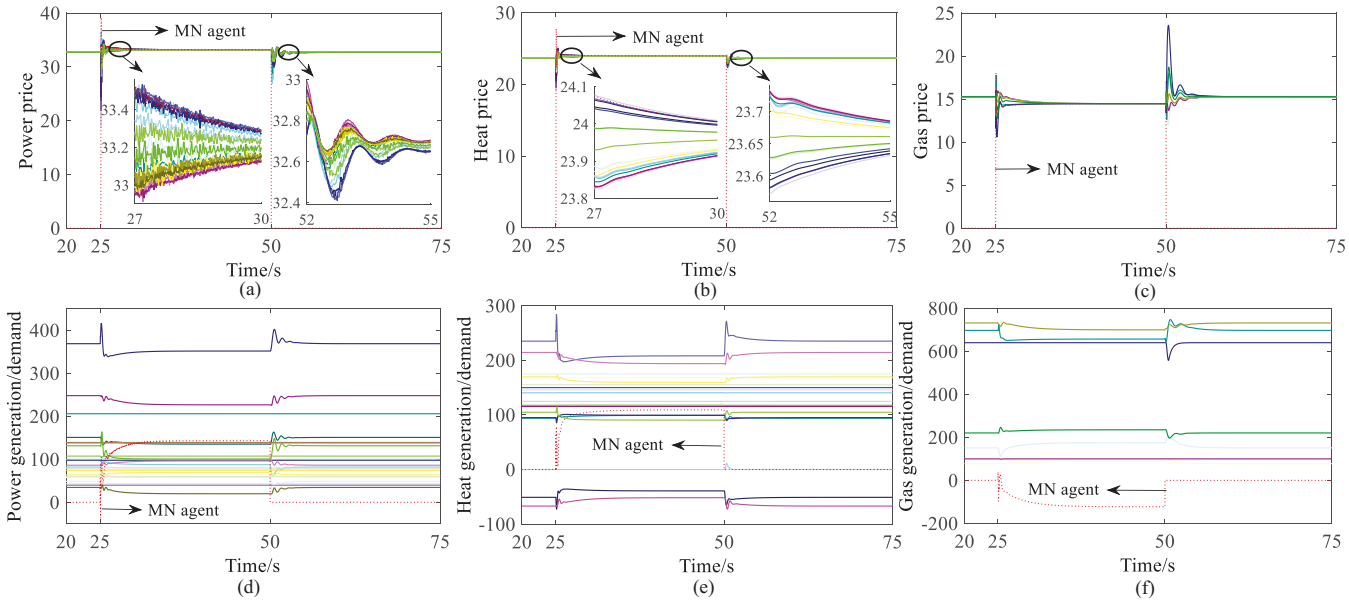

Fig. 7. Energy prices and energy generations/demands in case study C: (a) power price, (b) heat price, (c) gas price, (d) power generation/demand, (e) heat generation/demand, (f) gas generation/demand.

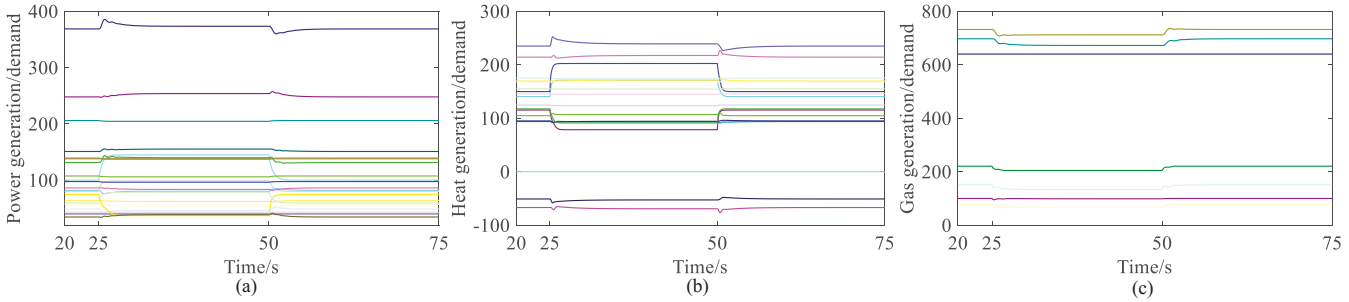

Fig. 8. Effectiveness test for fluctuations of renewable energy generations: (a) power generation/demand, (b) heat generation/demand, (c) gas generation/demand.

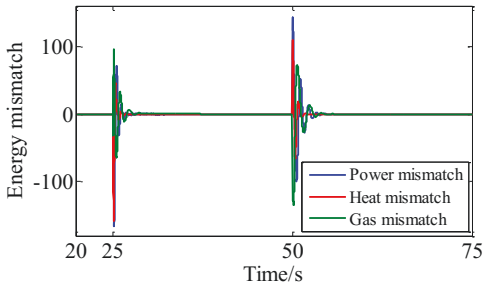

Fig. 6. Energy mismatch in case study C

energy generation/demand of each participant is shown in Fig. 8. At $t1$, it can be observed that each participant is able to automatically adjust its local output/demand in response to the fluctuations of renewable energy generations. After a few seconds, each participant finally converges to a new equilibrium point. Further, at $t2$, each participant again adjusts its local operation and gradually converges to a new solution responding to the new changes of RGs and RHDs. In addition, it can be seen that the final convergence solutions of all participants after $t2$ are the same as the original ones before $t1$. The results show that the proposed algorithm possesses better adaptability and flexibility to accomodate to the fluctuations of renewable energy generations.

E. Robustness Analysis under Model Change and Perturbations

In this case study, we focus on verifying the better robustness of the proposed algorithm under the change of system model and parameter perturbations by comparing with the most recently proposed method in [36]. From the results in the fourth case study, it has been verified that the proposed algorithm can better accomodate to renewable energy fluctuations. In this section, we will mianly analyze the model change caused by device fault and bounded perturbations for algorithm parameters.

