# A Distributed Dynamic Event-Triggered Algorithm With Linear Convergence Rate for the Economic Dispatch Problem

Ziwei Dong, Shuai Mao, Matjaž Perc, Wei Du, and Yang Tang, *Senior Member, IEEE*

***Abstract*—With the rapid development of the scale of distributed energy resources, the implementation of distributed algorithms imposes ever-increasing requirements on communication resources. With the aim of reducing the communication burden required to solve the economic dispatch problem, we here consider three key aspects, namely the amount of information exchanged per broadcast, the broadcast frequency per iteration, and the number of iterations needed to achieve a certain accuracy. The proposed primal-dual based algorithm is integrated with a discrete dynamic event-triggered scheme and enjoys significant advantages in all three mentioned aspects. We prove that the proposed algorithm converges to the optimal point at a linear convergence rate for suitable operating parameters and for cost functions that are strongly convex and smooth. We confirm the effectiveness and demonstrate the advantages of our approach by means of a set of simulation experiments.**

***Index Terms*—Distributed economic dispatch problem, linear convergence rate, dynamic event-triggered scheme, distributed optimization, power system.**

## I. Introduction

OVER the past few decades, the power system, have been undergoing transformation with the help of emerging technologies [1], [2], [3]. As the power system continues to modernize, systems composed of conventional generation power plants are gradually replaced by systems composed of a large number of distributed energy resources [4]. In the pursuit of better coordinations of the distributed energy resources, the economic dispatch problem (EDP) arises, abstracted as minimizing the total cost of electric power generation under output constraints and the total generation demand [5], [6]. To deal with EDP, an effective but traditional solution is centralized optimization, which suffers from performance limitations, such as single point of failure, high communication requirement, substantial computation burden and limited flexibility and scalability [7]. To overcome performance limitations, distributed optimization strategies have been proposed, which take advantages of local computation and communication between agents to realize the global optimization [7]. In the case where output constraints are excluded, the central-free algorithm (CFA) in [8], [9] works. However, for cases with output constraints, CFA cannot be directly applied because the balance between demand and supply will be damaged by operations, which are utilized to deal with constraints. To deal with EDP distributedly, the barrier-based method is employed to transform EDP to an unconstrained case [10]. Another mature method is to apply primal-dual theory and transform the original EDP to the dual problem [11], [12], which is a distributed optimization problem.

Manuscript received 23 March 2021; revised 7 April 2022; accepted 2 July 2022. Date of publication 25 October 2022; date of current version 6 January 2023. This work was supported in part by the National Natural Science Foundation of China, Basic Science Center Program under Grant 61988101, in part by the International (Regional) Cooperation and Exchange Project under Grant 1720106008, in part by the National Natural Science Foundation of China under Grant 62173144, in part by the Program of Shanghai Academic Research Leader under Grant 20XD1401300, and in part by Shanghai AI Lab and the Slovenian Research Agency under Grants P1-0403 and J1-2457. Recommended for acceptance by Prof. Tie Qiu. *(Corresponding author: Yang Tang)*

Ziwei Dong, Shuai Mao, Wei Du, and Yang Tang are with the Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China (e-mail: ZiweiDong1996@126.com; mshecust@163.com; duwei0203@gmail.com; tangtany@gmail.com).

Matjaž Perc is with the Faculty of Natural Sciences and Mathematics, University of Maribor, 2000 Maribor, Slovenia, also with Alma Mater Europaea, 2000 Maribor, Slovenia, and also with Department of Physics, Kyung Hee University, Dongdaemun-gu, Seoul, Republic of Korea (e-mail: matjaz.perc@gmail.com).

Digital Object Identifier 10.1109/TNSE.2022.3216572

### A. Motivations

Nowadays, some practical conditions have been taken into consideration, such as random noises in gradient and communication delays [13], unpredictable communication failures [14] and unbalanced directed communication topology [15]. However, few researches consider communication burden. Due to an increasing number of distributed generators that are integrated into the main grid, the implementation of distributed algorithms has higher requirements on communication resources. Considering the scale of smart grid and its communication network, the tremendous amount of exchanged information may exhaust communication resources and even lead to failure of convergence [16]. The impact of communication burden should be taken into the consideration when the algorithms are applied into practical cases [16].

Three aspects are considered and discussed as below:

- the number of iteration for achieving certain accuracy;
- the broadcast frequency per iteration;
- the amount of information exchanged per broadcast.

### B. Literature Review

First, the number of iteration for achieving certain accuracy is considered. Commonly discussed in both centralized and distributed optimization, it is directly reflected by the convergence rate of the algorithm, where the faster is preferred since the higher rate may mean fewer iterations and less communication time.

The aforementioned algorithms with diminishing step sizes can only guarantee the sublinear convergence rate or their exact convergence rate cannot be obtained. Employing fixed step sizes, the algorithm in [10] can asymptotically converge, but even the accelerated version can only improve the theoretical convergence rate to $\mathcal{O}(\frac{1}{k^2})$ (a sublinear rate). As shown in the researches of the dual problem, distributed optimization problem, diminishing step sizes influence the convergence rate [17], [18], [19] and the linear convergence rate can be obtained by some algorithms [18], [19] employing fixed step sizes. Moreover, it further reveals the difficulty for the meticulous design and analysis of linear convergent algorithms, i.e. if diminishing step sizes are simply replaced by fixed ones, the algorithm can not reach the exact optimal point [20]. As to EDP, counterparts of [18], [19] are proposed in [21], [22], lacking theoretical analysis for the linear convergence rates. A distributed ADMM-like method, called DPDA-D in [23], also includes fixed step sizes and lacks the exact convergence rate. Recently, there exist some successful attempts. Algorithms based on DIGing [18] are proposed in [24], [25], [26] and their linear convergent rates are proved along the small gain theorem used in [18].

Second, the aspect of the broadcast frequency per iteration is considered. The event-triggered scheme is widely used in multi-agent systems to reduce unnecessary communication per iteration [27], [28] by restricting transmissions to the case when the event-triggered condition is satisfied. It is verified that with a properly designed event-triggered scheme, the communication frequency decreases while maintaining key properties [10] and the core is the design of event-triggered condition [29], [30].

Some researches integrate the event-triggered scheme into linearly convergent algorithms. Based on the algorithm in [26], event-triggered versions and an accelerated one are provided in [31] and [32] respectively. Note that only static event-triggered scheme is integrated in the aforementioned algorithms. As one of future research prospect in [28], the dynamic event-triggered scheme is preferred since it employs an internal dynamic variable as the error threshold, which may increase at some steps and thus can allow for larger trigger intervals. It is theoretically proved in [33] that the lower bound of interval obtained in the dynamic event-triggered scheme must be larger than or equal to the one obtained in a traditional static event-triggered scheme.

Recently the dynamic event-triggered scheme has been integrated into continuous-time algorithms to deal with various problems in multi-agent systems [34], [35], [36], [37], [38]. However, few researches try to integrate the dynamic event-triggered scheme into distributed algorithms for the consensus optimization problem and EDP, probably because the internal variable, always larger than some $w^t$ used in [31], [32], brings difficulties into the proof along the small gain theorem in [31], [32], even though it may provide a looser threshold [39]. It is unknown whether the linear convergence rate can be maintained or not when a dynamic event-triggered scheme is integrated.

### C. Statement of Contributions

Motivated by the practical need for lessening communication burden in smart grid, this paper proposes a distributed algorithm with dynamic event-triggered scheme and provides a theoretical proof for its linear convergence rate when the objective function is Lipschitz smooth and strongly convex.

- The proposed event-triggered algorithm relieves the burden of communication from three aspects. Compared with asymptotically convergent algorithms in [10], [13], [14] that also broadcast one variable, the proposed algorithm guarantees faster convergence rate and less iterations. The integration of dynamic event-triggered scheme helps reduce the communication frequency. Compared with linearly convergent algorithms in [24], [25], [26], [31], [32], the proposed algorithm is carefully designed to use only half amount of exchanged information per broadcast, which touches on the third aspect, the amount of information exchanged per broadcast process.
- Furthermore, note that our work is a successful attempt to integrate dynamic event-triggered scheme into algorithms for EDP in smart grid. It is proved that the dynamic event-triggered scheme always works when the step sizes are chosen in the given ranges. Moreover, our work theoretically shows that the linear convergence rate can be maintained when a dynamic event-triggered scheme is integrated.
- The most similar counterpart in distributed consensus-based optimization problem is the one proposed in [40], which is a discrete-time counterpart of the continuous time algorithms in [41]. However for linear convergence, a matrix condition containing the step size is given in [40] rather than a specific range of step size. Furthermore, it is not clear whether there exists such step sizes that the matrix condition in [40] holds. In contrast, the proposed algorithm in this paper possesses flexibility and the analysis gives a chosen range for step size to guarantee the linear convergence.

### D. Paper Organization

Section II introduces notations, communication topology and the formulation of EDP. Section III reformulates EDP and proposes time-triggered and dynamic event-triggered algorithm. Theoretical results are shown in Section IV and verified by a simulation case in Section V. Section VI concludes the paper.

## II. Preliminaries

### A. Notations

Unless otherwise stated, subscripts and superscripts are used to distinguish nodes and to mark the time respectively. Denote $\mathbf{1}$, $\mathbf{0}$ as all one column vectors and zero vectors. For two vectors $\mathbf{x}$ and $\mathbf{y}$, their inner product is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$. Denote $I$ as the identity matrix. For a matrix $A$, $A^\top$, $\lambda_{max}(A)$ and $\sigma_{max}(A)$ represent its transpose matrix, the largest eigenvalue and the largest singular value respectively.

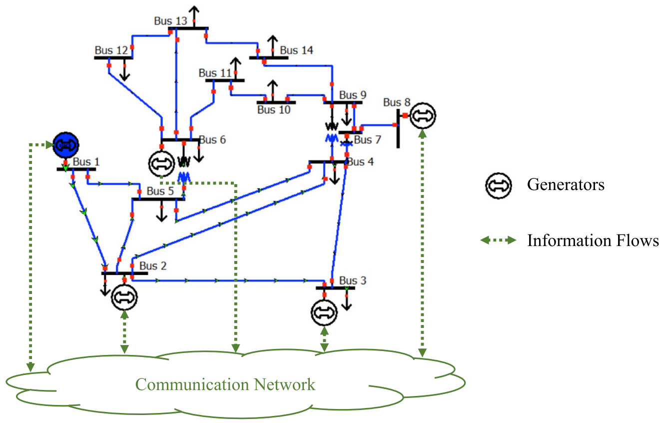


Fig. 1. EDP in the IEEE 14-bus system [42]. The IEEE 14-bus test case represents a simple approximation of the American Electric Power system as of February 1962, which has 14 buses, 5 generators and 11 loads. EDP arises among the generators to minimize the total cost of electric power generation while meeting the total generation demand.

### B. Communication Topology

In this paper, each generator is modeled as an agent that can communicate with some of others, as shown in Fig. 1. The communication topology in EDP is constructed as an undirected network $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, W\}$ composed of $N$ agents, where $\mathcal{V}$, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ and $W$ represent the set of agents, the set of communication links and communication weight matrix. $W_{ij} = W_{ij} > 0$ if there exists a communication link between agents $i$ and $j$, and otherwise $W_{ij} = 0$. Denote $\mathcal{N}_i = \{j : W_{ij} > 0, j \in \mathcal{E}\}$ as the neighbour set of agent $i$.

*Assumption 1:* The communication graph $\mathcal{G}$ is connected.

*Assumption 2:* 1) For all $i$, $W_{ii} > 0$;

2) For all $W_{ij} \neq 0$, there exists a constant $c > 0$ such that $W_{ij} > c$;

3) $W$ is doubly stochastic, i.e., $W\mathbf{1} = \mathbf{1}, \mathbf{1}^\top W = \mathbf{1}^\top$.