Firstly, on the basis of the first case study, the CHP of EB2 and the SEL of EB5 are in fault and removed from the system at $t1$ (after system convergence). It can be seen from Fig. 9 that each participant can automatically respond to the fault and converge to new states. More importantly, as shown in Fig. 10(a), the global power, heat and gas supply-demand mismatches gradually converge to zeros. With the same conditions, the consensus-based distributed method proposed in [36] is used to solve the same problem. We do not reset the initial values after the sudden model change. The trajectories of global energy supply-demand mismatches

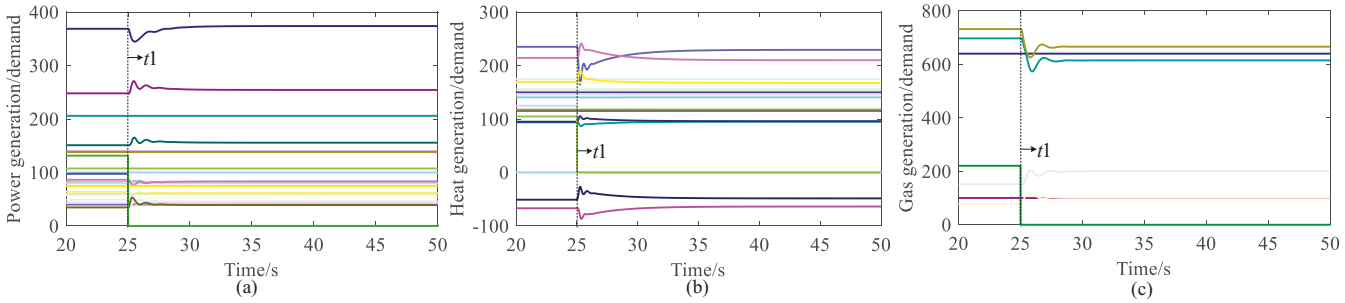

Fig. 9. Energy generation/demand after model change: (a) power generation/demand, (b) heat generation/demand, (c) gas generation/demand.

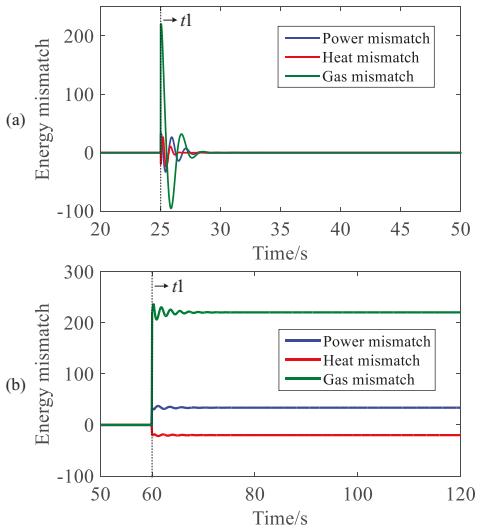

Fig. 10. Energy mismatch after model change: (a) distributed dynamic event-triggered Newton-Raphson algorithm, (b) the method proposed in [36] with setting 0.05s for each iteration.

are shown in Fig. 10(b). It can be observed that the global energy supply-demand mismatches cannot go back to zeros, which implies that the results are not feasible. This is because the method proposed in [36] requires strict initial conditions. Once the system model is changed, it is necessary to reset the global initial values. Otherwise, the corresponding algorithm may fail to provide a feasible solution. It results in weakened robustness. Compared with the method in [36], our proposed algorithm is free of initialization. The proposed algorithm can adaptively respond to the change without resetting initial values and re-running system dynamics. Thus, the proposed algorithm possesses strong robustness for the change of the system model.

Next, for each participant, we randomly added bounded perturbations for the parameters shown in Theorem 1 at $t2 = 2.5s$. The topology and load conditions are the same as those in the first case study. The trajectories of energy generation or demand are shown in Fig. 11. It can be seen that each participant can converge to the corresponding optimal solution. The results verify that the proposed algorithm also possesses strong robustness against bounded parameter perturbations.

V. Conclusion

In this paper, a double-mode energy management model is established to achieve the adaptive switching between islanded and network-connected modes. It is finally formulated as a distributed optimization problem by transforming some decision variables. To solve this problem, a distributed dynamic event-triggered Newton-Raphson algorithm has been proposed. The proposed algorithm possesses some satisfactory characteristics such as faster convergence speed, asynchronous implementation, no special initialization condition and reduced communication. By using Lyapunov technique, the convergence and optimality of the proposed algorithm have been proved. In the future, we would like to take more physical limits of power-heat-gas networks into consideration, explore effective relaxed method to expand the application of our proposed algorithm, and study the coordination control of electric, heat and gas energy systems.

Appendix

A. Choice of Parameters

As seen in Fig. A1, the main grid buy at price from $b\gamma_t^{p,min}$ to $b\gamma_t^p$ and sell at price from $b\gamma_t^p$ to $b\gamma_t^{p,max}$, where $b\gamma_t^{p,min}$ and $b\gamma_t^{p,max}$ are the minimum buying price and maximum selling price at time t. The lower the purchasing price, the more power the main grid would like to buy, and vice versa. According to the linear relationship between p_t^M and the market clearing price, we can set $a_p^M = (b\gamma_t^{p,max} - b\gamma_t^{p,min})/p_t^{M,max}$. The similar method can be applied to a_h^M and a_g^M.

B. Preliminary Knowledge

Based on the characteristics of the considered EMP in MES, some important properties are introduced in this section. Firstly, it should be noted that, for each participant or MN agent, the term of local objective function is of exponential-form, quadratic-form or cubic-form. Thus, it is not very difficult to verify that $W(x_{ij})$ is strongly convex within closed convex set Ω_{ij} for all $i = 0, \cdots, n$, $j = 1, \cdots, \wp_i$. Therefore, there exists positive constants $\emptyset_{ij}$ and ϵ_{ij} such that for any $x_{ij}, \breve{x}_{ij} \in \Omega_{ij}$, the following inequalities are satisfied, i.e.,

$$(x_{ij} - \breve{x}_{ij})^T(\nabla W(x_{ij}) - \nabla W(\breve{x}_{ij})) \geq \emptyset_{ij}||x_{ij} - \breve{x}_{ij}||^2, \quad (38)$$

$$\nabla^2 W(x_{ij}) \geq \epsilon_{ij} I_3, \quad (39)$$

where ∇ represents differentiation operator. In addition, we let $\nabla^2 W(x_{ij})$ be as an approximation of Hessian for CHP unit by employing Jacobi approximation. The purpose is to reduce

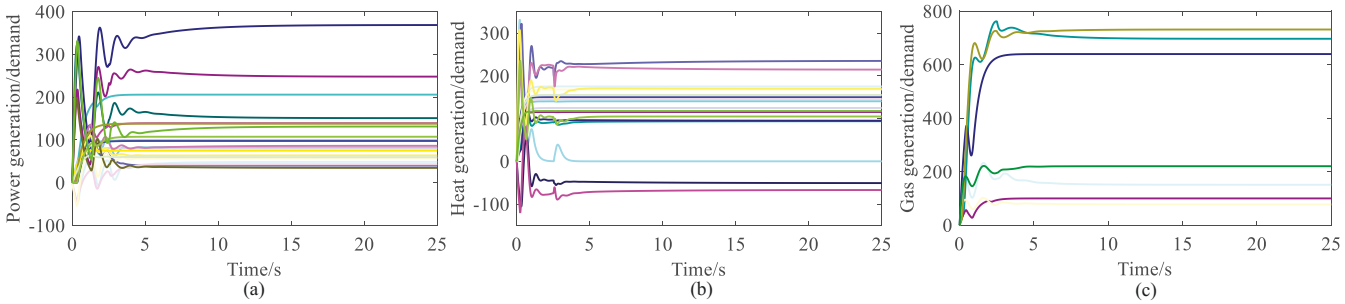

Fig. 11. Bounded parameter perturbations: (a) power generation/demand, (b) heat generation/demand, (c) gas generation/demand.

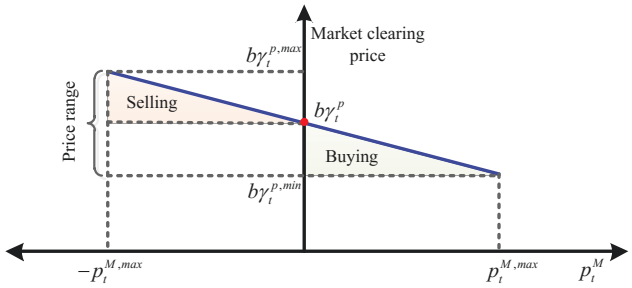

Fig. A1. Power price.

underlying computation accordingly. We define $\nabla^2 W(x_{ij})^{-1}$ as the inverse of $\nabla^2 W(x_{ij})$.

We use 1_d (resp. 0_d) to denote the d-dimensional column vector with all elements being 1 (resp. 0). Next, for each participant or MN agent, the normal cone of local Ω_{ij} is defined as $\Lambda(x_{ij}) = \{o | o^T(\breve{x}_{ij} - x_{ij}) \leq 0, \forall \breve{x}_{ij} \in \Omega_{ij}\}$. We further let $\widetilde{\Lambda}(x_{ij}) = \{o | o^T(\breve{x}_{ij} - x_{ij}) \leq 0, ||o||^2 = 1, \forall \breve{x}_{ij} \in \Omega_{ij}\}$ if $x_{ij} \in \Omega_{ij}^{bo}$, and $\widetilde{\Lambda}(x_{ij}) = 0_3$ if $x_{ij} \in \Omega_{ij}^{in}$. Therein, Ω_{ij}^{bo} and Ω_{ij}^{in} denote the boundary and inside of Ω_{ij}, respectively. Since Ω_{ij} is closed convex set, the tangent cone of Ω_{ij} at x_{ij} can be defined as $\Upsilon_{\Omega_{ij}}(x_{ij}) = \{\breve{x} | \breve{x}^T \mathfrak{d} \leq 0, \forall \mathfrak{d} \in \Lambda(x_{ij})\}$. Then, for $x_{ij} \in \Omega_{ij}$, the projection and differentiated projection operator are defined as $P_{\Omega_{ij}}(x_{ij}) = \arg \min_{\breve{x}_{ij}} ||x_{ij} - \breve{x}_{ij}||$ and $\Gamma_{\Omega_{ij}}(x_{ij}, \emptyset) = \lim_{\lambda \to 0} \big(P_{\Omega_{ij}}(x_{ij} + \lambda \emptyset) - x_{ij}\big)/\lambda$, respectively.

Lemma 1 [40]: 1) $\Gamma_{\Omega_{ij}}(x_{ij}, \emptyset) = \emptyset$ if $x_{ij} \in \Omega_{ij}^{in}$ (or $x_{ij} \in \Omega_{ij}^{bo}$ and $\max_{\rho_{ij} \in \widetilde{\Lambda}(x_{ij})} \emptyset^T \rho_{ij} \leq 0$); 2) $\Gamma_{\Omega_{ij}}(x_{ij}, \emptyset) = \emptyset - \emptyset^T \rho_{ij}^* \rho_{ij}^*$ if $x_{ij} \in \Omega_{ij}^{bo}$ and $\max_{\rho_{ij} \in \widetilde{\Lambda}(x_{ij})} \emptyset^T \rho_{ij} \geq 0$, where $\rho_{ij}^* = \arg \max_{\rho_{ij} \in \widetilde{\Lambda}(x_{ij})} \emptyset^T \rho_{ij}$.

Lemma 1 implies that $\Gamma_{\Omega_{ij}}(x_{ij}, \emptyset)$ is equivalent to $P_{\Upsilon_{\Omega_{ij}}(x_{ij})}(\emptyset)$.

C. Proof of Theorem 1

Firstly, the dynamics (33-35) shown in Algorithm 1 can be rewritten in state space as

$$\dot{X} = \Gamma_\Omega \Big(\nabla^2 W(X)^{-1} X, \nabla^2 W(X)^{-1} \big(-\nabla W(X) + B\hat{Y} \big) \Big), \tag{40}$$

$$\dot{Y} = -(L \otimes I_3)\hat{Y} - (L \otimes I_3)\hat{Z} + D - BX, \tag{41}$$

$$\dot{Z} = (L \otimes I_3)\hat{Y}, \tag{42}$$

where $\otimes$ denotes Kronecker product; $\Omega = \bigcup \{\Omega_{ij} | i = \aleph, \cdots, n; j = 1, \cdots, \wp_i\}$ represents the Cartesian product of the sets $\{\Omega_{ij}\}$; X is the column vector made from concatenation of x_{ij}'s. Similarly, Y, Z, $\nabla W(X)$ and D are the column stack vector forms of y_{ij}, z_{ij}, $\nabla W(x_{ij})$ and d_{ij}, respectively; $\nabla^2 W(X)^{-1} = \text{diag}\{\nabla^2 W(x_{ij})^{-1}\}$ and $B = \text{diag}\{B_{ij}\}$.

The symbol " $*$ " is used to represent equilibrium point. It follows that

$$0_{3N} = \Gamma_\Omega \Big(\nabla^2 W(X^*)^{-1} X^*, \nabla^2 W(X^*)^{-1} \big(-\nabla W(X^*) + BY^* \big) \Big), \tag{43}$$

$$0_{3N} = -(L \otimes I_3)Y^* - (L \otimes I_3)Z^* + D - BX^*, \tag{44}$$

$$0_{3N} = (L \otimes I_3)Y^*. \tag{45}$$

We consider that the implementation of the proposed algorithm is in a connected graph. Thus, it can be derived from (44) and (45) that

$$(1_N^T \otimes I_3)BX^* = (1_N^T \otimes I_3)D \tag{46}$$

$$Y^* = 1_N \otimes y^*. \tag{47}$$

In our considered model, $W(x_{ij})$ is a diagonal and positive definite matrix. According to the definition of normal cone, we have $\Lambda(\nabla^2 W(x_{ij})^{-1} x_{ij}) = \nabla^2 W(x_{ij})^{-1} \Lambda(x_{ij})$. It can be derived from Lemma 1 and (43) that

$$0_3 \in \nabla W(x_{ij}^*) - B_{ij} y_{ij}^* + \Lambda(x_{ij}^*). \tag{48}$$

Equations (46-48) are the optimality conditions of the studied problem by Theorem 3.34 in [41]. This implies that the equilibrium point is the optimal point.