The Laplace matrix $L = I - W$ is used in the remainder of this paper. It is easy to get that $\mathbf{1}^\top L = \mathbf{0}$ and $L^\top \mathbf{1} = \mathbf{0}$.

### C. Problem Formulation

Consider a power system including $N$ distributed generators (nodes), each of which is associated with a capacity limit and a local private objective function to measure the production cost of its power output. The task of EDP is to coordinate all distributed generators to jointly minimize the total generation cost within capacity limits while meeting the total power demand, i.e.,

$$\begin{aligned} \min_{x_i} \quad & \sum_{i=1}^{N} F_i(x_i) \\ \text{subject to} \quad & \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} d_i = d \\ & x_i \in \mathcal{X}_i \end{aligned} \tag{1}$$

where $x_i$, $\mathcal{X}_i \subseteq \mathbb{R}^m$ and $F_i(x_i) : \mathcal{X}_i \to \mathbb{R}$ are the local decision vector (power output), constrained set (capacity limit) and local private cost function of generator $i$, respectively. For each agent $i$, only virtual local demand $d_i$ is available. The sum $d = \sum_{i=1}^{N} d_i$ represents the total demand of the system.

Without loss of generality, we assume, that the variable $x$ has only one dimension, i.e., $m = 1$.

*Assumption 3:* The constrained set is nonempty and convex.

*Assumption 4:* For all $i \in \mathcal{V}$, the local cost function $F_i(x_i) : \mathcal{X}_i \to \mathbb{R}$ is differentiable and has a Lipschitz continuous gradient, i.e., there exists a Lipschitz constant $l \in (0, +\infty)$ such that

$$\|\nabla F_i(x) - \nabla F_i(y)\|_F \leq l \|x - y\|_F$$

for any $x, y \in \mathcal{X}_i$.

*Assumption 5:* For all $i \in \mathcal{V}$, the local cost function $F_i(x_i) : \mathcal{X}_i \to \mathbb{R}$ is differentiable and $\mu$-strongly convex, i.e.,

$$F_i(y) - F_i(x) \geq \langle \nabla F_i(x), y - x \rangle + \frac{\mu}{2} \|x - y\|_F^2$$

for any $x, y \in \mathcal{X}_i$.

*Assumption 6:* The optimal solution set of (1) is nonempty.

These assumptions are fairly standard and commonly used in related work, such as [10], [18], [19].

## III. Algorithm Design

In this section, we first reformulate EDP for the convenience of algorithm design, which is similar to [15]. Then we propose a time-triggered algorithm and integrate dynamic event-triggered scheme into it.

### A. Problem Reformulation

The Lagrange function of the EDP (1) is

$$\mathcal{L}(\mathbf{x}, \lambda) = \sum_{i=1}^{N} F_i(x_i) + \lambda \left( \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} d_i \right).$$

For a convex function $f : \mathcal{S} \to \mathbb{R}$, denote its conjugate function as $f^\perp(y) = \sup_{x \in \mathcal{S}} (y^\top x - f(x))$ for $y \in \mathbb{R}^{|S| \times 1}$.

Then the dual function of the EDP (1) is

$$\begin{aligned} D(\lambda) &= \min_{x_i \in \mathcal{X}_i} \left( \sum_{i=1}^{N} F_i(x_i) + \lambda \left( \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} d_i \right) \right) \\ &= -\lambda \sum_{i=1}^{N} d_i - \sum_{i=1}^{N} \sup_{x_i \in \mathcal{X}_i} (-F_i(x_i) - \lambda x_i) \\ &= \sum_{i=1}^{N} (-\lambda d_i - F_i^\perp(-\lambda)). \end{aligned}$$

Define $H_i(\lambda) = \lambda d_i + F_i^\perp(-\lambda)$ and $H(\lambda) = \sum_{i=1}^{N} H_i(\lambda) = -D(\lambda)$. Then we can get the dual problem of (1) is

$$\min_{\lambda \in \mathbb{R}} H(\lambda). \tag{2}$$

According to the definition of the conjugate function, the gradient of $H_i(\lambda)$ is $\nabla H_i(\lambda) = d_i - x_i$.

*Lemma 1:* [43] For a function $f : \mathcal{X} \to \mathbb{R}$ which is differentiable, $\mu$-strongly convex and has a $l$-Lipschitz continuous gradient, its conjugate function $f^{\perp}$ is differentiable, $\frac{1}{l}$-strongly convex and has a $\frac{1}{\mu}$-Lipschitz continuous gradient.

Until now, we have given the dual problem of the EDP, happened to be a distributed consensus-based optimization problem. When Assumptions 4 and 5 hold, Lemma 1 gives that the objective function in (2) has similar properties with that in (1) with different parameters.

### B. Time-Triggered Algorithm

Since we have given the dual problem (2) in the preceding subsection, here we consider solving a local dual problem followed by a subroutine that can distributedly find a consensus on the dual optimal point. Define $\mathbf{x}^k = [x_1^k, x_2^k, \ldots, x_N^k]^\top$ and $\mathbf{d} = [d_1, d_2, \ldots, d_N]^\top$. $\boldsymbol{\lambda}^k$, $\mathbf{z}^k$ and $\mathbf{g}^k$ are defined in a similar way. Then the primal-dual algorithm is designed as following:

$$\begin{aligned}\mathbf{x}^k &= \underset{\mathbf{x} \in \prod \mathcal{X}_i}{\arg\min} \left\{ \sum_{i=1}^{N} \left( F_i(\mathbf{x}_i) + \lambda_i^k \mathbf{x}_i \right) \right\};\\ \mathbf{g}^k &= -\mathbf{x}^k + \mathbf{d};\\ \boldsymbol{\lambda}^{k+1} &= \boldsymbol{\lambda}^k - \beta L \boldsymbol{\lambda}^k - \alpha\beta \mathbf{g}^k - \gamma\beta \mathbf{z}^k;\\ \mathbf{z}^{k+1} &= \mathbf{z}^k + \beta L \boldsymbol{\lambda}^k. \end{aligned} \tag{3}$$

where $\mathbf{z}^0 = 0$.

*Remark 1:* To implement the proposed algorithm, during time $t \in [k, k+1]$, each distributed generator $i$ maintains the original variable $x_i^k$, dual variable $\lambda_i^k$ and the compensation variable $z_i^k$, and then transmits dual variable $\lambda_i^k$ to its neighbours. Compared with [24], [25], [26], [34], [35], [36], [37], our algorithm requires half amount of exchanged variables per iteration while the amount of maintained variables is the same.

### C. Distributed Dynamic Event-Trigger Scheme

To further reduce the communication burden, we integrate event-triggered scheme into our algorithm.

Denote $\{k_{i,t}\}$ as the event-triggered time sequence of node $i$. Denote $\hat{\lambda}_i^k$ as the information used by the neighbours of node $i$ at time $k$. Node $i$ sends its latest estimation $\lambda_i^{k_{i,t}} = \lambda_i^k$ to its neighbours only at the event-triggered time $k = k_{i,t}$. Consequently, only at the event-triggered time $k = k_{i,t}$, neighbours of agent $i$ can get exact information of $\lambda_i^k$, while during time $k \in [k_{i,t}+1, k_{i,t+1}-1]$, neighbours of node $i$ can only use outdated information offered by node $i$ at time $k_{i,t}$, i.e., $\hat{\lambda}_i^k = \lambda_i^{k_{i,t}}$.

Under event-triggered scheme, our algorithm is designed as

$$\begin{aligned}\mathbf{x}^k &= \underset{\mathbf{x} \in \prod \mathcal{X}_i}{\arg\min} \left\{ \sum_{i=1}^{N} \left( F_i(\mathbf{x}_i) + \lambda_i^k \mathbf{x}_i \right) \right\};\\ \mathbf{g}^k &= -\mathbf{x}^k + \mathbf{d};\\ \boldsymbol{\lambda}^{k+1} &= \hat{\boldsymbol{\lambda}}^k - \beta L \hat{\boldsymbol{\lambda}}^k - \alpha\beta \mathbf{g}^k - \gamma\beta \mathbf{z}^k;\\ \mathbf{z}^{k+1} &= \mathbf{z}^k + \beta L \hat{\boldsymbol{\lambda}}^k. \end{aligned} \tag{4}$$

*Remark 2:* The event-triggered scheme means that each node inevitably has to maintain the latest received information of its neighbours for next iterations. Compared with [34], [35], [36], [37], as our algorithm requires half the amount of exchanged variables for updates, the amount of storage for neighbours' estimations is also reduced by half.

*Remark 3:* To guarantee the effectiveness of the algorithm, the broadcasted estimation $\hat{\lambda}_i^k$, rather than the latest estimation $\lambda_i^k$ is used in the update iteration of $\lambda_i^{k+1}$. One benefit is that if we rearrange the computation order, there is no need to maintain $\lambda_i^k$ any longer, saving a part of storage further.

As stated in [28], the key task in designing the event-triggered scheme mainly lies in the event-triggered condition.

To provide a better tradeoff between preserving convergence performance and reducing communication consumption, dynamic event-triggered scheme is used in this paper, even though it may make the convergence analysis more difficult [28].

For each node $i = 1, 2, \ldots, N$, denote the error as $e_i^k = \hat{\lambda}_i^k - \lambda_i^k$, an auxiliary variable as $p_i^k = -\frac{1}{2}\sum_{j \in \mathcal{N}_i} L_{ij} \|\hat{\lambda}_j^k - \hat{\lambda}_i^k\|^2 \geq 0$ and an internal dynamic variable as $\chi_i^k$ satisfying

$$\chi_i^{k+1} = (1-\tau_i)\chi_i^k - \delta_i(\|e_i^k\|^2 - \eta_i p_i^{k+1}), \chi_i^0 > 0. \tag{5}$$

Then, the event-triggered time for node $i$ is designed as

$$k_{i,t+1} = \min_k \{\theta_i \|e_i^k\|^2 \geq \chi_i^k, k > k_{i,t}\}. \tag{6}$$

This means that $\theta_i \|e_i^k\|^2 < \chi_i^k$ always holds for all $k$.

The proposed algorithm with dynamic event-triggered scheme from the view of node $i$ is summarized in Algorithm 1. Note that for each node $i$, we leave out the hat and the superscript of neighbours' information (representing inaccuracy and time respectively) as we only need to save and use the latest received estimates from each neighbour and there is no need to number them.

*Remark 4:* The advantages of dynamic event-triggered scheme against static event-triggered scheme used in [31], [32] is clear now. In detail, the researches [31], [32] employ $\|e_i^k\|^2 \geq cw^k$ for some $c > 0$ and $w \in (0,1)$ as a triggered condition. For any $c > 0$ and $w \in (0,1)$, we can simply set $\chi_i^0 > c$ and $1 - \tau - \frac{\delta_i}{\theta_i} = w$ and get that the upper bound for $\|e_i^{k+1}\|^2$ in the proposed event-triggered condition is always larger than that of [31], [32], i.e., $\chi_i^{k+1} > (1 - \tau - \frac{\delta_i}{\theta_i})\chi_i^k + \delta_i \eta_i p_i^{k+1} \geq w\chi_i^k \geq \cdots > cw^{k+1}$. Hence, the dynamic event-triggered scheme can provide larger intervals and further reduce the communication frequency.

## IV. Main Results

In this section, we first discuss the relationship between the fixed point of the proposed algorithm and the optimal point of EDP. Then we give analysis for linear convergence of the proposed algorithm and the effectiveness of dynamic event-triggered scheme.

**Algorithm 1:** for each node $i$.

**Step 1: Initialization**

• Choose arbitrary $\lambda_i^0$ as the local estimation of dual variable and arbitrary $\chi_i^0 > 0$ as internal dynamic variable. Set the correction item $z_i^0 = 0$ and $z_i^{-1} = 0$.