To study the convergence, we transfer the optimal point to the origin and make use of some change of variables

$$\mathcal{X} = X - X^* = \Theta\mu; \quad \mu = \Theta^T \mathcal{X}; \tag{49}$$

$$\mathcal{Y} = Y - Y^* = \Theta\nu; \quad \nu = \Theta^T \mathcal{Y}; \quad e = \hat{\nu} - \nu; \tag{50}$$

$$\mathcal{Z} = Z - Z^* = \Theta\omega; \quad \omega = \Theta^T \mathcal{Z}; \quad f = \hat{\omega} - \omega; \tag{51}$$

where X^*, Y^* and Z^* are the optimal point of X, Y and Z respectively. $\Theta^T = [r, \mathcal{R}]^T \otimes I_3$, $r = \frac{1}{\sqrt{N}} 1_N$, $r^T \mathcal{R} = 0_N^T$, $\mathcal{R}\mathcal{R}^T = I_N - \frac{1}{n} 1_N 1_N^T$ and $\mathcal{R}^T \mathcal{R} = I_{N-1}$. The new variables are partitioned as $\mu = col(\mu_1, \mu_{2:N})$, $\nu = col(\nu_1, \nu_{2:N})$, $\omega = col(\omega_1, \omega_{2:N})$, $e = col(e_1, e_{2:N})$ and $f = col(f_1, f_{2:N})$.

It can be derived from Lemma 1 that

$$\Gamma_{\Omega_{ij}} \Big(\nabla^2 W(x_{ij})^{-1} x_{ij}, \nabla^2 W(x_{ij})^{-1} \big(-\nabla W(x_{ij}) + B_{ij}^T \hat{y}_{ij} \big) \Big) = \nabla^2 W(x_{ij})^{-1} \big(-\nabla W(x_{ij}) + B_{ij}^T \hat{y}_{ij} - \zeta_{ij}(x_{ij}) \rho_{ij}(x_{ij}) \big), \tag{52}$$

where $\rho_{ij}(x_{ij}) \in \widetilde{\Lambda}(x_{ij})$, $\zeta_{ij}(x_{ij}) \geq 0$. Thus, at the optimal point, one has

$$-\nabla W(x_{ij}^*) + B^T y_{ij}^* = -\zeta_{ij}(x_{ij}^*)\rho_{ij}(x_{ij}^*). \tag{53}$$

We denote $\Xi_\Omega(X)$ and $\Xi_\Omega(X^*)$ as the column stack vector forms of $\zeta_{ij}(x_{ij})\rho_{ij}(x_{ij})$ and $\zeta_{ij}(x_{ij}^*)\rho_{ij}(x_{ij}^*)$, respectively. Then, in the new variables, dynamics (40-42) read as

$$\begin{aligned}\begin{bmatrix}\dot{\mu}_1\\ \dot{\mu}_{2:N}\end{bmatrix} =&\Theta^T\nabla^2 W(\mathcal{X}+X^*)^{-1}B^T\Theta(\nu+e)\\ &-\Theta^T\Big(\nabla^2 W(\mathcal{X}+X^*)^{-1}\big(\nabla W(\mathcal{X}+X^*)\\ &-\nabla W(X^*)+\Xi_\Omega(X)-\Xi_\Omega(X^*)\big),\end{aligned} \tag{54}$$

$$\begin{aligned}\begin{bmatrix}\dot{\nu}_1\\ \dot{\nu}_{2:N}\end{bmatrix} =&-\begin{bmatrix}0_3\\ \big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)(\nu_{2:N}+e_{2:N}+\omega_{2:N}+f_{2:N})\end{bmatrix}\\ &-\Theta^T B\Theta\mu,\end{aligned} \tag{55}$$

$$\begin{bmatrix}\dot{\omega}_1\\ \dot{\omega}_{2:N}\end{bmatrix} = \begin{bmatrix}0_3\\ \big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)(\nu_{2:N}+e_{2:N})\end{bmatrix}. \tag{56}$$

In this paper, we use Jacobi as an approximation of Hessian for CHP units. For simplification, we define Ξ as the set of CHP units and $\{\mathcal{V}-\Xi\}$ as the set without CHP units. Note that each CHP unit has a convex cost function in quadratic-form. Hence, the corresponding Jacobi matrix is a positive definite matrix with all elements being constants. Recalling (39), it follows that

$$\begin{aligned}V_{ij}^a =&\frac{1}{2}(x_{ij}-x_{ij}^*)^T\nabla^2 W(x_{ij})(x_{ij}-x_{ij}^*)\\ \geq&\frac{1}{2}\epsilon_{ij}||(x_{ij}-x_{ij}^*)||^2,\ \forall ij\in\Xi.\end{aligned} \tag{57}$$

Moreover, in light of (38), we have the following

$$\begin{aligned}V_{ij}^b =&W(x_{ij}^*)-W(x_{ij})-(x_{ij}^*-x_{ij})^T\nabla W(x_{ij})\\ \geq&\frac{1}{2}\emptyset_{ij}||x_{ij}-x_{ij}^*||^2,\ \forall ij\in\{\mathcal{V}-\Xi\}.\end{aligned} \tag{58}$$

Based on the change of variables shown in (49), one has

$$V_{one} = \sum_{ij\in\Xi} V_{ij}^a + \sum_{ij\in\{\mathcal{V}-\Xi\}} V_{ij}^b \geq \overline{\emptyset\epsilon}||\Theta\mu||^2. \tag{59}$$

where $\overline{\emptyset\epsilon} = \min\big\{\frac{1}{2}\emptyset_{ij}|ij\in\Xi, \frac{1}{2}\epsilon_{ij}|ij\in\{\mathcal{V}-\Xi\}\big\}$. In addition, it should be pointed out that V_{one} is also a function related to the new variable μ.

Next, we construct the following candidate Lyapunov function,

$$\begin{aligned}V =&\pounds_1 V_{one}+\frac{1}{2}\pounds_1||\nu||^2+\frac{1}{2}(\pounds_1+\pounds_2)||\omega_{2:N}||^2\\ &+\frac{1}{2}\pounds_2||\nu_{2:N}+\omega_{2:N}||^2+\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\mathcal{Q}_{ij},\end{aligned} \tag{60}$$

where $\pounds_1$ and $\pounds_2$ are positive constants.

According to (36) and (37), we have

$$\dot{\mathcal{Q}}_{ij} \geq -\vartheta_{1,ij}\mathcal{Q}_{ij} - \frac{\vartheta_{2,ij}}{\vartheta_{5,ij}}\mathcal{Q}_{ij}. \tag{61}$$

It follows that

$$\mathcal{Q}_{ij}(t) \geq \mathcal{Q}_{ij}(t0)e^{-(\vartheta_{1,ij}+\frac{\vartheta_{2,ij}}{\vartheta_{5,ij}})t} > 0,\ \forall t > t0. \tag{62}$$

The Lie derivative of V along (54-56) and (36) is

$$\begin{aligned}\dot{V} =&\pounds_1\mu^T\Theta^T\nabla^2 W(\Theta\mu+X^*)\Theta\dot{\mu}\\ &+\pounds_1\nu^T\dot{\nu}+(\pounds_1+\pounds_2)\omega_{2:N}^T\dot{\omega}_{2:N}\\ &+\pounds_2(\nu_{2:N}+\omega_{2:N})^T(\dot{\nu}_{2:N}+\dot{\omega}_{2:N})+\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\dot{\mathcal{Q}}_{ij}\\ =&-\pounds_1\mathcal{X}^T\big(\nabla W(X)-\nabla W(X^*)-\Xi_\Omega(X^*)+\Xi_\Omega(X)\big)\\ &+\pounds_1\mu^T\Theta^T B^T\Theta e\\ &-\frac{1}{2}\pounds_1\nu_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)\nu_{2:N}\\ &-\pounds_1\nu_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)(\frac{1}{2}\nu_{2:N}+e_{2:N}+f_{2:N})\\ &+\pounds_1\omega_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)e_{2:N}\\ &-\frac{1}{2}\pounds_2\omega_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)\omega_{2:N}\\ &-\pounds_2\omega_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)(\frac{1}{2}\omega_{2:N}-e_{2:N}+f_{2:N})\\ &-\pounds_2\nu_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)f_{2:N}\\ &-\pounds_2(\nu_{2:N}^T+\omega_{2:N}^T)(\mathcal{R}^T\otimes I_3)B(\mathcal{R}\otimes I_3)\mu_{2:N}\\ &+\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\dot{\mathcal{Q}}_{ij}.\end{aligned} \tag{63}$$

Since $B_{ij} = I_3$ or $-I_3$ and $\mathcal{R}^T\mathcal{R} = I_{N-1}$, we have

$$\begin{aligned}&\pounds_2(\nu_{2:N}^T+\omega_{2:N}^T)(\mathcal{R}^T\otimes I_3)B(\mathcal{R}\otimes I_3)\mu_{2:N} \leq\\ &\frac{\pounds_2\pounds_3}{2}(||\nu_{2:N}||^2+||\omega_{2:N}||^2)+\frac{\pounds_2}{2\pounds_3}\mathcal{X}^T\mathcal{X},\end{aligned} \tag{64}$$

$$\mu^T\Theta^T B^T\Theta e \leq \frac{1}{2}\mathcal{X}^T\emptyset\mathcal{X}+\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\frac{1}{2\emptyset_{ij}}||(\hat{y}_{ij}-y_{ij})||, \tag{65}$$

where $\emptyset = \text{diag}\{\emptyset_{ij}\}$.

Based on the definition of normal cone, we have

$$\begin{aligned}\mathcal{X}^T\Xi_\Omega(X) =&-\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\zeta_{ij}(x_{ij})(x_{ij}^*\\ &-x_{ij})^T\rho_{ij}(x_{ij})\geq 0,\end{aligned} \tag{66}$$

$$\begin{aligned}-\mathcal{X}^T\Xi_\Omega(X^*) =&-\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\zeta_{ij}(x_{ij}^*)(x_{ij}\\ &-x_{ij}^*)^T\rho_{ij}(x_{ij}^*)\geq 0.\end{aligned} \tag{67}$$

In addition, we also have the following facts

$$\begin{aligned}&-\nu_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)(\frac{1}{2}\nu_{2:N}+e_{2:N})\\ &=\frac{1}{2}(\hat{Y}-Y)^T(L\otimes I_3)(\hat{Y}-Y)\\ &\quad-\frac{1}{2}\hat{Y}^T(L\otimes I_3)\hat{Y},\end{aligned} \tag{68}$$

$$\begin{aligned}&-\omega_{2:N}^T\big((\mathcal{R}^T L\mathcal{R})\otimes I_3\big)(\frac{1}{2}\omega_{2:N}+f_{2:N})\\ &=\frac{1}{2}(\hat{Z}-Z)^T(L\otimes I_3)(\hat{Z}-Z)\\ &\quad-\frac{1}{2}(\hat{Z}-Z^*)^T(L\otimes I_3)(\hat{Z}-Z^*),\end{aligned} \tag{69}$$

$$
\begin{aligned}
&(\hat{Y}-Y)^T(L\otimes I_3)(\hat{Y}-Y)\\
&\quad\le 2(\hat{Y}-Y)^T(D\otimes I_3)(\hat{Y}-Y)\\
&\quad=\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}2|\mathcal{N}_{ij}|\cdot||\hat{y}_{ij}-y_{ij}||^2, \qquad (70)\\
&(\hat{Z}-Z)^T(L\otimes I_3)(\hat{Z}-Z)\\
&\quad\le\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}2|\mathcal{N}_{ij}|\cdot||\hat{z}_{ij}-z_{ij}||^2, \qquad (71)\\
&\nu_{2:N}^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)f_{2:N}\\
&\quad\le\frac{1}{8}\nu_{2:N}^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)\nu_{2:N}\\
&\qquad+2(\hat{Z}-Z)^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)(\hat{Z}-Z), \qquad (72)\\
&\omega_{2:N}^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)e_{2:N}\\
&\quad\le\frac{1}{8}\omega_{2:N}^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)\omega_{2:N}\\
&\qquad+2(\hat{Y}-Y)^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)(\hat{Y}-Y), \qquad (73)\\
&\qquad\nu_{2:N}^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)\nu_{2:N}\ge\lambda_2||\nu_{2:N}||^2, \qquad (74)\\
&\qquad\omega_{2:N}^T\big((\mathcal{R}^TL\mathcal{R})\otimes I_3\big)\omega_{2:N}\ge\lambda_2||\omega_{2:N}||^2. \qquad (75)
\end{aligned}
$$