• Set $k = 0$. Broadcast its own estimation $\lambda_i^k$ to neighbours and save it as $\hat{\lambda}_i$. Let the error $e_i^k = \hat{\lambda}_i - \lambda_i^k = 0$.

• Receive and save $\lambda_j$ from all neighbours $j$.

**Step 2: Computation and Communication**

1: **for** time $k = 0, 1, \ldots, K$ **do**

2: Compute along (4) from the view of each node $i$:

$$\begin{aligned} x_i^k =& \arg\min_{x_i \in \mathcal{X}_i} \{F_i(x_i) + \lambda_i^k x_i\}; \\ g_i^k =& - x_i^k + d_i; \\ \lambda_i^{k+1} =& \hat{\lambda}_i + \beta \sum_{j \in \mathcal{N}_i} W_{ij}(\lambda_j - \hat{\lambda}_i) - \alpha\beta g_i^k - \gamma\beta z_i^k; \\ z_i^{k+1} =& z_i^k - \beta \sum_{j \in \mathcal{N}_i} W_{ij}(\lambda_j - \hat{\lambda}_i). \end{aligned}$$

3: Keep receiving and updating the neighbours' state as $\lambda_j = \lambda_j^{k+1}$. Update the internal dynamic variable of event-triggered scheme along (5) and check whether the event-triggered condition holds. That is:

$$\begin{aligned} e_i^{k+1} =& \hat{\boldsymbol{\lambda}}_i - \boldsymbol{\lambda}_i^{k+1}; \\ p_i^{k+1} =& -\frac{1}{2} \sum_{j \in \mathcal{N}_i} L_{ij} \|\hat{\lambda}_j - \hat{\lambda}_i\|^2; \\ \chi_i^{k+1} =& (1 - \tau_i)\chi_i^k - \delta_i(\|e_i^k\|^2 - \eta_i p_i^{k+1}). \end{aligned}$$

**if** $\theta_i \|e_i^{k+1}\|^2 \geq \chi_i^{k+1}$ **then**

Broadcast $\lambda_i^{k+1}$ to neighbours.

Update $\hat{\lambda}_i = \lambda_i^{k+1}$ and $e_i^{k+1} = 0$.

**end if**

4: **end for**

**Step 3: Output of the local estimation**

• Compute

$$x_i^{K+1} = \arg\min_{x_i \in \mathcal{X}_i} \{F_i(x_i) + \lambda_i^{K+1} x_i\}.$$

• $x_i^{K+1}$ is the local estimation of the optimal point of the problem (1) after $K$ steps.

### A. Optimal Point

In this part, we show that the fixed-point of our algorithm is just the optimal point of (1). Define $\mathbf{x}^*$ and $\lambda^*$ as the optimal solutions of the primal problemd and the dual problem.

Suppose that the proposed algorithm (3) converges to a fixed point $(\mathbf{x}^\infty, \boldsymbol{\lambda}^\infty, \mathbf{z}^\infty)$. Then according to the algorithm, it is clear that the fixed point satisfies that

$$x_i^\infty = \arg\min_{x_i \in \mathcal{X}_i} \{F_i(x_i) + \lambda_i^\infty x_i\};$$

$$L\boldsymbol{\lambda}^\infty = \mathbf{0}; \quad \alpha \mathbf{g}(\boldsymbol{\lambda}^\infty) + \gamma \mathbf{z}^\infty = 0.$$

Since $\mathbf{1}^\top L = \mathbf{0}$, we have

$$\mathbf{1}^\top \mathbf{z}^{k+1} = \mathbf{1}^\top \mathbf{z}^k = \ldots = \mathbf{1}^\top \mathbf{z}^0 = 0 \text{ for any } k = 0, 1, \ldots.$$

Thus $\sum_{i=1}^N x_i^\infty - \sum_{i=1}^N d_i = \mathbf{1}^\top \mathbf{g}(\boldsymbol{\lambda}^\infty) = -\frac{\gamma}{\alpha} \mathbf{1}^\top \mathbf{z}^\infty = 0$. According to Theorem 6.2.5 in Ref.[44], $(\mathbf{x}^\infty, \boldsymbol{\lambda}^\infty)$ is the saddle point of the Lagrange function $\mathcal{L}(\mathbf{x}, \lambda)$ and $\mathbf{x}^\infty$ and $\boldsymbol{\lambda}$ are the optimal solutions to the primal and dual problems with no duality gap, i.e., $(\mathbf{x}^*, \lambda^*) = (\mathbf{x}^\infty, \boldsymbol{\lambda}^\infty)$.

Denote $\mathbf{z}^* = \mathbf{z}^\infty$ and we have $\alpha \mathbf{g}(\boldsymbol{\lambda}^*) + \gamma \mathbf{z}^* = 0$.

### B. Convergence Analysis

Define $U = \sqrt{\gamma}\beta L^{1/2}$. Then we have the following lemma for the dual problem.

*Lemma 2:* [19] Under Assumptions 2, 4, 5 and 6, $\lambda_1^* = \lambda_2^* = \cdots = \lambda_N^*$ is the optimal solution of the dual problem (2) if and only if there exists $\mathbf{b}^* = U\mathbf{a}$ for some $\mathbf{a} \in \mathbb{R}^N$ such that

$$\begin{cases} U\mathbf{b}^* + \alpha\beta \mathbf{g}(\boldsymbol{\lambda}^*) = \mathbf{0}, \\ U\boldsymbol{\lambda}^* = \mathbf{0}. \end{cases}$$

Next we will transform the proposed algorithm for the convergence analysis. After the transformation, it will be clear to see the similarity of the conditions between Lemma 2 and the fixed point of the proposed algorithm.

According to the iteration, $\mathbf{z}^k = \beta \sum_{t=0}^{k-1} L\hat{\lambda}^t$. Thus the proposed algorithm can be written as

$$\boldsymbol{\lambda}^{k+1} = \hat{\lambda}^k - \beta L \hat{\lambda}^k - \alpha\beta \mathbf{g}^k - \gamma\beta^2 \sum_{t=0}^{k-1} L\hat{\lambda}^t, \tag{7}$$

where

$$\begin{aligned} \mathbf{g}^k =& - \mathbf{x}^k + \mathbf{d}, \\ \mathbf{x}^k =& \sum_{i=1}^N \arg\min_{x_i \in \mathcal{X}_i} \{F_i(\mathbf{x}_i) + \lambda_i^k \mathbf{x}_i\}. \end{aligned}$$

Define $\mathbf{b}^k = U \sum_{t=0}^k \hat{\lambda}^t$, then $\mathbf{z}^k = \frac{1}{\gamma\beta} U \mathbf{b}^{k-1}$ and (7) can be written as

$$\boldsymbol{\lambda}^{k+1} = \hat{\lambda}^k - \beta L \hat{\lambda}^k - \alpha\beta \mathbf{g}^k - U\mathbf{b}^{k-1}. \tag{8}$$

Now we have been ready to show the main theorems.

*Theorem 1:* Suppose that Assumptions 1-6 hold. The parameters are properly chosen as below: the step size $\gamma > 0$, $\beta$ satisfying

$$\beta \in \left(0, \frac{\sigma_{\max}(L) - \sqrt{\sigma_{\max}^2(L) - 4\sigma_{\max}(L)\gamma}}{2\sigma_{\max}(L)\gamma}\right),$$

$\zeta_1 \in (0, \frac{2}{l})$ and $\alpha$ satisfying

$$\alpha \in \left(0, \frac{\zeta_1 \mu^2 (1 - \beta(1 - \gamma\beta)\sigma_{\max}(L))}{\beta}\right),$$

and event-triggered parameter $\tau_i \in (0, 1)$, $\delta_i \in (0, +\infty)$, $\eta_i \in [0, \frac{1}{A_5}]\ \kappa$ and $\theta_i$ satisfying

$$0 < \frac{A_5(1 - \tau_i)}{\delta_i} \leq \kappa < \frac{A_5}{\delta_i}, \quad \theta_i > \max\left\{\frac{A_4}{\frac{A_5}{\delta_i} - \kappa}, \frac{\delta_i}{1 - \tau_i}\right\},$$

where $A_4$ and $A_5$ are defined as

$$A_4 = \zeta_2 + \zeta_8\gamma^2\beta^4\sigma_{\max}(L) \times \left(\left(\frac{\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)} + \frac{1}{\zeta_2}\right)\frac{\zeta_4}{\zeta_4 - 1} + \frac{1}{\zeta_3}\right),$$

and

$$A_5 = \frac{\zeta_8\sigma_{\max}(I - \beta(1-\gamma\beta)L)}{\zeta_8 - 1}\left(\left(\frac{\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)} + \frac{1}{\zeta_2}\right)\frac{\zeta_4}{\zeta_4 - 1} + \frac{1}{\zeta_3}\right)$$

for some $\zeta_2$, $\zeta_3$ in $(0, +\infty)$ and $\zeta_4$, $\zeta_8$ in $(1, +\infty)$. Then there exists $\epsilon_1 > 0$, $\epsilon_2 > 0$ such that the sequence $\{\mathbf{x}^k, \boldsymbol{\lambda}^k\}$ generated by Algorithm 1 satisfies

$$\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*\|^2 \leq \left(\frac{1}{(1 + \min\{\epsilon_1, \epsilon_2\})}\right)^k C,$$
$$\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq \left(\frac{1}{(1 + \min\{\epsilon_1, \epsilon_2\})}\right)^k \frac{C}{\mu^2},$$

where the constant $C = \sigma_{\min}^{-1}(I - \beta(1-\gamma\beta)L)M^0$.

In detail,

$$0 < \epsilon_1 < \min\{\epsilon_{11}, \epsilon_{12}\}$$

and

$$0 < \epsilon_2 \leq \frac{1}{\kappa}\left(-\frac{A_4}{\theta_i} + \frac{A_5}{\delta_i} - \kappa\right),$$

where $\zeta_5 \in (1, +\infty)$,

$$\epsilon_{11} = \frac{\gamma\beta^2\lambda_{\min}(L)(\zeta_5 - 1)\left((1 - \frac{\alpha\beta}{\zeta_1\mu^2}) - \beta(1-\gamma\beta)\sigma_{\max}(L)\right)}{\zeta_4\zeta_5\epsilon_1\left(\sqrt{\sigma_{\beta,\gamma}} + \frac{\alpha\beta}{\mu}\right)^2},$$
$$\epsilon_{12} = \frac{\alpha\beta(\frac{2}{l} - \zeta_1) + 2\beta(1 - 2\gamma\beta)\sigma_{\max}(L)}{\sigma_{\beta,\gamma} + (\frac{\zeta_4\zeta_5\beta^2}{\gamma\beta^2\lambda_{\min}(L)})\left((1 - 2\gamma\beta)\sqrt{\sigma_{\max}(L)} + \frac{\alpha}{\mu}\right)^2},$$

and $\sigma_{\beta,\gamma} = \sigma_{\max}(I - \beta(1-\gamma\beta)L)$.

*Proof:* See Appendix A. ∎

*Remark 5:* With the most similar structure for distributed consensus optimization problem, [40] provides a matrix condition for the linear convergence rate property, which may not have any feasible step size. Note that here in Theorem 1, explicit ranges for parameters of the proposed dynamic event-triggered algorithm are given to guarantee the linear convergence rate.

*Remark 6:* The parameters $\alpha$, $\beta$ and $\gamma$ adjust the weight of the local gradient, other nodes' estimations and the correction item respectively. It is shown that the convergence rate has a complex relation with the parameters and no qualitative relationship can be given from the view of theoretical analysis. For the event-triggered algorithm, the parameters regarding the event-triggered scheme also influence the convergence rate. In detail, bigger $\tau_i$, $\theta_i$ and smaller $\delta_i$ tend to indicate higher convergence rate and more frequent triggering.