We let $\pounds_4=\frac{\pounds_2}{\pounds_1}$. Based on the selected coefficients in Theorem 1 and the facts (38, 66-75), we have

$$
\begin{aligned}
\dot{V}\le&-\frac{1}{2}\pounds_1\mathcal{X}^T\emptyset\mathcal{X}+\frac{\pounds_2}{2\pounds_3}\mathcal{X}^T\mathcal{X}\\
&-\frac{\pounds_2}{2}(\hat{Z}-Z^*)^T(L\otimes I_3)(\hat{Z}-Z^*)\\
&-\big((\frac{3}{8}\pounds_1-\frac{1}{8}\pounds_2)\lambda_2-\frac{\pounds_2\pounds_3}{2}\big)||\nu_{2:N}||^2\\
&-\big((\frac{3}{8}\pounds_2-\frac{1}{8}\pounds_1)\lambda_2-\frac{\pounds_2\pounds_3}{2}\big)||\omega_{2:N}||^2\\
&+\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\big(\frac{1}{2\emptyset_{ij}}+5|\mathcal{N}_{ij}|+4\frac{\pounds_2}{\pounds_1}|\mathcal{N}_{ij}|\big)||(\hat{y}_{ij}-y_{ij})||^2\\
&+\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}(5\frac{\pounds_2}{\pounds_1}|\mathcal{N}_{ij}|+4|\mathcal{N}_{ij}|)||(\hat{z}_{ij}-z_{ij})||^2\\
&-\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\sum_{\overline{ij}\in\mathcal{N}_{ij}}\frac{1}{4}a_{ij,\overline{ij}}||\hat{y}_{ij}-\hat{y}_{\overline{ij}}||^2\\
&-\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\vartheta_{1,ij}\mathcal{Q}_{ij} \qquad (76)\\
&-\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}\vartheta_{2,ij}\big(\vartheta_{3,ij}||(\hat{y}_{ij}-y_{ij})||^2\\
&+\vartheta_{4,ij}||(\hat{z}_{ij}-z_{ij})||^2-\frac{1}{4}\sum_{\overline{ij}\in\mathcal{N}_{ij}}a_{ij,\overline{ij}}||\hat{y}_{ij}-\hat{y}_{\overline{ij}}||^2\big)\\
\le&-\frac{1}{2}\mathcal{X}^T\mathrm{diag}\{\pounds_1\emptyset_{ij}-\frac{\pounds_2}{\pounds_3}\}\mathcal{X}\\
&-\frac{\pounds_2}{2}(\hat{Z}-Z^*)^T(L\otimes I_3)(\hat{Z}-Z^*)\\
&-\big((\frac{3}{8}\pounds_1-\frac{1}{8}\pounds_2)\lambda_2-\frac{\pounds_2\pounds_3}{2}\big)||\nu_{2:N}||^2\\
&-\big((\frac{3}{8}\pounds_2-\frac{1}{8}\pounds_1)\lambda_2-\frac{\pounds_2\pounds_3}{2}\big)||\omega_{2:N}||^2\\
&-\pounds_1\sum_{i=0}^{n}\sum_{j=1}^{\wp_i}(\vartheta_{1,ij}-\frac{1-\vartheta_{2,ij}}{\vartheta_{5,ij}})\mathcal{Q}_{ij}<0. \qquad (77)
\end{aligned}
$$

Based on the Lyapunov stability theory by Theorem 4.1 in [42], it is concluded that dynamics (33-36) can converge to the equilibrium point, i.e., optimal point. The proof is thus completed.

D. Lower Bound of $t_{ij}^{k+1}-t_{ij}^k$

It can be driven from (34) that the right-hand derivative of $||(\hat{y}_{ij}-y_{ij})||$, denoted as $\frac{d^+}{dt}||(\hat{y}_{ij}-y_{ij})||$, satisfies

$$
\begin{aligned}
\frac{d^+}{dt}||(\hat{y}_{ij}-y_{ij})||\le&||\sum_{\overline{ij}\in\mathcal{N}_{ij}}a_{ij,\overline{ij}}(\hat{y}_{ij}-\hat{y}_{\overline{ij}})||\\
&+||\sum_{\overline{ij}\in\mathcal{N}_{ij}}a_{ij,\overline{ij}}(\hat{z}_{ij}-\hat{z}_{\overline{ij}})||\\
&+||d_{ij}-B_{ij}x_{ij}||=\pounds_{ij,y}. \qquad (78)
\end{aligned}
$$

Similarly, in light of (35), we have

$$
\frac{d^+}{dt}||(\hat{z}_{ij}-z_{ij})||\le||\sum_{\overline{ij}\in\mathcal{N}_{ij}}a_{ij,\overline{ij}}(\hat{y}_{ij}-\hat{y}_{\overline{ij}})||=\pounds_{ij,z}. \qquad (79)
$$

According to (78) and (79), one has

$$
\begin{aligned}
&\vartheta_{3,ij}||(\hat{y}_{ij}-y_{ij})||^2+\vartheta_{4,ij}||(\hat{z}_{ij}-z_{ij})||^2\\
&\qquad\le(\vartheta_{3,ij}\pounds_{ij,y}^2+\vartheta_{4,ij}\pounds_{ij,z}^2)(t-t_{ij}^k)^2. \qquad (80)
\end{aligned}
$$

Note that $\sum_{\overline{ij}\in\mathcal{N}_{ij}}a_{ij,\overline{ij}}||\hat{y}_{ij}-\hat{y}_{\overline{ij}}||^2\ge 0$ and (62) hold. Recalling the triggering mechanism (37), when the next triggering occurs, it follows that

$$
\begin{aligned}
&\vartheta_{5,ij}(\vartheta_{3,ij}\pounds_{ij,y}^2+\vartheta_{4,ij}\pounds_{ij,z}^2)(t_{ij}^{k+1}-t_{ij}^k)^2\\
&\qquad\ge\mathcal{Q}_{ij}(t0)e^{-(\vartheta_{1,ij}+\frac{\vartheta_{2,ij}}{\vartheta_{5,ij}})t_{ij}^{k+1}}. \qquad (81)
\end{aligned}
$$

Thus, the estimated lower bound of $t_{ij}^{k+1}-t_{ij}^k$ can be obtained form (81), that is

$$
\sqrt{\frac{\mathcal{Q}_{ij}(t0)}{\vartheta_{5,ij}(\vartheta_{3,ij}\pounds_{ij,y}^2+\vartheta_{4,ij}\pounds_{ij,z}^2)}}e^{-\frac{1}{2}(\vartheta_{1,ij}+\frac{\vartheta_{2,ij}}{\vartheta_{5,ij}})t_{ij}^{k+1}}. \qquad (82)
$$

REFERENCES

[1] A. Q. Huang, M. L. Crow, and G. T. Heydt, et al., "The future renewable electric energy delivery and management(FREEDM) system: The energy internet," *Proc. IEEE*, vol. 99, no. 1, pp. 133-148, Nov. 2011.

[2] T. Li, R. Huang, and L. Chen, et al., "Compression of uncertain trajectories in road networks," *Proc. VLDB Endowment,* vol. 13, no. 7, pp. 1050-1063, 2020.

[3] R. Wang, Q. Sun, and D. Ma, et al., "Line impedance cooperative stability region identification method for grid-tied inverters under weak grids," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 2856-2866, 2020.

[4] R. Wang, Q. Sun, and P. Zhang, et al, "Reduced-order transfer function model of the droop-controlled inverter via Jordan continued-fraction expansion," *IEEE Trans. Energy Convers*, DOI: 10.1109/TEC.2020.2980033, 2020.

[5] J. Zhou, Y. Xu, and H. Sun, "Distributed power management for networked AC/DC microgrids with unbalanced microgrids," *IEEE Trans. Ind. Inf.*, DOI:10.1109/TII.2019.2925133, 2019.

[6] H. Hua, Y. Qin, and Hao C, et al., "Optimal energy management strategies for energy Internet via deep reinforcement learning approach," *Appl. Energy,* vol,239, pp. 598-609, 2019.

[7] C. E. Lin, and G. L. Viviani, "Hierarchical economic dispatch for piecewise quadratic cost functions," *IEEE Trans. Power App. Syst.,* vol. 103, no.6, pp. 1170-1175, 1984.

[8] C. E. Lin, S. T. Chen, and C. L. Huang, "A direct Newton-Raphson economic dispatch," *IEEE Trans. Powers Syst.*, vol. 7, no. 3, pp. 1149-1154, 1992.
[9] AsdS. Wright, Primal-Dual Interior-Point Methods. Philadelphia, PA, USA: SIAM, 1997.
[10] M. Moeini-Aghtaie, P. Dehghanian, and M. Fotuhi-Firuzabad, et al., "Multiagent genetic algorithm: An online probabilistic view on economic dispatch of energy hubs constrained by wind availability," *IEEE Trans. Sustain. Energy,* vol. 5, no. 2, pp. 699-708, 2014.
[11] J. Sun, V. Palade, X. J. Wu, et al., "Solving the power economic dispatch problem with generator constraints by random drift particle swarm optimization," *IEEE Trans. Ind. Informat.,* vol. 10, no. 1, pp. 222-232, 2014.
[12] Y. Liang, F. Liu and S. Mei, "Distributed real-time economic dispatch in smart grids: a state-based potential game approach," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4194-4208, 2018.
[13] H. Pourbabak, J. Luo, T. Chen and W. Su, "A novel consensus-based distributed algorithm for economic dispatch based on local estimation of power mismatch," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5930-5942, 2018.
[14] S. Yang, S. Tan, and J. Xu, "Consensus based approach for economic dispatch problem in a smart grid," *IEEE Trans. Powers Syst.*, vol. 28, no. 4, pp. 4416-4426, Nov. 2013.
[15] H. Xing, Y. Mou, M. Fu, et al., "Distributed bisection method for economic power dispatch in smart grid," *IEEE Trans. Powers Syst.*, vol. 30, no. 6, pp. 3024-3035, Nov. 2015.
[16] F. Guo, C. Wen, and J. Mao, et al., "Distributed economic dispatch for smart grids with random wind power," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1572-1583, May. 2016.
[17] T. Xu, W. Wu, and W. Zheng, et al., "Fully distributed quasi-Newton multi-area dynamic economic dispatch method for active distribution networks," *IEEE Trans. Power Syst.,* vol. 33, no. 4, pp. 4253-4263, 2018.
[18] W. Lu, M. Liu, and S. Lin, et al., "Fully decentralized optimal power flow of multi-area interconnected power systems based on distributed interior point method," *IEEE Trans. Power Syst.,* vol. 33, no. 1, pp.901-910, 2018.
[19] W. Lu, M. Liu, and Lin S, et al., "Incremental-oriented ADMM for distributed optimal power flow with discrete variables in distribution networks," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6320-6331, 2019.
[20] G. Chen and Z. Zhao., "Delay effects on consensus-based distributed economic dispatch algorithm in microgrid," *IEEE Trans. Power Syst.,* vol. 33, no. 1, pp. 602-612, Jan. 2018.
[21] B. Huang, L. Liu, and H. Zhang, et al., "Distributed optimal economic dispatch for microgrids considering communication delays," *IEEE Trans. Syst., Man, Cybern., Syst.,* DOI:10.1109/TSMC.2019.2900722, 2019.
[22] W. Ma, J. Wang, and V. Gupta, et al., "Distributed energy management for networked microgrids using online ADMM with regret," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 847-856, 2018.
[23] J. Zeng, Q. Wang, and J. Liu, et al., "A potential game approach to distributed operational optimization for microgrid energy management with renewable energy and demand response," *IEEE Trans. Ind. Electron.,* vol. 66, no. 6, pp. 4479-4489, 2019.
[24] J. Duan and M. Chow. "A resilient consensus-based distributed energy management algorithm against data integrity attacks," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 4729-4740, 2018.
[25] J. Duan and M. Chow, "A novel data integrity attack on consensus-based distributed energy management algorithm using local information," *IEEE Trans. Ind. Inf.*, vol. 15, no. 3, pp. 1544-1553, 2019.
[26] Z. Wu, W. Gao, and T. Gao, et al. , "State-of-the-art review on frequency response of wind power plants in power systems," *J. Modern Power Syst. Clean Energy* , vol. 6, no. 1, pp. 1-16, 2018.
[27] Q. Li, D. W. Gao, and H. Zhang, et al., "Consensus-based distributed economic dispatch control method in power systems," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 941-954, 2019.
[28] L. Bai, M. Ye, and C. Sun, et al., "Distributed economic dispatch control via saddle point dynamics and consensus algorithms," *IEEE Trans. Control Syst. Technol.,* vol. 27, no. 2, pp. 898-905, 2019.
[29] C. Li, X. Yu, and W. Yu, et al., "Distributed event-triggered scheme for economic dispatch in smart grids," *IEEE Trans. Ind. Inf.*, vol.12 no. 4, pp. 1775-1785, 2015.
[30] T. Zhao, Z. Li, and Z. Ding, "Consensus-based distributed optimal energy management with less communication in a microgrid," *IEEE Trans. Ind. Inf.*, vol. 15, no. 6, pp. 3356-3367, 2019.
[31] H. Zhang, Y. Li, and D. W. Gao, et al., "Distributed optimal energy management for energy internet," *IEEE Trans. Ind. Inf.*, vol. 13, no. 6, pp. 3081-3097, 2017.
[32] W. Huang, N. Zhang, and J. Yang, et al., "Optimal configuration planning of multi-energy systems considering distributed renewable energy," *IEEE Trans. Smart Grid,* vol. 10, no. 2, pp. 1452-1464, 2019.
[33] E. A. Martinez Cesena and P. Mancarella, "Energy systems integration in smart districts: robust optimisation of multi-energy flows in integrated electricity, heat and gas networks," *IEEE Trans. Smart Grid,* vol. 10, no. 1, pp. 1122-1131, 2019.
[34] Y. Li, H. Zhang, and X. Liang, et al., "Event-triggered based distributed cooperative energy management for multienergy systems," *IEEE Trans.Ind. Inf.*, vol. 15, no. 14, pp. 2008-2022, 2019.
[35] X. Le, S. Chen, and F. Li, et al., "Distributed neurodynamic optimization for energy internet management," *IEEE Trans. Syst., Man, Cybern., Syst.,* DOI: 10.1109/TSMC.2019.2898551,2019.
[36] Q. Sun, R. Fan, and Y. Li, et al. "A distributed double-consensus algorithm for residential We-Energy," *IEEE Trans. Ind. Inf.*, vol.15, no. 8, pp. 4830-4842, 2019.
[37] E. Foruzan, L. K. Soh, and S. Asgarpoor, "Reinforcement learning approach for optimal distributed energy management in a microgrid," *IEEE Trans. Power Syst.,* vol. 33, no. 5, pp. 5749-5758, 2018.
[38] J. Qiu, Z. Y. Dong, and J. H. Zhao, et al., "Multi-stage flexible expansion co-planning under uncertainties in a combined electricity and gas market," *IEEE Trans. Power Syst.,* vol. 30, no. 4, pp. 2119-2129, 2015.
[39] A. S. Zamzam, N. D. Sidiropoulos, and E. DallAnese, "Beyond relaxation and Newton-Raphson: Solving AC OPF for multi-phase systems with renewables," *IEEE Trans. Smart Grid,* vol. 9, no. 5, pp. 3966-3975, 2018.
[40] B. Brogliato, A. Daniilidis, C. Lemarchal, and V. Acary, "On the equivalence between complementarity systems, projected systems and differential inclusions," *Syst. Control Lett.,* vol. 55, pp. 45-51, 2006.
[41] A. P. Ruszczynski and A. Ruszczynski, *Nonlinear optimization*, Princeton university press, 2006.
[42] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2001.