*Theorem 2:* Suppose that Assumptions 1-6 hold. The sequence $\{\mathbf{x}^k, \boldsymbol{\lambda}^k\}$ generated by the proposed (time-triggered) distributed algorithm (3) satisfies

$$\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*\|^2 \leq (1 + \epsilon_1)^{-k} C,$$
$$\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq (1 + \epsilon_1)^{-k} \frac{C}{\mu^2},$$

where $\epsilon_1 > 0$ and the constant $C = \sigma_{\min}^{-1}(I - \beta(1 - \gamma\beta)L)M^0$. The parameters are properly chosen as below: the step size $\gamma > 0$, $\beta$ satisfying

$$\beta \in \left(0, \frac{\sigma_{\max}(L) - \sqrt{\sigma_{\max}^2(L) - 4\sigma_{\max}(L)\gamma}}{2\sigma_{\max}(L)\gamma}\right),$$

$\zeta_1 \in (0, \frac{2}{l})$ and $\alpha$ satisfying

$$\alpha \in \left(0, \frac{\zeta_1\mu^2(1 - \beta(1-\gamma\beta)\sigma_{\max}(L))}{\beta}\right).$$

*Proof:* The proof is omitted due to its similarity to the proof for Theorem 1. ∎

### C. Effectiveness of Event-Triggered Scheme

*Theorem 3:* In the same setting of Theorem 1, the event-triggered scheme always works when the parameter $\tau_i$ is chosen as $\tau_i < \frac{\epsilon_1}{1+\epsilon_1}$.

*Proof:* See Appendix B. ∎

*Remark 7:* A successful integration of event-triggered scheme can reach a balance between the reduction of event-triggered time and the influence on convergence rate. As to the convergence rate, Theorem 1 shows that the integration of the dynamic event-triggered scheme does not influence the convergence rate greatly even if we cannot draw an exact conclusion on the decaying rate of the event-triggered bound $\boldsymbol{\chi}$. As to the triggered time, an undesirable case is that the event-triggered bound decays faster. For example, if we integrate a trigger scheme with linearly decaying bound into a sublinearly convergent algorithm, the triggered condition will always be satisfied after some time, which means the event-triggered scheme does not work anymore. Theorem 3 shows that a properly chosen parameter will promise the effectiveness of the proposed event-triggered scheme in the long run.

## V. Simulation

In this section, we illustrate theoretical results via numerical examples and demonstrate comparisons with algorithms in recent work [15], [29], [30], [31], [32].

We choose the case adopted in [31] and [32] where the economic dispatch problem in the IEEE 14 bus system with five generators is considered as described in Fig 1. The cost function of each generator is provided as

$$F_i(x_i) = a_i x_i^2 + b_i x_i, x_i^{\min} \leq x_i \leq x_i^{\max}.$$

Note that the capacity constraints are not tight at the optimal point in the case in [31], causing the problem degenerating into

TABLE I
GENERATOR PARAMETERS IN IEEE 14 BUS SYSTEM

| Bus | $a_i$(\$/MW$^2$) | $b_i$(\$/MW) | $[x_i^{\min}, x_i^{\max}]$(MW) | $x_i^*$ |
|---|---|---|---|---|
| 1 | 0.04 | 2.0 | [0,80] | 80 |
| 2 | 0.03 | 3.0 | [0,90] | 90 |
| 3 | 0.035 | 4.0 | [0,70] | 194/3 |
| 6 | 0.03 | 4.0 | [0,70] | 70 |
| 8 | 0.04 | 2.5 | [0,80] | 226/3 |

an unconstrained problem at the neighbourhood of the optimal point. To demonstrate the main results, we employ the generator parameters and power demand $\sum_{i=1}^{N} d_i = 380$MW adopted in [32], which are summarized in Table I.

## A. Performance From the View of Figures

As shown in Fig 2, the proposed algorithm can find the optimal point while the event-triggered scheme works and the nodes communicate intermittently. To vividly show the convergence rate and the performance of event-triggered scheme, comparisons are carried out. In the comparisons, the algorithms in [15], [29], [30] are slightly adjusted to solve the case of EDP with event-triggered scheme. For example, the static event-triggered scheme used in [31] is integrated into the algorithm in [15]. It is easy to get that the adjustments will not change the property of convergence and convergence rate of the original algorithm. In order to enable algorithms to realize their full potential, step sizes are carefully selected based on the recommended choice in [15], [29], [30], [31], [32].

Two representative generators are selected in Fig. 3. Since the optimal point of generator 2 is at the constraint boundary, the exact solution is found after some step $k$ for all algorithms, which is exhibited as a line with an ending. On the contrary, the constraints of generator 5 are not tight at the optimal point. Thus as shown in Fig. 3(b), the generator 5 constantly seeks for the optimal point. It is shown that algorithms in this paper, [31] and [32] all converge linearly to the optimal point, while the algorithms in [15], [29], [30] do not. It is verified that ET-DAPDA in [32], as an accelerate version of ET-DPDA in [31], is faster than ET-DPDA, while our algorithm is fastest in this case.

Slight wave exists in the proposed Algorithm 1 and [32] in Fig. 3 as a result of the event-triggered shceme. It has been shown in various work, such as [20], [44], that with inexact information, algorithms can only reach an neighbourhood of the exact solution. As a important feature of the event-triggered scheme, algorithms always use outdated information as approximation of neighbours' states. When the event-triggered condition is satisfied, the neighbours receive new information as stimulations and reach a higher accuracy. Thus a well designed event-triggered condition will be satisfied when necessary and help the algorithm constantly seek for the exact solution. It is shown in Section IV and verified in Fig. 3 that linear convergence rate maintains when the dynamic event-triggered scheme is integrated into the proposed algorithm. The fact in the proof of Theorem 1 is also verified in Fig. 3 that the error is bounded by a linearly decreased bound.

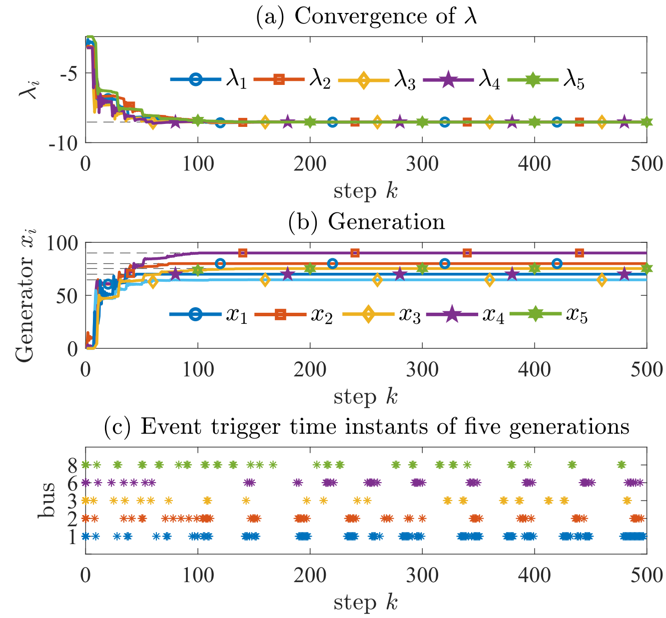


Fig. 2. Performance of Algorithm 1. (a) The evolution of generators' estimations on dual variable $\lambda$; (b) the evolution of generators' output $\{x_i\}$; (c) the triggered time of the proposed dynamic event-triggered algorithm.

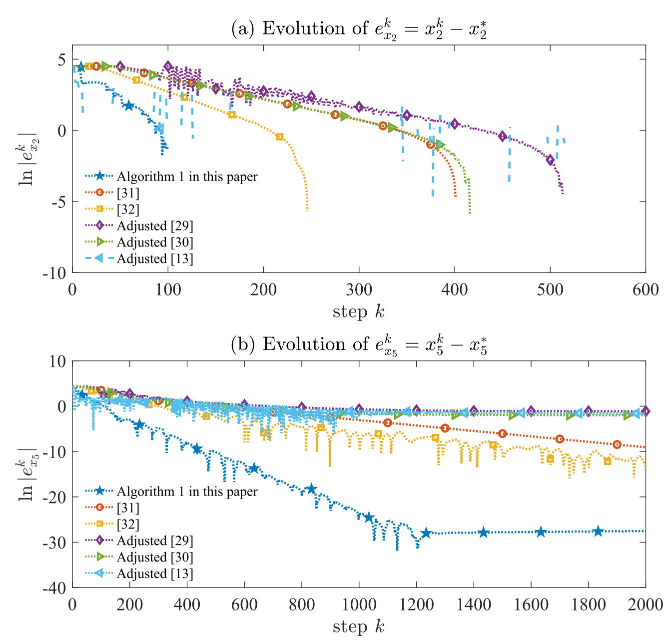


Fig. 3. The comparison of relative errors across different algorithms. (a) The evolution of error for the second generator (bus 2); (b) the evolution of error for the fifth generator (bus 8).

## B. Performance From the View of Data

Two indices are considered in the simulation to measure the effect on reducing communication burden. One commonly used index is the trigger frequency. When the error between the estimation $x_i^k$ and the optimal solution $x_i^*$ reaches the accuracy $e^{-2}$ for all generators, the average communication rate of the proposed algorithm is 77/745 = 10.34%, while that of [15] [30] [31] and [32] are 5720/17445 = 32.79%, 4792/8675 = 55.24%, 701/4245 = 16.51% and 470/2885 = 16.29%

TABLE II
NUMBER OF TRIGGER OF ALL NODES

| Accuracy | $e^{-1}$ | $e^{-2}$ | $e^{-6}$ | $e^{-10}$ | $e^{-14}$ |
|---|---|---|---|---|---|
| **[30]** | 1738 | 2744 | - [1] | - | - |
| **[29]** | 4144 | - | - | - | - |
| **[15]** | 6065 | 17445 | 5302735 | - | - |
| **[31]** | 661 | 701 | 869 | 1039 | 1215 |
| **[32]** | 444 | 523 | 639 | 742 | 847 |
| (3)[2] | 475 | 625 | 1190 | 1835 | 2385 |
| **Algorithm 1** | 65 | 77 | 173 | 298 | 397 |

[1] *The accuracy can not be achieved within ten million iterations.*
[2] *(3) is a time-triggered version of Algorithm 1.*

TABLE III
ACHIEVED ACCURACY OF ALL NODES ($\max_i\{\log_{10}|x_i^k - X_i^*|\}$)

| Number of Trigger | 50 | 100 | 500 | 1000 |
|---|---|---|---|---|
| **[30]**[1] | 1.9542 | 1.9542 | 1.5337 | 0.9078 |
| **[29]** | 1.6942 | 1.3400 | 0.1782 | 0.0038 |
| **[15]** | 1.3481 | 1.3481 | 0.6757 | −0.2160 |
| **[31]** | 1.9197 | 1.8324 | 0.7658 | −3.9567 |
| **[32]** | 1.8924 | 1.7923 | −1.6215 | −9.6353 |
| (3)[2] | 0.9989 | 0.7270 | −0.5711 | −2.0002 |
| **Algorithm 1** | 0.1636 | −1.2867 | −7.7111 | −12.1150 |

[1] *One node is called triggered if it sends its current information to at least one of its neighbours.*
[2] *(3) is a time-triggered version of Algorithm 1.*

respectively. The average communication rate of the algorithm in [29] is absent here since the algorithm fails to reach the accuracy $e^{-2}$ within ten million iterations.

As the event-triggered scheme extends the communication intervals, it leads to more iteration steps and may slow down the convergence rate. Thus to measure the effect on reducing the communication, the relation between the number of trigger and the achieved accuracy is also considered as an index and shown in Table II and Table III. By comparing the results of time-triggered algorithm (3) and event-triggered Algorithm 1, it is clear that the integration of dynamic event-triggered scheme in the proposed algorithm does reduce communication burden in this case. In Table II we can see that the proposed algorithm uses fewer steps than [15], [29], [30], [31], [32] to meet four different accuracy requirements. In Table III we can see that the proposed algorithm can achieve higher accuracy than [15], [29], [30], [31], [32] with four different trigger numbers. These all show the advantage of the proposed dynamic event-triggered algorithm. Furthermore, if we compare the algorithms by the amount of variable transmitted, the proposed algorithm enjoys even smaller one than [31] and ET-DAPDA [32].

## VI. Conclusion

This paper concentrates on dealing with economic dispatch problem distributedly and reducing communication burden without increasing computation complexity. Three aspects are considered and implemented, that includes the iteration number under certain accuracy requirement, the broadcast frequency per iteration, and the amount of information exchanged per broadcast. Based on the primal-dual method, this paper proposes a linearly convergent algorithm which only needs half amount of transmitted variables compared with existing algorithms. Furthermore, dynamic event-triggered scheme is integrated while the linear convergence rate is retained. The effectiveness and advantage of the proposed algorithm are shown by numerical simulations.

## Appendix A
## Proof of Theorem 1

First we need to prove the following Lemma.

*Lemma 3:* For the sequence $\{\boldsymbol{\lambda}^k, \mathbf{b}^k, \mathbf{x}^k, \mathbf{e}^k\}$, the following equation always holds:

$$\begin{aligned}&\beta(1-2\gamma\beta)L(\hat{\boldsymbol{\lambda}}^{k+1}-\boldsymbol{\lambda}^*)+(I-\beta(1-\gamma\beta)L)(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k})\\&=-\alpha\beta(\mathbf{g}^k-\mathbf{g}^*)-U(\mathbf{b}^{k+1}-\mathbf{b}^*)+\mathbf{e}^{k+1}.\end{aligned}$$

*Proof:* According to the definition, we have $U^\top U = \gamma\beta^2 L$ and $\mathbf{b}^{k+1} - \mathbf{b}^k = U\hat{\boldsymbol{\lambda}}^{k+1}$. Thus the iteration (8) can be written as

$$\begin{aligned}\boldsymbol{\lambda}^{k+1} =&\hat{\boldsymbol{\lambda}}^{k}-\beta L\hat{\boldsymbol{\lambda}}^{k}-\alpha\beta\mathbf{g}^k-U\mathbf{b}^{k-1}\\ =&\hat{\boldsymbol{\lambda}}^{k}-\beta L\hat{\boldsymbol{\lambda}}^{k}-\alpha\beta\mathbf{g}^k-U\mathbf{b}^{k+1}+\gamma\beta^2 L(\hat{\boldsymbol{\lambda}}^{k}+\hat{\boldsymbol{\lambda}}^{k+1}).\end{aligned}$$

Thus,

$$\begin{aligned}&\beta(1-2\gamma\beta)L\hat{\boldsymbol{\lambda}}^{k+1}+(I-\beta L+\gamma\beta^2 L)(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k})\\&=-\alpha\beta\mathbf{g}^k-U\mathbf{b}^{k+1}+\mathbf{e}^{k+1}.\end{aligned}$$

By noting that $U\mathbf{b}^* + \alpha\beta\mathbf{g}(\boldsymbol{\lambda}^*) = \mathbf{0}$, we come to the conclusion that

$$\begin{aligned}&\beta(1-2\gamma\beta)L(\hat{\boldsymbol{\lambda}}^{k+1}-\boldsymbol{\lambda}^*)+(I-\beta(1-\gamma\beta)L)(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k})\\&=-\alpha\beta(\mathbf{g}^k-\mathbf{g}^*)-U(\mathbf{b}^{k+1}-\mathbf{b}^*)+\mathbf{e}^{k+1}.\end{aligned}$$

■

Now we are ready to prove Theorem 1:

*Proof:* According to the definition of $\chi_i^k$ and the event-triggered condition, we have

$$\chi_i^{k+1} > (1-\tau-\frac{\delta_i}{\theta_i})\chi_i^k + \delta_i\eta_i p_i^{k+1}.$$

As $p_i^k$ and $\chi_i^0$ are nonnegative and positive respectively, $\chi_i^k > 0$ will always hold by setting $1-\tau-\frac{\delta_i}{\theta_i} \geq 0$.

Define

$$V^k = \|\mathbf{q}^k - \mathbf{q}^*\|_G^2,$$

where

$$\mathbf{q}^k = \begin{pmatrix}\mathbf{b}^k\\ \boldsymbol{\lambda}^k\end{pmatrix} \text{ and } G = \begin{pmatrix} I & 0\\ 0 & I-\beta(1-\gamma\beta)L\end{pmatrix}.$$

Thus $V^k = \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*\|^2_{I-\beta(1-\gamma\beta)L} + \|\mathbf{b}^k - \mathbf{b}^*\|^2$.

Next we focus on

$$V^k - V^{k+1} = \|\mathbf{q}^k - \mathbf{q}^*\|_G^2 - \|\mathbf{q}^{k+1} - \mathbf{q}^*\|_G^2.$$

As the equation

$$\begin{aligned}&2\langle \mathbf{q}^{k+1} - \mathbf{q}^k, G(\mathbf{q}^* - \mathbf{q}^{k+1})\rangle \\ =& - \|\mathbf{q}^{k+1} - \mathbf{q}^*\|_G^2 + \|\mathbf{q}^k - \mathbf{q}^*\|_G^2 - \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2\end{aligned}$$

always holds, it results into that

$$\begin{aligned}&V^k - V^{k+1} \\ =& 2\langle \mathbf{q}^{k+1} - \mathbf{q}^k, G(\mathbf{q}^* - \mathbf{q}^{k+1})\rangle + \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 \\ =& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 + 2\langle \mathbf{b}^{k+1} - \mathbf{b}^k, \mathbf{b}^* - \mathbf{b}^{k+1}\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle.\end{aligned}$$

Recalling the definition of $\mathbf{b}^k$ and $U\boldsymbol{\lambda}^* = \mathbf{0}$, it is easy to get $\mathbf{b}^{k+1} - \mathbf{b}^k = U\hat{\lambda}^{k+1}$ and obtain

$$\begin{aligned}&V^k - V^{k+1} \\ =& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 + 2\langle \hat{\lambda}^{k+1} - \boldsymbol{\lambda}^*, U(\mathbf{b}^* - \mathbf{b}^{k+1})\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle.\end{aligned}$$

One can further utilize the definition $\mathbf{e}^k = \hat{\lambda}^k - \boldsymbol{\lambda}^k$ to derive the following equations:

$$\begin{aligned}&V^k - V^{k+1} \\ =& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 + 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, U(\mathbf{b}^* - \mathbf{b}^{k+1})\rangle \\ &+ 2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^* - \mathbf{b}^{k+1})\rangle \\ &+ 2\langle \hat{\lambda}^{k+1} - \hat{\lambda}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle -\mathbf{e}^{k+1} + \mathbf{e}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ =& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 + 2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^* - \mathbf{b}^{k+1})\rangle \\ &+ 2\langle -\mathbf{e}^{k+1} + \mathbf{e}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, -U(\mathbf{b}^{k+1} - \mathbf{b}^*)\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, -(I - \beta(1-\gamma\beta)L)(\hat{\lambda}^{k+1} - \hat{\lambda}^k)\rangle.\end{aligned}$$

Then applying Lemma 3 gives

$$\begin{aligned}&V^k - V^{k+1} \\ =& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 + 2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^* - \mathbf{b}^{k+1})\rangle \\ &+ 2\langle -\mathbf{e}^{k+1} + \mathbf{e}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \beta(1-2\gamma\beta)L(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*)\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \beta(1-2\gamma\beta)L\mathbf{e}^{k+1}\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \alpha\beta(\mathbf{g}^k - \mathbf{g}^*) - \mathbf{e}^{k+1}\rangle \\ =& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 + 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \alpha\beta(\mathbf{g}^k - \mathbf{g}^*)\rangle \\ &+ 2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^* - \mathbf{b}^{k+1}) + \gamma\beta^2 L(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \mathbf{e}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \beta(1-2\gamma\beta)L(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*)\rangle,\end{aligned}$$

where the last equation reorganizes the item related to errors $\mathbf{e}^k$ and $\mathbf{e}^{k+1}$.

By spliting $\mathbf{g}^k - \mathbf{g}^*$ into $\mathbf{g}^k - \mathbf{g}^{k+1} + \mathbf{g}^{k+1} - \mathbf{g}^*$, we have

$$\begin{aligned}&V^k - V^{k+1} \\ =& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 \\ &+ 2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^* - \mathbf{b}^{k+1}) + \gamma\beta^2 L(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \mathbf{e}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \beta(1-2\gamma\beta)L(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*)\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \alpha\beta(\mathbf{g}^k - \mathbf{g}^{k+1})\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \alpha\beta(\mathbf{g}^{k+1} - \mathbf{g}^*)\rangle.\end{aligned}$$

Under Assumptions 4 and 5, applying Lemma 1 gives that

$$\begin{aligned}\|g_i(x) - g_i(y)\|_F &\leq \frac{1}{\mu}\|x - y\|_F, \\ \langle g_i(x_i) - g_j(x_j), x_i - x_j\rangle &\geq \frac{1}{l}\|x_i - x_j\|^2\end{aligned}$$

holds for any $x, y \in \mathbb{R}$. Thus

$$2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \alpha\beta(\mathbf{g}^{k+1} - \mathbf{g}^*)\rangle \geq \frac{2\alpha\beta}{l}\|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2.$$

As to $2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \alpha\beta(\mathbf{g}^k - \mathbf{g}^{k+1})\rangle$, we have

$$\begin{aligned}&2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \alpha\beta(\mathbf{g}^k - \mathbf{g}^{k+1})\rangle \\ \geq& -\alpha\beta\zeta_1\|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2 - \frac{\alpha\beta}{\zeta_1}\|\mathbf{g}^k - \mathbf{g}^{k+1}\|^2 \\ \geq& -\alpha\beta\zeta_1\|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2 - \frac{\alpha\beta}{\zeta_1\mu^2}\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|^2.\end{aligned}$$

It follows that

$$\begin{aligned}&V^k - V^{k+1} \\ \geq& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 \\ &+ 2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^* - \mathbf{b}^{k+1}) + \gamma\beta^2 L(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \mathbf{e}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \beta(1-2\gamma\beta)L(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*)\rangle \\ &- \alpha\beta\zeta_1\|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2 - \frac{\alpha\beta}{\zeta_1\mu^2}\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|^2 \\ &+ \frac{2\alpha\beta}{l}\|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2\end{aligned}$$

and that

$$\begin{aligned}&V^k - V^{k+1} \\ \geq& \|\mathbf{q}^{k+1} - \mathbf{q}^k\|_G^2 \\ &+ 2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^* - \mathbf{b}^{k+1}) + \gamma\beta^2 L(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ 2\langle \mathbf{e}^k, (I - \beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1})\rangle \\ &+ \|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2_{\alpha\beta(\frac{2}{l} - \zeta_1)I + 2\beta(1-2\gamma\beta)L} \\ &- \frac{\alpha\beta}{\zeta_1\mu^2}\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|^2.\end{aligned}$$

Recalling the definition of $\mathbf{q}^k$ and $G$, it is easy to get that

$$\begin{aligned}&V^k - V^{k+1}\\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{(1-\frac{\alpha\beta}{\zeta_1\mu^2})I-\beta(1-\gamma\beta)L}\\ &+2\langle \mathbf{e}^{k+1}, U(\mathbf{b}^*-\mathbf{b}^{k+1})+\gamma\beta^2 L(\boldsymbol{\lambda}^*-\boldsymbol{\lambda}^{k+1})\rangle\\ &+2\langle \mathbf{e}^k, (I-\beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^*-\boldsymbol{\lambda}^{k+1})\rangle\\ &+\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{\alpha\beta(\frac{2}{l}-\zeta_1)I+2\beta(1-2\gamma\beta)L}.\end{aligned}$$