Yushuai Li (M'19) received the B. S. degree in electrical engineering and automation, and the Ph.D degrees in control theory and control engineering from the Northeastern University, Shenyang, China, in 2014 and 2019, respectively. He is currently a postdoctoral research scholar in the Department of Electrical and Computer Engineering, University of Denver, Colorado, USA.

His current research interests include distributed control and optimization, machine learning with applications in microgrids, smart grid and Energy Internet.

David Wenzhong Gao (S'00-M'02-SM'03) received his M.S. and Ph.D. degrees in electrical and computer engineering, specializing in electric power engineering, from Georgia Institute of Technology, Atlanta, USA, in 1999 and 2002, respectively. He is now with the Department of Electrical and Computer Engineering, University of Denver, Colorado, USA. His current teaching and research interests include renewable energy and distributed generation, microgrid, smart grid, power system protection, power electronics applications in power systems, power system modeling and simulation, and hybrid electric propulsion systems. He is an Associate Editor for IEEE Journal of Emerging and Selected Topics in Power Electronics. He was an editor of IEEE Transactions on Sustainable Energy. He is the General Chair for the 48th North American Power Symposium (NAPS 2016) and the IEEE Symposium on Power Electronics and Machines in Wind Applications (PEMWA 2012).

Wei Gao (S17) received his bachelors degree in automation from Hebei University of Technology, Tianjin, China, in 2017. He is studying in the Department of Electrical and Computer Engineering for his Ph.D. degree at University of Denver, Denver, Colorado, USA. His research interests are microgrid control, renewable energy, and power system stability.

Huaguang Zhang (M'03-SM'04-F'14) received the B.S. degree and the M.S. degree in control engineering from Northeast Dianli University of China, Jilin City, China, in 1982 and 1985, respectively. He received the Ph.D. degree in thermal power engineering and automation from Southeast University, Nanjing, China, in 1991. He joined the Department of Automatic Control, Northeastern University, Shenyang, China, in 1992, as a Postdoctoral Fellow for two years. Since 1994, he has been a Professor and Head of the Institute of Electric Automation, School of Information Science and Engineering, Northeastern University, Shenyang, China. His main research interests are fuzzy control, stochastic system control, neural networks based control, nonlinear control, and their applications. He has authored and coauthored over 280 journal and conference papers, six monographs and co-invented 90 patents.

Dr. Zhang is the fellow of IEEE, the E-letter Chair of IEEE CIS Society, the former Chair of the Adaptive Dynamic Programming & Reinforcement Learning Technical Committee on IEEE Computational Intelligence Society. He is an Associate Editor of AUTOMATICA , IEEE TRANSACTIONS ON NEURAL NETWORKS, IEEE TRANSACTIONS ON CYBERNETICS, and NEUROCOMPUTING, respectively. He was an Associate Editor of IEEE TRANSACTIONS ON FUZZY SYSTEMS (2008-2013). He was awarded the Outstanding Youth Science Foundation Award from the National Natural Science Foundation Committee of China in 2003. He was named the Cheung Kong Scholar by the Education Ministry of China in 2005. He is a recipient of the IEEE Transactions on Neural Networks 2012 Outstanding Paper Award.

Jianguo Zhou (M'19) received the B.S. degree in automation and the M.S. and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, China, in 2011, 2013, and 2018, respectively. Since 2018, he has been a Postdoctoral Researcher with the Tsinghua-Berkeley Shenzhen Institute (TBSI), Tsinghua Shenzhen International Graduate School (TsinghuaSIGS), Tsinghua University, Shenzhen, China.

His current research interests include distributed control and optimization with applications in microgrids, energy Internet and integrated energy systems.