Reorganizing the items and applying $2\langle a,b\rangle \geq -\zeta\|a\|^2 - \frac{1}{\zeta}\|b\|^2$ gives

$$\begin{aligned}&V^k - V^{k+1}\\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{(1-\frac{\alpha\beta}{\zeta_1\mu^2})I-\beta(1-\gamma\beta)L}\\ &+\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{(\alpha\beta(\frac{2}{l}-\zeta_1)-\zeta_3)I+2\beta(1-2\gamma\beta)L}\\ &-\zeta_2\|\mathbf{e}^{k+1}\|^2-\frac{1}{\zeta_2}\|U(\mathbf{b}^*-\mathbf{b}^{k+1})\|^2\\ &-\frac{1}{\zeta_3}\|\gamma\beta^2\ L\mathbf{e}^{k+1}+(I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2.\end{aligned}$$

As we want to get a relation between $V^k - V^{k+1}$ and $V^{k+1}$, next we consider the term $V^k-(1+\epsilon_1)V^{k+1}$ where $\epsilon_1 > 0$. It follows directly from the previous inequality that

$$\begin{aligned}&V^k-(1+\epsilon_1)V^{k+1}\\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{(1-\frac{\alpha\beta}{\zeta_1\mu^2})I-\beta(1-\gamma\beta)L}\\ &+\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{(\alpha\beta(\frac{2}{l}-\zeta_1)-\zeta_3)I+2\beta(1-2\gamma\beta)L-\epsilon_1(I-\beta(1-\gamma\beta)L)}\\ &-\epsilon_1\|\mathbf{b}^{k+1}-\mathbf{b}^*\|^2-\frac{1}{\zeta_2}\|U(\mathbf{b}^*-\mathbf{b}^{k+1})\|^2-\zeta_2\|\mathbf{e}^{k+1}\|^2\\ &-\frac{1}{\zeta_3}\|\gamma\beta^2\ L\mathbf{e}^{k+1}+(I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2.\end{aligned}$$

To give a upper bound on $\|U(\mathbf{b}^{k+1}-\mathbf{b}^*)\|^2$ and $\|\mathbf{b}^{k+1}-\mathbf{b}^*\|^2$, we apply Lemma 3 again,

$$\begin{aligned}&\|U(\mathbf{b}^{k+1}-\mathbf{b}^*)\|^2\\ =&\|\beta(1-2\gamma\beta)L(\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*)+\alpha\beta(\mathbf{g}^{k+1}-\mathbf{g}^*)\\ &+(I-\beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^k)-\gamma\beta^2\ L\mathbf{e}^{k+1}\\ &+\alpha\beta(\mathbf{g}^k-\mathbf{g}^{k-1})-(I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2.\end{aligned}$$

Apply the following inequality repeatedly

$$\|a+b\|^2 \leq \zeta\|a\|^2+\frac{\zeta}{\zeta-1}\|b\|^2, \text{ for any } \zeta > 1,$$

then it yields that

$$\begin{aligned}&\|U(\mathbf{b}^{k+1}-\mathbf{b}^*)\|^2\\ \leq&\frac{\zeta_4}{\zeta_4-1}\|\gamma\beta^2\ L\mathbf{e}^{k+1}+(I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2\\ &+\zeta_4\zeta_5\zeta_6\beta^2(1-2\gamma\beta)^2\|L(\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*)\|^2\\ &+\frac{\zeta_4\zeta_5\zeta_6\alpha^2\beta^2}{\zeta_6-1}\|\mathbf{g}^{k+1}-\mathbf{g}^*\|^2\\ &+\frac{\zeta_4\zeta_5\zeta_7\alpha^2\beta^2}{(\zeta_5-1)(\zeta_7-1)}\|\mathbf{g}^k-\mathbf{g}^{k+1}\|^2\\ &+\frac{\zeta_4\zeta_5\zeta_7}{\zeta_5-1}\|(I-\beta(1-\gamma\beta)L)(\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^k)\|^2),\end{aligned}$$

where $\zeta_4 > 1$, $\zeta_5 > 1$, $\zeta_6 > 1$, $\zeta_7 > 1$. Then apply Lemma 1 under Assumptions 4 and 5 and choose $\zeta_6 = 1 + \frac{\alpha}{\mu(1-2\gamma\beta)\sqrt{\sigma_{\max}(L)}}$ and $\zeta_7 = 1+\frac{\alpha\beta}{\mu\sqrt{\sigma_{\max}(I-\beta(1-\gamma\beta)L)}}$, we can obtain that

$$\begin{aligned}&\|U(\mathbf{b}^{k+1}-\mathbf{b}^*)\|^2\\ \leq&\frac{\zeta_4}{\zeta_4-1}\|\gamma\beta^2\ L\mathbf{e}^{k+1}+(I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2\\ &+\zeta_4\zeta_5\beta^2\left((1-2\gamma\beta)\sqrt{\sigma_{\max}(L)}+\frac{\alpha}{\mu}\right)^2\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2\\ &+\frac{\zeta_4\zeta_5}{\zeta_5-1}\left(\sqrt{\sigma_{\max}(I-\beta(1-\gamma\beta)L)}+\frac{\alpha\beta}{\mu}\right)^2\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^k\|^2.\end{aligned}$$

Now $\|U(\mathbf{b}^{k+1}-\mathbf{b}^*)\|^2$ is proved to be bounded. To derive an upper bound on $\|\mathbf{b}^{k+1}-\mathbf{b}^*\|^2$, recall the definition $\mathbf{b}^k = U\sum_{t=0}^k \hat{\lambda}^t$ and $U^\top U = \gamma\beta^2\ L$, then we have

$$\|\mathbf{b}^{k+1}-\mathbf{b}^*\|^2 \leq \frac{1}{\gamma\beta^2\lambda_{\min}(L)}\|U(\mathbf{b}^{k+1}-\mathbf{b}^*)\|^2.$$

With the given upper bounds on $\|U(\mathbf{b}^{k+1}-\mathbf{b}^*)\|^2$ and $\|\mathbf{b}^{k+1}-\mathbf{b}^*\|^2$ we have

$$\begin{aligned}&V^k-(1+\epsilon_1)V^{k+1}\\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{(1-\frac{\alpha\beta}{\zeta_1\mu^2})I-\beta(1-\gamma\beta)L}\\ &+\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{(\alpha\beta(\frac{2}{l}-\zeta_1)-\zeta_3)I+2\beta(1-2\gamma\beta)L-\epsilon_1(I-\beta(1-\gamma\beta)L)}\\ &-\frac{1}{\zeta_3}\|\gamma\beta^2\ L\mathbf{e}^{k+1}+(I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2-\zeta_2\|\mathbf{e}^{k+1}\|^2\\ &-\left(\frac{\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)}+\frac{1}{\zeta_2}\right)\|U(\mathbf{b}^*-\mathbf{b}^{k+1})\|^2\\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2+\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{A_1}+\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{A_2}\\ &-A_3\|\gamma\beta^2\ L\mathbf{e}^{k+1}+(I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2-\zeta_2\|\mathbf{e}^{k+1}\|^2\end{aligned}$$

where

$$\begin{aligned}A_1 =& (1-\frac{\alpha\beta}{\zeta_1\mu^2})I-\beta(1-\gamma\beta)L\\ &-\left(\frac{\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)}+\frac{1}{\zeta_2}\right)\frac{\zeta_4\zeta_5}{\zeta_5-1}\\ &*\left(\sqrt{\sigma_{\max}(I-\beta(1-\gamma\beta)L)}+\frac{\alpha\beta}{\mu}\right)^2,\end{aligned} \tag{9}$$

$$\begin{aligned}A_2 =& (\alpha\beta(\tfrac{2}{l}-\zeta_1)-\zeta_3)I+2\beta(1-2\gamma\beta)L\\ &-\epsilon_1(I-\beta(1-\gamma\beta)L)-\left(\tfrac{\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)}+\tfrac{1}{\zeta_2}\right)\\ &*\zeta_4\zeta_5\beta^2\left((1-2\gamma\beta)\sqrt{\sigma_{\max}(L)}+\tfrac{\alpha}{\mu}\right)^2,\end{aligned} \tag{10}$$

$$A_3 = \left(\tfrac{\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)}+\tfrac{1}{\zeta_2}\right)\tfrac{\zeta_4}{\zeta_4-1}+\tfrac{1}{\zeta_3}. \tag{11}$$

Applying the inequality

$$\|a+b\|^2 \leq \zeta\|a\|^2 + \frac{\zeta}{\zeta-1}\|b\|^2, \text{ for any } \zeta > 1,$$

again for $\|\gamma\beta^2\ L\mathbf{e}^{k+1} + (I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2$ gives

$$\begin{aligned}&\|\gamma\beta^2\ L\mathbf{e}^{k+1} + (I-\beta(1-\gamma\beta)L)\mathbf{e}^k\|^2 \\ \leq&\zeta_8\gamma^2\beta^4\sigma_{\max}(L)\|\mathbf{e}^{k+1}\|^2 + \frac{\zeta_8\sigma_{\max}(I-\beta(1-\gamma\beta)L)}{\zeta_8-1}\|\mathbf{e}^k\|^2.\end{aligned}$$

Hence,

$$\begin{aligned}&V^k - (1+\epsilon_1)V^{k+1} \\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|^2_{A_1} + \|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{A_2} \\ &- A_4\|\mathbf{e}^{k+1}\|^2 - A_5\|\mathbf{e}^k\|^2,\end{aligned}$$

where

$$A_4 = \zeta_2 + \zeta_8\gamma^2\beta^4\sigma_{\max}(L)A_3, \tag{12}$$

$$A_5 = \tfrac{\zeta_8\sigma_{\max}(I-\beta(1-\gamma\beta)L)}{\zeta_8-1}A_3. \tag{13}$$

Define $\chi^k = \sum_i \chi_i^k$ and $M^K = V^k + \kappa\chi^k$, then

$$\begin{aligned}&M^k - (1+\min\{\epsilon_1,\epsilon_2\})M^{k+1} \\ \geq&V^k - (1+\epsilon_1)V^{k+1} + \kappa\chi^k - (1+\epsilon_2)\kappa\chi^{k+1} \\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{A_1} + \|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{A_2} \\ &- A_4\|\mathbf{e}^{k+1}\|^2 - A_5\|\mathbf{e}^k\|^2 + \kappa\chi^k - (1+\epsilon_2)\kappa\chi^{k+1}\end{aligned}$$

Recall that the iteration of the internal dynamic variable $\chi_i^k$ is designed as

$$\chi_i^{k+1} = (1-\tau_i)\chi_i^k - \delta_i(\|e_i^k\|^2 - \eta_i p_i^{k+1}), \chi_i^0 > 0.$$

and the event-triggered time for node $i$ is designed as:

$$k_{i,t+1} = \min_k\{\theta_i\|e_i^k\|^2 \geq \chi_i^k, k > k_{i,t}\}.$$

It results that

$$\begin{aligned}&M^k - (1+\min\{\epsilon_1,\epsilon_2\})M^{k+1} \\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{A_1} + \|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{A_2} \\ &- A_4\sum_i\frac{1}{\theta_i}\chi_i^{k+1} - A_5\sum_i(\|e_i^k\|^2 - \eta_i p_i^{k+1}) \\ &- A_5\sum_i\eta_i p_i^{k+1} + \kappa\chi^k - (1+\epsilon_2)\kappa\chi^{k+1} \\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{A_1} + \|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{A_2} \\ &- A_4\sum_i\frac{1}{\theta_i}\chi_i^{k+1} + A_5\sum_i\frac{1}{\delta_i}(\chi_i^{k+1} - (1-\tau_i)\chi_i^k) \\ &- A_5\sum_i\eta_i p_i^{k+1} + \kappa\chi^k - (1+\epsilon_2)\kappa\chi^{k+1}.\end{aligned}$$

Reorganizing the items gives

$$\begin{aligned}&M^k - (1+\min\{\epsilon_1,\epsilon_2\})M^{k+1} \\ \geq&\|\mathbf{b}^{k+1}-\mathbf{b}^k\|^2 + \|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{A_1} + \|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{A_2} \\ &+ \sum_i\left(-\frac{A_4}{\theta_i} + \frac{A_5}{\delta_i} - (1+\epsilon_2)\kappa\right)\chi_i^{k+1} \\ &+ \sum_i\left(\kappa - \frac{A_5(1-\tau_i)}{\delta_i}\right)\chi_i^k - A_5\sum_i\eta_i p_i^{k+1}.\end{aligned}$$

It follows from the definition of $p_i^k$ that

$$\begin{aligned}\sum_{i=1}^N p_i^k =& \sum_{i=1}^N -\frac{1}{2}\sum_{j=1}^N L_{ij}\|\hat{\lambda}_j^k - \hat{\lambda}_i^k\|^2 = \sum_{i=1}^N\sum_{j=1}^N \hat{\lambda}_i^k L_{ij}\hat{\lambda}_j^k \\ =&\hat{\lambda}^{k\top}L\hat{\lambda}^k = \|U\hat{\lambda}^k\|^2 = \|\mathbf{b}^k - \mathbf{b}^{k-1}\|^2.\end{aligned}$$

Thus we have

$$\begin{aligned}&M^k - (1+\min\{\epsilon_1,\epsilon_2\})M^{k+1} \\ \geq&\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2_{A_1} + \|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2_{A_2} \\ &+ \sum_i\left(-\frac{A_4}{\theta_i} + \frac{A_5}{\delta_i} - (1+\epsilon_2)\kappa\right)\chi_i^{k+1} \\ &+ \sum_i\left(\kappa - \frac{A_5(1-\tau_i)}{\delta_i}\right)\chi_i^k + \sum_i(1-A_5\eta_i)p_i^{k+1}.\end{aligned}$$

If we can show that there exists $\epsilon_1 > 0$ and $\epsilon_2 > 0$ so that the following inequality holds:

$$M^k - (1+\min\{\epsilon_1,\epsilon_2\})M^{k+1} \geq 0,$$

then we have that $M^k$ converges to 0 at the linear rate $O((1+\min\{\epsilon_1,\epsilon_2\})^{-k})$, that is

$$M^{k+1} \leq \frac{1}{1+\min\{\epsilon_1,\epsilon_2\}}M^k \leq (1+\min\{\epsilon_1,\epsilon_2\})^{-(k+1)}M^0.$$

As $\chi > 0$ and $\kappa > 0$, we come to the conclusion that

$$\begin{aligned}\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^*\|^2 \leq&\sigma_{\min}(I-\beta(1-\gamma\beta)L)^{-1}V^k \\ \leq&\sigma_{\min}(I-\beta(1-\gamma\beta)L)^{-1}M^k \\ \leq&(1+\min\{\epsilon_1,\epsilon_2\})^{-k}\sigma_{\min}^{-1}(I-\beta(1-\gamma\beta)L)M^0.\end{aligned}$$

Recall that by applying Lemma 1, the function $H_i(\lambda)$ of the dual problem (2) has $\frac{1}{\mu}$-Lipschitz continuous gradient under the Assumptions 4 and 5. Define $\nabla\mathbf{H} = [\nabla H_1, \nabla H_2, \ldots, \nabla H_N]$. Recall that the gradient of $H$ is $\nabla H_i(\lambda) = d_i - x_i$. Thus,

$$\begin{aligned}&\|\mathbf{x}^k - \mathbf{x}^*\|^2 \\ =&\|\nabla\mathbf{H}(\lambda^k) - \nabla\mathbf{H}(\lambda^*)\|^2 \leq \frac{1}{\mu^2}\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^*\| \\ \leq&\frac{1}{\mu^2}(1+\min\{\epsilon_1,\epsilon_2\})^{-k}\sigma_{\min}^{-1}(I-\beta(1-\gamma\beta)L)M^0,\end{aligned}$$

which means that the proposed algorithm is guaranteed to find the global optimal point at a linear rate $O((1+\min\{\epsilon_1,\epsilon_2\})^{-k})$ for all agents.

Next we concentrate on the existence of $\epsilon_1 > 0$ and $\epsilon_2 > 0$ so that

$$M^k - (1+\min\{\epsilon_1,\epsilon_2\})M^{k+1} \geq 0.$$

It needs $A_1$ and $A_2$ to be positive and

$$-\frac{A_4}{\theta_i} + \frac{A_5}{\delta_i} - (1+\epsilon_2)\kappa \geq 0; \tag{14}$$

$$\kappa - \frac{A_5(1-\tau_i)}{\delta_i} \geq 0; \tag{15}$$

$$1 - A_5\eta_i \geq 0. \tag{16}$$

Recall that $L = I - W$ is positive semidefinite and that $1 - \tau - \frac{\delta_i}{\theta_i}$ has to be nonnegative. The aforementioned conditions are discussed in detail as below:

1) $A_1 \succ 0$

$$\begin{aligned} A_1 =& (1 - \frac{\alpha\beta}{\zeta_1\mu^2})I - \beta(1-\gamma\beta)L \\ &- \frac{\zeta_4\zeta_5}{\zeta_2(\zeta_5-1)}\left(\sqrt{\sigma_{\max}(I-\beta(1-\gamma\beta)L)} + \frac{\alpha\beta}{\mu}\right)^2 I \\ &- \frac{\zeta_4\zeta_5\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)(\zeta_5-1)} \\ &* \left(\sqrt{\sigma_{\max}(I-\beta(1-\gamma\beta)L)} + \frac{\alpha\beta}{\mu}\right)^2 I. \end{aligned}$$

To set $A_1$ as a positive matrix we can choose $\alpha$ and $\beta$ so that

$$1 - \frac{\alpha\beta}{\zeta_1\mu^2} - \beta(1-\gamma\beta)\sigma_{\max}(L) > 0,$$

and set sufficiently large $\zeta_2 > 0$ and sufficiently small $\epsilon_1 > 0$.

2) $A_2 \succ 0$

$$\begin{aligned} &A_2 \\ =& \alpha\beta(\frac{2}{l} - \zeta_1)I + 2\beta(1-2\gamma\beta)L - \epsilon_1(I - \beta(1-\gamma\beta)L) - \zeta_3 I \\ &- \left(\frac{\epsilon_1}{\gamma\beta^2\lambda_{\min}(L)} + \frac{1}{\zeta_2}\right)\zeta_4\zeta_5\beta^2\left((1-2\gamma\beta)\sqrt{\sigma_{\max}(L)} + \frac{\alpha}{\mu}\right)^2. \end{aligned}$$

To set $A_2$ as a positive matrix we can set $\zeta_1 \in (0, \frac{2}{l})$ and choose $\alpha$ and $\beta$ so that

$$\alpha\beta(\frac{2}{l} - \zeta_1)I + 2\beta(1-2\gamma\beta)L \succ 0,$$

$$\beta(1-2\gamma\beta) > 0,$$

$$1 - \beta(1-\gamma\beta)\sigma_{\max}(L) \geq 0,$$

and set sufficiently large $\zeta_2 > 0$ and sufficiently small $\zeta_3 > 0$, $\epsilon_1 > 0$.

3) To make (14) (15) and (16) hold, we can choose arbitrary $\delta_i > 0$, $\tau_i \in (0,1)$, $\kappa$ satisfying

$$0 < \frac{A_5(1-\tau_i)}{\delta_i} \leq \kappa < \frac{A_5}{\delta_i},$$

sufficiently large $\theta_i$ satisfying

$$\theta_i > \frac{A_4}{\frac{A_5}{\delta_i} - \kappa},$$

sufficiently small $\epsilon_2$ satisfying

$$\epsilon_2 \leq \frac{1}{\kappa}\left(-\frac{A_4}{\theta_i} + \frac{A_5}{\delta_i} - \kappa\right),$$

and sufficiently small $\eta_i$ satisfying

$$\eta_i \leq \frac{1}{A_5}.$$

In conclusion, when we set $\zeta_1 \in (0, \frac{2}{l})$ and choose arbitrary $\gamma > 0$, $\beta$ satisfying

$$\beta \in \left(0, \frac{\sigma_{\max}(L) - \sqrt{\sigma_{\max}^2(L) - 4\sigma_{\max}(L)\gamma}}{2\sigma_{\max}(L)\gamma}\right),$$

$\alpha$ satisfying

$$\alpha \in \left(0, \frac{\zeta_1\mu^2(1 - 2\beta(1-\gamma\beta)\sigma_{\max}(L))}{\beta}\right),$$

and other parameters chosen as stated before, we can get some $\epsilon_1 > 0$ and $\epsilon_2 > 0$ so that

$$M^k - (1+\min\{\epsilon_1,\epsilon_2\})M^{k+1} \geq 0.$$

Then the conclusion of linear convergence follows.

Finally we show the relationship between the convergence rate and the chosen stepsizes. As it needs $A_1 \succ 0$, $A_2 \succ 0$ and (14) holds, we have

$$0 < \epsilon_1 < \min\{\epsilon_{11}, \epsilon_{12}\}$$

and

$$0 < \epsilon_2 \leq \frac{1}{\kappa}\left(-\frac{A_4}{\theta_i} + \frac{A_5}{\delta_i} - \kappa\right).$$

where

$$\epsilon_{11} = \frac{\gamma\beta^2\lambda_{\min}(L)(\zeta_5-1)\left((1-\frac{\alpha\beta}{\zeta_1\mu^2}) - \beta(1-\gamma\beta)\sigma_{\max}(L)\right)}{\zeta_4\zeta_5\epsilon_1\left(\sqrt{\sigma_{\beta,\gamma}} + \frac{\alpha\beta}{\mu}\right)^2},$$

$$\epsilon_{12} = \frac{\alpha\beta(\frac{2}{l} - \zeta_1) + 2\beta(1-2\gamma\beta)\sigma_{\max}(L)}{\sigma_{\beta,\gamma} + (\frac{\zeta_4\zeta_5\beta^2}{\gamma\beta^2\lambda_{\min}(L)})\left((1-2\gamma\beta)\sqrt{\sigma_{\max}(L)} + \frac{\alpha}{\mu}\right)^2},$$

and $\sigma_{\beta,\gamma} = \sigma_{\max}(I - \beta(1-\gamma\beta)L)$.

The convergence rate $(\frac{1}{1+\min\{\epsilon_1,\epsilon_2\}})^k$ have a complex relation with the chosen parameters including the stepsizes $\alpha$, $\beta$, $\gamma$ and the event-triggered parameters $\delta_i$, $\tau_i$, $\kappa$, $\theta_i$ and $\eta_i$. ■

## APPENDIX B
## PROOF OF THEOREM 3

Suppose that after time $k = T$, the event-triggered conditions always hold for all nodes, which means that after time

$k = T$ the event-triggered scheme does not work any more. As the iteration without event-triggered scheme in (3) can be seen as a particular situation of (1), we can similarly prove that for all $k \geq T$

$$V^k - (1+\epsilon_1)V^{k+1} \geq 0,$$

when $\alpha$ and $\beta$ are properly chosen as stated in Theorem 1. Furthermore, the following inequalities always hold for $k \geq T$:

$$\begin{aligned}\|\lambda_i^{k+1} - \lambda_i^k\|^2 \leq& 2(\|\lambda_i^{k+1} - \lambda_i^*\|^2 + \|\lambda_i^k - \lambda_i^*\|^2) \\ \leq& 2\ V^{k+1} + 2\ V^k \leq \frac{2(2+\epsilon_1)}{1+\epsilon_1} V^k \\ \leq& \frac{2(2+\epsilon_1)}{(1+\epsilon_1)^{k-T+1}} V^T.\end{aligned}$$

Recall that the internal dynamic variable $\chi_i$ iterates as

$$\chi_i^{k+1} = (1-\tau_i)\chi_i^k + \delta_i \eta_i p_i^{k+1},$$

for $k \geq T$. Thus if we set $\tau_i$ satisfying

$$1 - \tau_i > \frac{1}{1+\epsilon_1},$$

there must exist a time $k = T_1 > T$ so that

$$\|\lambda_i^{k+1} - \lambda_i^k\|^2 \leq \frac{2(2+\epsilon_1)}{(1+\epsilon_1)^{T_1-T+1}} V^T \leq (1-\tau_i)^{T_1-T} \chi_i^T < \chi_i^{T_1},$$

which means that there is a contradiction as the event-triggered condition does not hold for node $i$ at time $k = T_1 > T$.

## REFERENCES


[1] U. Nakarmi, M. Rahnamay-Naeini, and H. Khamfroush, "Critical component analysis in cascading failures for power grids using community structures in interaction graphs," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 1079–1093, Jul.–Sep. 2020.

[2] X. Lou, D. K. Yau, R. Tan, and P. Cheng, "Cost and pricing of differential privacy in demand reporting for smart grids," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 2037–2051, Jul.–Sep. 2020.

[3] K. Meng, W. Zhang, J. Qiu, Y. Zheng, and Z. Y. Dong, "Offshore transmission network planning for wind integration considering AC and DC transmission options," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4258–4268, Nov. 2019.

[4] R. Lasseter et al., "Integration of distributed energy resources. the CERTS microgrid concept," Lawrence Berkeley Nat. Lab., Berkeley, CA, USA, Tech. Rep., 2002.

[5] B. Y. Qu, Y. S. Zhu, Y. C. Jiao, M. Y. Wu, P. N. Suganthan, and J. J. Liang, "A survey on multi-objective evolutionary algorithms for the solution of the environmental/economic dispatch problems," *Swarm Evol. Comput.*, vol. 38, pp. 1–11, 2018.

[6] W. Lin and E. Bitar, "A structural characterization of market power in electric power networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 987–1006, Jul.–Sep. 2020.

[7] T. Yang et al., "A survey of distributed optimization," *Annu. Rev. Control*, vol. 47, pp. 278–305, 2019.

[8] Y. C. Ho, L. Servi, and R. Suri, "A class of center-free resource allocation algorithms," *IFAC Proc. Volumes*, vol. 13, no. 6, pp. 475–482., 1980.

[9] L. Xiao and S. Boyd, "Optimal scaling of a gradient method for distributed resource allocation," *J. Optim. Theory Appl.*, vol. 129, no. 3, pp. 469–488, 2006.

[10] C. Li, X. Yu, W. Yu, T. Huang, and Z. W. Liu, "Distributed event-triggered scheme for economic dispatch in smart grids," *IEEE Trans. Ind. Inform.*, vol. 12, no. 5, pp. 1775–1785, Oct. 2016.

[11] H. Li, Q. Lu, and T. Huang, "Convergence analysis of a distributed optimization algorithm with a general unbalanced directed communication network," *IEEE Trans. Netw. Sci. Eng.*, vol. 6, no. 3, pp. 237–248, Jul.–Sep. 2019.

[12] Q. Lu, X. Liao, H. Li, and T. Huang, "Achieving acceleration for distributed economic dispatch in smart grids over directed networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 1988–1999, Jul.–Sep. 2020.

[13] H. Zhang, H. Li, Y. Zhu, Z. Wang, and D. Xia, "A distributed stochastic gradient algorithm for economic dispatch over directed network with communication delays," *Int. J. Elect. Power Energy Syst.*, vol. 110, pp. 759–771, 2019.

[14] T. Yang et al., "A distributed algorithm for economic dispatch over time-varying directed networks with delays," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5095–5106, Jun. 2017.

[15] H. Li, Z. Wang, G. Chen, and Z. Y. Dong, "Distributed robust algorithm for economic dispatch in smart grids over general unbalanced directed networks," *IEEE Trans. Ind. Inform.*, vol. 16, no. 7, pp. 4322–4332, Jul. 2020.

[16] V. C. Gungor et al., "Smart grid technologies: Communication technologies and standards," *IEEE Trans. Ind. Inform.*, vol. 7, no. 4, pp. 529–539, Nov. 2011.

[17] K. Yuan, Q. Ling, and W. Yin, "On the convergence of decentralized gradient descent," *SIAM J. Optim.*, vol. 26, no. 3, pp. 1835–1854, 2016.

[18] A. Nedić, A. Olshevsky, and W. Shi, "Achieving geometric convergence for distributed optimization over time-varying graphs," *SIAM J. Optim.*, vol. 27, no. 4, pp. 2597–2633, 2017.

[19] W. Shi, Q. Ling, G. Wu, and W. Yin, "Extra: An exact first-order algorithm for decentralized consensus optimization," *SIAM J. Optim.*, vol. 25, no. 2, pp. 944–966, 2015.

[20] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Proc. IEEE Conf. Decis. Control*, 2008, pp. 4177–4184.

[21] Z. Tang, D. J. Hill, and T. Liu, "A novel consensus-based economic dispatch for microgrids," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3920–3922, Jul. 2018.

[22] A. Nedić, A. Olshevsky, and W. Shi, "Improved convergence rates for distributed resource allocation," in *Proc. IEEE Conf. Decis. Control*, 2018, pp. 172–177.

[23] N. S. Aybat and E. Y. Hamedani, "A distributed ADMM-like method for resource sharing over time-varying networks," *SIAM J. Optim.*, vol. 29, no. 4, pp. 3036–3068, 2019.

[24] T. Yang et al., "Distributed energy resource coordination over time-varying directed communication networks," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 3, pp. 1124–1134, Sep. 2019.

[25] J. Zhang, K. You, and K. Cai, "Distributed dual gradient tracking for resource allocation in unbalanced networks," *IEEE Trans. Signal Process.*, vol. 68, pp. 2186–2198, 2020.

[26] P. Liu, H. Li, X. Dai, and Q. Han, "Distributed primal-dual optimisation method with uncoordinated time-varying step-sizes," *Int. J. Syst. Sci.*, vol. 49, no. 6, pp. 1256–1272, Mar. 2018.

[27] Y. Li, K. Meng, and Z. Y. Dong, "Distributed consensus control with event-triggered communication for multi-microgrid cluster," in *Proc. IEEE Milan PowerTech*, 2019, pp. 1–6.

[28] L. Ding, Q. L. Han, X. Ge, and X. M. Zhang, "An overview of recent advances in event-triggered consensus of multiagent systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 1110–1123, Apr. 2018.

[29] C. Liu, H. Li, Y. Shi, and D. Xu, "Distributed event-triggered gradient method for constrained convex minimization," *IEEE Trans. Autom. Control*, vol. 65, no. 2, pp. 778–785, Feb. 2020.

[30] Y. Kajiyama, N. Hayashi, and S. Takai, "Distributed subgradient method with edge-based event-triggered communication," *IEEE Trans. Autom. Control*, vol. 63, no. 7, pp. 2248–2255, Jul. 2018.

[31] J. Wang, H. Li, and Z. Wang, "Distributed event-triggered scheme for economic dispatch in power systems with uncoordinated step-sizes," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 16, pp. 3612–3622, 2019.

[32] K. Zhang, J. Xiong, X. Dai, and Q. Lü, "On the convergence of event-triggered distributed algorithm for economic dispatch problem," *Int. J. Elect. Power Energy Syst.*, vol. 122, 2020, Art. no. 106159.

[33] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1680–1685, Sep. 2007.

[34] C. Yang et al., "Dynamic event-triggered robust secondary frequency control for islanded AC microgrid," *Appl. Energy*, vol. 242, pp. 821–836, 2019.

[35] X. Shi, R. Zheng, T. Yang, Z. Lin, and G. Yan, “A distributed algorithm with event-triggered communication for resource allocation problem,” in *Proc. IEEE Amer. Control Conf. (ACC)*, 2019, pp. 3916–3921.

[36] A. Girard, “Dynamic triggering mechanisms for event-triggered control,” *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1992–1997, Jul. 2015.

[37] X. Yi, K. Liu, D. V. Dimarogonas, and K. H. Johansson, “Dynamic event-triggered and self-triggered control for multi-agent systems,” *IEEE Trans. Autom. Control*, vol. 64, no. 8, pp. 3300–3307, Aug. 2019.

[38] X. Ge, Q. L. Han, L. Ding, Y. L. Wang, and X. M. Zhang, “Dynamic event-triggered distributed coordination control and its applications: A survey of trends and techniques,” *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 9, pp. 3112–3125, Sep. 2020.

[39] X. Ge, Q. Han, and Z. Wang, “A dynamic event-triggered transmission scheme for distributed set-membership estimation over wireless sensor networks,” *IEEE Trans. Cybern.*, vol. 49, no. 1, pp. 171–183, Jan. 2019.

[40] J. Yin, Y. Chen, V. Gupta, and D. Wang, “Distributed convex optimization of discrete-time multi-agent systems: A new model,” in *Proc. IEEE 14th Conf. Ind. Electron. Appl. (ICIEA)*, 2019, pp. 2417–2422..

[41] J. Wang and N. Elia, “Control approach to distributed optimization,” in *Proc. IEEE 48th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, 2010, pp. 557–561.

[42] “IEEE 14 bus system,” 2021. [Online]. Available: http://icseg.iti.illinois.edu/ieee-14-bus-system/

[43] T. T. Doan and A. Olshevsky, “On the geometric convergence rate of distributed economic dispatch/demand response in power networks,” 2016, *arXiv:1609.06660*.

[44] M. G. Rabbat and R. D. Nowak, “Quantized incremental algorithms for distributed optimization,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 798–808, Apr. 2005.

**Ziwei Dong** received the B.S. degree in mathematics and applied mathematics from the China University of Mining and Technology, Beijing, China, in 2018. She is currently working toward the Ph.D. degree with the East China University of Science and Technology, Shanghai, China. Her research interests include include distributed optimization and its application in power systems.

**Shuai Mao** received the B.S. degree in automation in 2017 from the East China University of Science and Technology, Shanghai, China, where he is currently working toward the Ph.D. degree. His research interests include multiagent systems, distributed optimization, and their applications in practical engineering.

**Matjaž Perc** is currently a Professor of physics with the University of Maribor, Maribor, Slovenia. He is a Member of Academia Europaea, London, U.K. and the European Academy of Sciences and Arts, Salzburg, Austria, and among top 1% most cited physicists according to 2020 Clarivate Analytics data. He is also the 2015 recipient of the Young Scientist Award for Socio and Econophysics from the German Physical Society, 2017 USERN Laureate, and Zois Award, which is the highest National Research Award in Slovenia in 2018. In 2019 he became Fellow of the American Physical Society.

**Wei Du**, photograph and biography not available at the time of publication.

**Yang Tang** (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Donghua University, Shanghai, China, in 2006 and 2010, respectively. From 2008 to 2010, he was a Research Associate with The Hong Kong Polytechnic University, Hong Kong. From 2011 to 2015, he was a Postdoctoral Researcher with the Humboldt University of Berlin, Berlin, Germany, and Potsdam Institute for Climate Impact Research, Potsdam, Germany. He is currently a Professor with the East China University of Science and Technology, Shanghai, China. His research interests include distributed estimation/control/optimization, cyber-physical systems, hybrid dynamical systems, computer vision, and reinforcement learning and their applications. Prof. Tang was the recipient of the Alexander von Humboldt Fellowship and ISI Highly Cited Researchers Award by Clarivate Analytics in 2017. He is a Senior Board Member of *Scientific Reports*, an Associate Editor for IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-I: REGULAR PAPERS, IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE, and IEEE SYSTEMS JOURNAL AND ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE (IFAC JOURNAL). He is the Leading Guest Editor of special issues in IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE and IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS.