Automatica 123 (2021) 109356

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Distributed economic dispatch via a predictive scheme: Heterogeneous delays and privacy preservation ☆

Fei Chen [a,b,*], Xiaozheng Chen [c], Linying Xiang [b], Wei Ren [d]

[a] State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, 110004, China
[b] School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao, 066004, China
[c] Department of Automation, Xiamen University, Xiamen, Fujian 361005, China
[d] Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521, USA

A R T I C L E I N F O

Article history:
Received 12 November 2018
Received in revised form 11 April 2020
Accepted 15 October 2020
Available online 11 November 2020

Keywords:
Economic dispatch
Heterogeneous time-delay
Privacy preservation
Predictive control
Smart grid

A B S T R A C T

This paper studies distributed economic dispatch problems for smart grids, in which a quadratic generation cost is to be minimized over a feasible set that is determined jointly by an equality constraint and a box constraint. Our primary objective is to seek a distributed design that can handle heterogeneous time-delays, while preserving agents' privacy—a fundamental prerequisite that has become gradually important for cyber–physical systems. For this purpose, we design a state predictor for each agent to compensate for the effect of heterogeneous time-delays, which allows the agents to predict the missing states between two consecutive update times. Based upon the predictor, we present a distributed gradient-descent algorithm to locally update the outputs of the generators, which guarantees that the optimal solution is attained in an asymptotic manner. Among other things, we incorporate a privacy preservation scheme to the proposed algorithm in order to preserve agents' privacy and delicately characterize its convergence, differential privacy properties, as well as accuracy.

1. Introduction

Cooperative control of multi-agent systems (MASs) has received intensive attention in recent years. Typically research problems in MASs include consensus (Olfati-Saber & Murray, 2004), formation control (Ji et al., 2008), and distributed optimization (Liu et al., 2017; Lü et al., 2017). In particular, the last few years have witnessed a growing interest in dispatching power generation for smart grids (Abedini et al., 2013; He et al., 2018; Nagata et al., 2012; Yan et al., 2013). The economic dispatch (ED) problem is one of the most fundamental problems for smart grids, which aims at distributing the total power demand among the generators to minimize the operating cost under certain security constrains.

Earlier works on ED problems could be roughly classified into two categories: convex optimization based approaches and non-convex optimization based approaches. Convex optimization methods include the gradient search approach (Wood & Wollenberg, 2012) and the Newton approach (Yorino et al., 2012). In the non-convex case, heuristic algorithms such as genetic algorithms (Chen & Chang, 1995) and particle swarm optimization algorithms (Park et al., 2010) have been employed. However, most of the existing economic dispatch schemes are centralized, which require global information over the entire power grid. It has been revealed in one way or another that centralized methods are usually expensive to implement and susceptible to one-point failures—a fact known generally; therefore, it is necessary and appealing to develop distributed algorithms for the ED problem.

It has been customary to employ consensus-based algorithms to solve ED problems in a distributed manner, which have led to fully distributed algorithms employing solely local information. There has been a variety of research works along this line of reach in recent years (see, e.g., Li et al., 2018; Wen et al., 2016; Yang et al., 2013; Yu et al., 2015 and the references therein). Ref. Yu et al. (2015) discusses the relationship between the optimal solution of economic dispatch and consensus in smart grids. Ref. Yang et al. (2013) proposes a distributed consensus-based

☆ This work was supported in part by the National Science Foundation of China under Grants 61973061 and 61973064, Natural Science Foundation of Hebei Province of China under Grants F2019501043 and F2019501126, and National Science Foundation under Grant ECCS-1920798. The material in this paper was partially presented at the 43rd Annual Conference of the IEEE Industrial Electronics Society, IECON 2017, October 29–November 1, 2017, Beijing, China (Chen & Chen, 2017). This paper was recommended for publication in revised form by Associate Editor Giancarlo Ferrari-Trecate under the direction of Editor Ian R. Petersen.

* Corresponding author at: State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, 110004, China.

E-mail addresses: fei.chen@ieee.org (F. Chen), dankechan@stu.xmu.edu.cn (X. Chen), xianglinying@neuq.edu.cn (L. Xiang), ren@ee.ucr.edu (W. Ren).

https://doi.org/10.1016/j.automatica.2020.109356
0005-1098/

algorithm to maintain the balance between power generation and demand. In order to handle communication uncertainties in the ED problem, an adaptive consensus-based scheme is proposed and analyzed in Wen et al. (2016) . In Li et al. (2014), a fast distributed gradient-descent method consisting of the θ-logarithmic barrier function is proposed to solve the ED problem. Moreover, an event-triggering scheme is proposed in Li et al. (2016), which seeks to reduce the computational and communication cost for smart grids. The aforementioned works all assume that the network is ideal and reliable; more practical network models (Binetti et al., 2014; Nedić & Olshevsky, 2015; Zhang & Chow, 2012) are also discussed.

Despite the efficiency and autonomy offered by the distributed optimization scheme, smart grids face a challenge in employing the delayed information, subject to possibly different computational time-delays, that are inevitable in distributed systems, particularly when generators are hard to synchronize (see, e.g., Yang et al., 2015; Zhao et al., 2016). If the computation is implemented via a circuit, the delay might be incurred by sensing or circuit wiring. If the algorithm is implemented via a digital computer, there also exists non-ignorable latency or response time, since the computational task needs to compete for the CPU time with other computer tasks, e.g., package routing. Additionally, there are significant privacy risks associated with delayed gradients for applications dealing with sensitive data, such as distributed ED (Mandal, 2016; Zhao et al., 2017). The optimal operating outputs and power demand are significant private information that deserves careful protection, which should be taken into account in distributed ED. This motivates the problem of designing distribute ED algorithms which are robust to heterogeneous delayed gradients and are secure to potential adversaries. Additionally, it is worth pointing out that there exist fundamental challenges due to the coupling amongst the different techniques for prediction, optimization, and privacy preservation, since the separation principle generally fails to work for nonlinear systems.

In this paper, we propose a fully distributed algorithm for the ED problem under heterogeneous time-delays while meeting the requirement of privacy preservation. Specifically, we make the following contributions.

- We construct a state predictor for each agent to attenuate the effect of heterogeneous time-delays. The predictor allows each agent to predict the missing states between two consecutive update times. The predicted values are then employed in the distributed gradient-descent algorithm to guarantee the desired convergence properties.
- We propose a privacy-preserved predictive scheme by adding noise to the output information to meet the requirement of privacy preservation under heterogeneous time-delays. Among other things, we carefully characterize its convergence, differential privacy properties, as well as accuracy.

The rest of this paper is organized as follows. In Section 2, we provide mathematical preliminaries. In Section 3, we describe the ED problem of this paper. In Section 4, we propose a predictive control scheme to compensate for the effect of heterogeneous time-delays and investigate its convergence. In Section 5, we present the privacy-preserved predictive scheme to meet the requirement of privacy preservation and analyze its performance. Finally, Section 6 concludes this paper.

2. Mathematical preliminaries

The notation used throughout this paper is fairly standard. We denote the set of real numbers by $\mathbb{R}$, the set of positive real numbers by $\mathbb{R}^+$, the set of N-dimensional real vectors by $\mathbb{R}^N$, and the set of $N\times N$ real matrices by $\mathbb{R}^{N\times N}$. We denote the identity matrix by I. We denote the column vectors of all ones and all zeros by $\mathbf{1}$ and $\mathbf{0}$, respectively. For a vector $x\in\mathbb{R}^N$, we denote the two norm by $\|x\|_2=(|x_1|^2+\cdots+|x_N|^2)^{1/2}$. We denote by $\mathrm{diag}\{a_1,\ldots,a_N\}$ the diagonal matrix with a_i, $i=1,\ldots,N$, being its ith diagonal element. We denote the transpose over a vector or matrix by the superscript T. We denote the gradient and Hessian of a function by ∇f and $\nabla^2 f$, respectively. A twice differentiable function $f:\mathbb{R}\to\mathbb{R}$ is said to be strongly convex if there exists a constant $m>0$ such that $\nabla^2 f\geq m$. For simplicity, we denote events of the type $E_r=\{\omega\in\Omega\mid r(\omega)\}$ by $\{r\}$, where r is a logical statement on the elements of Ω, and Ω denotes the total sample space of the noise sequence ω. We denote the probability of E_r by $\mathbb{P}\{E_r\}$. For a random variable X, we denote its expectation and variance by $\mathbb{E}(X)$ and $\mathrm{Var}(X)$, respectively.

We denote a graph by $\mathcal{G}=\{\mathcal{V},\mathcal{E}\}$, where $\mathcal{V}=\{1,2,\ldots,N\}$ is the node set, and $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ is the edge set. The graph with the property that $(i,j)\in\mathcal{E}\Leftrightarrow(j,i)\in\mathcal{E}$ is said to be undirected. The set of neighbors of the ith node is denoted by $\mathcal{N}_i=\{j\in\mathcal{V}\mid(j,i)\in\mathcal{E}\}$. The adjacency matrix $\mathcal{A}=[a_{ij}]\in\mathbb{R}^{N\times N}$ of $\mathcal{G}$ is defined as follows: $a_{ij}=1$ if $(i,j)\in\mathcal{E}$; $a_{ij}=0$ otherwise. The degree of node i is defined as $d_i=\sum_{j=1}^N a_{ij}$, and the degree matrix is defined as $\mathcal{D}=\mathrm{diag}\{d_1,\ldots,d_N\}\in\mathbb{R}^{N\times N}$. The Laplacian matrix of $\mathcal{G}$ is given by $\mathcal{L}=\mathcal{D}-\mathcal{A}\in\mathbb{R}^{N\times N}$. In this paper, we use an undirected graph $\mathcal{G}$ to describe the information flows among the power grid, where each node represents a generator, and an edge (i,j) denotes an information channel between generators i and j. To design distributed ED algorithms, the connectedness of the graph is necessary to postulate.

Assumption 1. Graph $\mathcal{G}$ is connected.

3. Problem formulation

Consider an N-generator power grid. The generation cost for the ith generator is given by the quadratic function $C_i(P_i)=\alpha_i {P_i}^2+\beta_i P_i+\gamma_i$, where P_i denotes the output of the ith power generator, and α_i, β_i, and $\gamma_i\in\mathbb{R}$ are some non-negative constants (Wood & Wollenberg, 2012). The ED problem is defined as

$$\min_{P_i}\quad C(P)=\sum_{i=1}^N C_i(P_i) \tag{1a}$$

$$\text{s.t.}\quad \sum_{i=1}^N P_i=P_d \tag{1b}$$

$$P_{i,\min}\leq P_i\leq P_{i,\max}, \tag{1c}$$

where P_d is the total power demand, and $P_{i,\min}$ and $P_{i,\max}$ are, respectively, the lower and upper bound of the ith generator's output capacity. Each generator should work under its capacity constrain, i.e., $P_i\in[P_{i,\min},P_{i,\max}]$. Particularly, we denote the global optimal solution of the ED problem (1) by $P^*=[P_1^*,\ldots,P_N^*]^T$. In practice, there always exist time-delays. It is now well recognized that time-delays are not negligible and might destroy system stability (Gu et al., 2003). Motivated by the fact, we consider the following distributed ED model with heterogeneous time-delays:

$$P_i(K+1)=P_i(K)+U_i\left(\{F_j(K-\tau_j)\}_{j\in\mathcal{N}_i\cup\{i\}}\right), \tag{2}$$

where $P_i(K)$ is the output of the ith generator at time K, U_i is the control input of the ith generator, which depends solely upon $\{F_j(K-\tau_j)\}_{j\in\mathcal{N}_i\cup\{i\}}$, the delayed information of generator i and its neighbors, and τ_j is the time-delay of agent j. Note that for discrete-time dynamical systems, computation time-delay is also a central issue that needs to be addressed before the algorithm can be implemented in real-time. We are interested in designing the control input U_i for the delayed system (2) such that

$$\lim_{K\to\infty} \left\| P(K) - P^* \right\|_2 \le \xi, \tag{3}$$

where $P(K) = [P_1(K), \dots, P_N(K)]^T$, and ξ can be arbitrarily small.

In what follows, we relax the box constraint (1c) by employing the θ-logarithmic barrier function (Fiacco & McCormick, 1990) and transform the optimization problem (1) to

$$\min_{P_i} \quad f(P) = \sum_{i=1}^{N} f_i(P_i) \tag{4a}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} P_i = P_d \tag{4b}$$

where $f_i(P_i) = C_i(P_i) - \frac{1}{\theta}\big(\ln(P_i - P_{i,\min}) + \ln(P_{i,\max} - P_i)\big)$. It is noticed that the penalized objective function (4a) is finite only if the box constraints (1c) are strictly satisfied. Under the barrier function method, each generator should be initialized in a feasible domain with a sufficiently large parameter θ. We denote the global optimal solution to the ED problem (4) by $P^\star = [P_1^\star, \dots, P_N^\star]^T$. Under the proposed method, P_i will never reach $P_{i,\min}$ or $P_{i,\max}$. Consequently, if the optimal solution P_i^* to (4) is $P_{i,\min}$ or $P_{i,\max}$, the optimal point will not be achieved. Instead, a suboptimal point $P_i^\star$ will be achieved, whose difference with the true optimum is upper bounded by $\frac{2N}{\theta}$ (see Li et al., 2016). Note that when θ goes to infinity, the optimal solution $P^\star$ of (4) will converge to the optimal solution P^* of the original problem (1).

The Hessian Matrix of the penalized objective function (4a) is given by $\nabla^2 f = \begin{bmatrix} 2\alpha_1+M_1 & 0 & \cdots & 0 \\ 0 & 2\alpha_2+M_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2\alpha_N+M_N \end{bmatrix}$, where $2\alpha_i + M_i$, $i \in \{1, 2, \dots, N\}$ is the second-order derivative of $f_i(P_i)$ with respect to P_i, and $M_i = \frac{1}{\theta}\left(\frac{1}{(P_i-P_{i,\min})^2} + \frac{1}{(P_{i,\max}-P_i)^2}\right)$. Since α_i is positive, and P_i is an interior point of the feasible domain, $2\alpha_i + M_i$ is lower bounded by $2\alpha_i$. In practical implementation, one should guarantee that P_i does not get arbitrarily close to the bounds $P_{i,\min}$ and $P_{i,\max}$. This is usually achieved by setting up tighter bounds, e.g., $P_{i,\min}+\epsilon$ and $P_{i,\max}-\epsilon$, where ϵ is a small positive constant. In this way, the output variable P_i will keep a distance of, at least, ϵ from $P_{i,\min}$ and $P_{i,\max}$, which insures that there exists $\overline{M}_i > 0$ such that $2\alpha_i + M_i \le \overline{M}_i$ (Li et al., 2014). Therefore, we have

$$\Pi \le \nabla^2 f \le \Psi, \tag{5}$$

with $\Pi = \mathrm{diag}\{2\alpha_1, \dots, 2\alpha_N\}$, $\Psi = \mathrm{diag}\{\overline{M}_1, \dots, \overline{M}_N\}$.

4. The predictive control scheme

In this section, we design the following predictive control law to attenuate the effect of time-delays:

$$U_i(K) = -\zeta \sum_{j=1}^{N} l_{ij} \nabla f_j\big(\hat{P}_j(K|K-\tau_j)\big), \tag{6}$$

where $U_i(K)$ is the control input of agent i, $\hat{P}_j(K|K-\tau_j)$ is the prediction of agent j's output at time K using the available information up to $K-\tau_j$, and $\zeta > 0$ is a control gain. Substituting (6) into (2) yields the following closed-loop system

$$P_i(K+1) = P_i(K) - \zeta \sum_{j=1}^{N} l_{ij} \nabla f_j\big(\hat{P}_j(K|K-\tau_j)\big) \tag{7}$$

where the predictor is designed as

$$\begin{aligned} &\hat{P}_i(K-\tau_i+1|K-\tau_i) \\ =&\hat{P}_i(K-\tau_i|K-\tau_i-1) + k_f \Big(P_i(K-\tau_i) \\ &- \hat{P}_i(K-\tau_i|K-\tau_i-1) \Big) + U_i(K-\tau_i). \end{aligned} \tag{8}$$

In (8), τ_i denotes the time-delay for agent i to process the information and calculate its incremental cost $\nabla f_i(\cdot)$, whereas $\hat{P}_i(K-\tau_i+1|K-\tau_i)$ denotes agent i's prediction of its next output value given the information up to time $K-\tau_i$. The predictor consists of three terms: the prediction at the last time $\hat{P}_i(K-\tau_i|K-\tau_i-1)$, the prediction error feedback $k_f(P_i(K-\tau_i) - \hat{P}_i(K-\tau_i|K-\tau_i-1))$, and agent i's input $U_i(K-\tau_i)$, where k_f is a predictor gain to be designed. In practical implementation, the prediction $\hat{P}_i(K-\tau_i|K-\tau_i-1)$ is calculated at time $K-\tau_i-1$. Consequently, the prediction error feedback can also be obtained at time K. The input $U_i(K-\tau_i)$ is obtained by agent i by summing up all its neighboring delayed information.

The outputs during the time interval $[K-\tau_i+2, K]$ can be predicted by

$$\begin{aligned} \hat{P}_i(K-\tau_i+k|K-\tau_i) =& \hat{P}_i(K-\tau_i+k-1|K-\tau_i) \\ &+U_i(K-\tau_i+k-1), \quad k = 2, \dots, \tau_i, \end{aligned} \tag{9}$$

where $\hat{P}_i(K-\tau_i+k|K-\tau_i)$ and $U_i(K-\tau_i+k-1)$ are defined similarly. Instead of sending the delayed information directly, generator i will send its predicted information (8) and (9) to its neighbors, which will compensate for the effect of the time-delay. Note that the prediction $\hat{P}_i(K|K-\tau_i)$ is calculated in the interval $[K-\tau_i, K-1]$, during which the inputs $U_i(K-\tau_i), \dots U_i(K-1)$ are available.

Replacing K with $K+\tau_i$ in (8) yields

$$\begin{aligned} \hat{P}_i(K+1|K) =& \hat{P}_i(K|K-1) + k_f \Big(P_i(K) \\ &- \hat{P}_i(K|K-1) \Big) + U_i(K). \end{aligned} \tag{10}$$

Let $\delta_i(K) = P_i(K) - \hat{P}_i(K|K-1)$. Subtracting (10) from (7) leads to

$$\delta_i(K+1) = (1-k_f)\delta_i(K). \tag{11}$$

Applying (9) recursively gives

$$\begin{aligned} &\hat{P}_i(K|K-\tau_i) \\ =&\hat{P}_i(K-\tau_i|K-\tau_i-1) + \sum_{k=1}^{\tau_i} U_i(K-\tau_i+k-1) \\ &+ k_f\big(P_i(K-\tau_i) - \hat{P}_i(K-\tau_i|K-\tau_i-1)\big). \end{aligned} \tag{12}$$

Similarly, applying (7) recursively yields

$$P_i(K) = P_i(K-\tau_i) + \sum_{k=1}^{\tau_i} U_i(K-\tau_i+k-1). \tag{13}$$

Subtracting (13) from (12), we have

$$\begin{aligned} &\hat{P}_i(K|K-\tau_i) \\ =&P_i(K) + \hat{P}_i(K-\tau_i|K-\tau_i-1) - P_\tau(K-\tau_i) \\ &+ k_f\big(P_i(K-\tau_i) - \hat{P}_i(K-\tau_i|K-\tau_i-1)\big) \\ =&P_i(K) - \delta_i(K-\tau_i) + k_f\delta_i(K-\tau_i) \\ =&P_i(K) - \delta_i(K-\tau_i+1). \end{aligned} \tag{14}$$

The main results of this section are stated as follows.

Theorem 1. *If Assumption 1 holds,*

$$\zeta < \frac{4\lambda_2(\mathcal{L})}{(\max_{i=1,\dots,N} \overline{M}_i + 1)\lambda_N^2(\mathcal{L})},$$

where $\lambda_2(\mathcal{L})$ and $\lambda_N(\mathcal{L})$ are the smallest and largest nonzero eigenvalues of the Laplacian matrix $\mathcal{L}$, and $1-\frac{\sqrt{2}\min_{i=1,\ldots,N}\alpha_i}{\max_{i=1,\ldots,N}\overline{M}_i} \le k_f < 1$, then under the system (7) and the predictor (8) and (9), the control objective (3) is achieved asymptotically where $\xi = \sqrt{\frac{N\max_{i=1,\ldots,N}\overline{M}_i}{\theta\min_{i=1,\ldots,N}\alpha_i^2}}$.

Proof. Let $\triangle P_i(K) = P_i(K+1) - P_i(K)$ and $\triangle P(K) = [\triangle P_1(K), \ldots, \triangle P_N(K)]^T$. It follows from (7) that

$$\triangle P(K) = P(K+1) - P(K) = -\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \tag{15}$$

where $\hat{P}_\tau(K) = [\hat{P}_1(K|K-\tau_1), \ldots, \hat{P}_N(K|K-\tau_N)]^T$, and $\nabla f\left(\hat{P}_\tau(K)\right) = \left[\nabla f_1\left(\hat{P}_1(K|K-\tau_1)\right), \ldots, \nabla f_N\left(\hat{P}_N(K|K-\tau_N)\right)\right]^T$. A Taylor expansion of $f(\cdot)$ yields

$$\begin{aligned} f(P(K+1)) = &f\left(P(K)\right) + \nabla f\left(P(K)\right)^T \triangle P(K) \\ &+\frac{1}{2}\triangle P(K)^T \nabla^2 f\left(z(K)\right)\triangle P(K), \end{aligned} \tag{16}$$

where $z(K) = [z_1(K), \ldots, z_N(K)]^T$ with $z_i(K) \in [P_i(K), P_i(K+1)]$, and $\nabla^2 f\left(z(K)\right)$ is the Hessian Matrix of the penalized objective function (4a) at $z(K)$. Substituting (15) into (16) yields

$$\begin{aligned} f\left(P(K+1)\right) = &f\left(P(K)\right) - \nabla f\left(P(K)\right)^T \zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &+\frac{1}{2}\nabla f\left(\hat{P}_\tau(K)\right)^T \mathcal{L}^T\zeta\nabla^2 f(z(K))\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right). \end{aligned} \tag{17}$$

Rewrite (14) in a vector form as

$$\hat{P}_\tau(K) = P(K) - \delta_\tau(K), \tag{18}$$

where $\delta_\tau(K) = [\delta_1(K-\tau_1+1), \ldots, \delta_N(K-\tau_N+1)]^T$. Due to (18), (17) can be written as $f\left(P(K+1)\right) = f\left(P(K)\right) - \nabla f\left(\hat{P}_\tau(K)+\delta_\tau(K)\right)^T \zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) + \frac{1}{2}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta\nabla^2 f\left(z(K)\right)\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right)$. It follows from the mean-value theorem (Khalil, 1996) that

$$\begin{aligned} &\nabla f_i\left(\hat{P}_i(K|K-\tau_i) + \delta_i(K-\tau_i+1)\right) \\ &-\nabla f_i\left(\hat{P}_i(K|K-\tau_i)\right) = \varepsilon_i(K)\delta_i(K-\tau_i+1), \end{aligned} \tag{19}$$

where $2\alpha_i \le \varepsilon_i(K) = \nabla^2 f_i(y_i(K)) \le \overline{M}_i$, and $y_i(K)$ is some value between $P_i(K)$ and $\hat{P}_i(K|K-\tau_i)$. Rewrite (19) in a vector form as

$$\nabla f\left(P(K)\right) = \nabla f\left(\hat{P}_\tau(K)\right) + \varepsilon(K)\delta_\tau(K), \tag{20}$$

where $\varepsilon(K) = \mathrm{diag}\{\varepsilon_i(K), \ldots, \varepsilon_N(K)\}$. It follows from (17) and (20) that

$$\begin{aligned} &f\left(P(K+1)\right) - f\left(P(K)\right) \\ =& -\nabla f\left(\hat{P}_\tau(K)\right)^T\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &-\left(\varepsilon(K)\delta_\tau(K)\right)^T\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &+\frac{1}{2}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta\nabla^2 f\left(z(K)\right)\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right). \end{aligned} \tag{21}$$

Define the Lyapunov function candidate $V(K) = f\left(P(K)\right) - f(P^\star) + 2\|\varepsilon(K)\delta_\tau(K)\|_2^2$ and $\triangle V(K+1) = V(K+1) - V(K)$. It follows that

$$\begin{aligned} &\triangle V(K+1) = f\left(P(K+1)\right) - f\left(P(K)\right) \\ &+2\|\varepsilon(K+1)\delta_\tau(K+1)\|_2^2 - 2\|\varepsilon(K)\delta_\tau(K)\|_2^2. \end{aligned} \tag{22}$$

Substituting (21) into (22) yields

$$\begin{aligned} &\triangle V(K+1) \\ =& -\nabla f\left(\hat{P}_\tau(K)\right)^T\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &-\left(\varepsilon(K)\delta_\tau(K)\right)^T\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &+\frac{1}{2}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta\nabla^2 f\left(z(K)\right)\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &+2\|\varepsilon(K+1)\delta_\tau(K+1)\|_2^2 - 2\|\varepsilon(K)\delta_\tau(K)\|_2^2. \end{aligned} \tag{23}$$

It thus follows that $\left(\varepsilon(K)\delta_\tau(K)\right)^T\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) + \|\varepsilon(K)\delta_\tau(K)\|_2^2 = \left\|\varepsilon(K)\delta_\tau(K) + \frac{1}{2}\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right)\right\|_2^2 - \frac{1}{4}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta^2\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right)$. Hence, (23) can be written as $\triangle V(K+1) = -\nabla f\left(\hat{P}_\tau(K)\right)^T\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) - \left\|\varepsilon(K)\delta_\tau(K) + \frac{1}{2}\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right)\right\|_2^2 + \frac{1}{4}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta^2\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) + \frac{1}{2}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta\nabla^2 f\left(z(K)\right)\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) + 2\|\varepsilon(K+1)\delta_\tau(K+1)\|_2^2 - \|\varepsilon(K)\delta_\tau(K)\|_2^2$, leading to

$$\begin{aligned} &\triangle V(K+1) \\ \le& -\nabla f\left(\hat{P}_\tau(K)\right)^T\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &+\frac{1}{4}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta^2\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &+\frac{1}{2}\nabla f\left(\hat{P}_\tau(K)\right)^T\mathcal{L}^T\zeta\Psi\zeta\mathcal{L}\nabla f\left(\hat{P}_\tau(K)\right) \\ &+2\|\varepsilon(K+1)\delta_\tau(K+1)\|_2^2 - \|\varepsilon(K)\delta_\tau(K)\|_2^2 \\ =& -\frac{1}{4}\nabla f\left(\hat{P}_\tau(K)\right)^T\left(4\zeta\mathcal{L} - 2\mathcal{L}^T\zeta\Psi\zeta\mathcal{L} - \mathcal{L}^T\zeta^2\mathcal{L}\right) \\ &\nabla f\left(\hat{P}_\tau(K)\right) + 2\|\varepsilon(K+1)\delta_\tau(K+1)\|_2^2 \\ &-\|\varepsilon(K)\delta_\tau(K)\|_2^2. \end{aligned} \tag{24}$$

By (11), (24) can be written as

$$\begin{aligned} &\triangle V(K+1) \\ \le& -\frac{1}{4}\nabla f\left(\hat{P}_\tau(K)\right)^T Q\nabla f\left(\hat{P}_\tau(K)\right) \\ &+2(1-k_f)^2\|\varepsilon(K+1)\delta_\tau(K)\|_2^2 - \|\varepsilon(K)\delta_\tau(K)\|_2^2, \end{aligned} \tag{25}$$

where $Q = \zeta(4\mathcal{L} - 2\zeta\mathcal{L}^T M\mathcal{L} - \zeta\mathcal{L}^T\mathcal{L})$. By Assumption 1, $\mathcal{L}$ is positive semi-definite with a single zero eigenvalue, i.e., $0 = \lambda_1(\mathcal{L}) < \lambda_2(\mathcal{L}) \le \cdots \le \lambda_N(\mathcal{L})$. Additionally, notice that $2\zeta\mathcal{L}^T M\mathcal{L} + \zeta\mathcal{L}^T\mathcal{L} = \zeta\mathcal{L}^T(2M+I)\mathcal{L}$. Because $0 < \zeta(2\min_{i=1,\ldots,N}\overline{M}_i+1)\lambda_i^2(\mathcal{L}) \le \lambda_i\left(\zeta\mathcal{L}^T(2M+I)\mathcal{L}\right) \le \zeta(2\max_{i=1,\ldots,N}\overline{M}_i+1)\lambda_i^2(\mathcal{L})$ for $i = 2, \ldots, N$, we have $4\zeta\lambda_i(\mathcal{L}) - \zeta^2(2\max_{i=1,\ldots,N}\overline{M}_i+1)\lambda_N^2(\mathcal{L}) \le \lambda_i(Q) \le 4\zeta\lambda_i(\mathcal{L}) - \zeta^2(2\min_{i=1,\ldots,N}\overline{M}_i+1)\lambda_2^2(\mathcal{L})$ for $i = 2, \ldots, N$. If $\zeta < \frac{4\lambda_2(\mathcal{L})}{(\max_{i=1,\ldots,N}\overline{M}_i+1)\lambda_N^2(\mathcal{L})}$, we have $4\zeta\lambda_i(\mathcal{L}) - \zeta^2(\max_{i=1,\ldots,N}\overline{M}_i+1)\lambda_N^2(\mathcal{L}) > 0$, which indicates that $0 = \lambda_1(Q) < \lambda_2(Q) \le \cdots \le \lambda_N(Q)$, i.e., the matrix Q can be made positive semi-definite.

Let $e(K)$ denote the projection of $\nabla f\left(\hat{P}(K)\right)$ to $\mathbf{1}$, i.e.,

$$\begin{aligned} e(K) =& \left(I - \frac{\mathbf{1}\mathbf{1}^T}{N}\right)\nabla f\left(\hat{P}_\tau(K)\right) \\ =& \nabla f\left(\hat{P}_\tau(K)\right) - \frac{\mathbf{1}^T\nabla f\left(\hat{P}_\tau(K)\right)}{N}\mathbf{1}. \end{aligned} \tag{26}$$

Noting that $\mathcal{L}\mathbf{1} = \mathbf{0}$, the substitution of (26) into (25) yields $\triangle V(K+1) \le -\frac{1}{4}e^T(K)Qe(K) + 2(1-k_f)^2\|\varepsilon(K+1)\delta_\tau(K)\|_2^2 - \|\varepsilon(K)\delta_\tau(K)\|_2^2$. If k_f is chosen to satisfy $1 - \frac{\sqrt{2}\min_{i=1,\ldots,N}\alpha_i}{\max_{i=1,\ldots,N}\overline{M}_i} \le k_f < 1$, we have

$$2(1-k_f)^2\|\varepsilon(K+1)\delta_\tau(K)\|_2^2 - \|\varepsilon(K)\delta_\tau(K)\|_2^2 \le 0,$$

which implies that $\triangle V(K+1) \le 0$, and $\triangle V(K+1) < 0$ only if $e(K) \ne \mathbf{0}$ or $\delta_\tau(K) \ne \mathbf{0}$. Consequently, we have $e(K) \to \mathbf{0}$ and $\delta_\tau(K) \to \mathbf{0}$. It follows from $e(K) \to \mathbf{0}$ and (26) that

$$\nabla f\left(\hat{P}_\tau(K)\right) \to \alpha\mathbf{1}, \tag{27}$$

for some α. Additionally, Eq. (18) and $\delta_\tau(K) \to \mathbf{0}$ yield

$$\hat{P}_\tau(K) \to P(K) \text{ as } K \to \infty. \tag{28}$$

Hence, we can conclude from (27) and (28) that $\nabla f\left(P(K)\right) \to \alpha\mathbf{1}$, as $K \to \infty$, which indicates that $P^\star$ is attained asymptotically (Wood & Wollenberg, 2012).

Next, we drive the optimization error bound ξ. It follows from (5) that (cf. Li et al., 2016)

$$\frac{2}{\max_{i=1,\ldots,N}\overline{M}_i}\left\|\nabla f\big(P(K)\big)\right\|_2^2 \le f\big(P(K)\big)-f(P^*). \tag{29}$$

The strong convexity property (5) also yields

$$\begin{aligned} f(P^*) \ge& f\big(P(K)\big)-\left\|\nabla f\big(P(K)\big)\right\|_2\left\|P(K)-P^*\right\|_2 \\ &+\min_{i=1,\ldots,N}\alpha_i\left\|P(K)-P^*\right\|_2^2. \end{aligned} \tag{30}$$

Noting that $f(P^*)\le f\big(P(K)\big)$, it follows from (30) that

$$\left\|P(K)-P^*\right\|_2 \le \frac{1}{\min_{i=1,\ldots,N}\alpha_i}\left\|\nabla f\big(P(K)\big)\right\|_2, \tag{31}$$

which leads to

$$\left\|P(K)-P^*\right\|_2 \le \sqrt{\frac{\max_{i=1,\ldots,N}\overline{M}_i}{2\min_{i=1,\ldots,N}\overline{\alpha}_i^2}\big(f\big(P(K)\big)-f(P^*)\big)}.$$

Noting $f\big(P(K)\big)\to f(P^\star)$ and $\limsup_{K\to\infty}\big(f(P^\star)-f(P^*)\big)=\frac{2N}{\theta}$, one has $\limsup_{K\to\infty}\left\|P(K)-P^*\right\|_2 \le \limsup_{K\to\infty}\sqrt{\frac{\max_{i=1,\ldots,N}\overline{M}_i}{2\min_{i=1,\ldots,N}\alpha_i^2}\big(f\big(P(K)\big)-f(P^*)\big)} \le \sqrt{\frac{N\max_{i=1,\ldots,N}\overline{M}_i}{\theta\min_{i=1,\ldots,N}\alpha_i^2}}=\xi$. □

It follows from Theorem 1 that the control error ξ can be made arbitrarily small by choosing sufficiently large θ. Additionally, the more convex the objective function, the smaller the control error.

5. Privacy preservation

When the generators do not fully trust each other or the communication channels are insecure, sending the exact (predicted) gradient information to neighbors induces a risk on the system. That is, the initial states, which contain significant information about the optimal solution of the system (see Binetti et al., 2014), can be inferred by an adversary from their execution. In this section, we extend the results in the last section by adding random noise to the system (7) in order to protect agent privacy. Specifically, we propose the following privacy preserving scheme for the ED problem.

Algorithm 1 Differential Privacy Preserving Scheme

(1) At time K, each agent generates a random noise $\omega_i(K)\sim \mathrm{Lap}(\varphi_i^K)$. That is, $\omega_i(K)$ follows the Laplace distribution $\mathrm{PDF}\big(\omega_i(K)\big)=\frac{1}{2\varphi_i^K}\exp\left(-\frac{|\omega_i(K)|}{\varphi_i^K}\right)$, where $0<\varphi_i<1$ and PDF denotes probability density function.

(2) Each agent uses a "noisy" version of the incremental cost, i.e.,

$$\nabla\widetilde{f}_i(K)=\nabla f_i\big(\hat{P}_i(K|K-\tau_i)\big)+\omega_i(K). \tag{32}$$

(3) Each agent updates its output by

$$P_i(K+1)=P_i(K)-\zeta\sum_{j=1}^N l_{ij}\nabla\widetilde{f}_j(K). \tag{33}$$

(4) Increase K by one, and go to step 1).

Define the observation sequence executed from $P(0)$ as follows: $\mathcal{I}^{out}_{P(0)}\big(\omega(K)\big)=[\nabla\widetilde{f}_1(K),\ldots,\nabla\widetilde{f}_N(K)]^T$, where $\omega(K)=[\omega_1(K),\ldots,\omega_N(K)]^T$. Let $\mathcal{I}^{out}_{P(0)}(\omega)=\{\mathcal{I}^{out}_{P(0)}\big(\omega(0)\big),\mathcal{I}^{out}_{P(0)}\big(\omega(1)\big),\mathcal{I}^{out}_{P(0)}\big(\omega(2)\big),\ldots\}$. In the following, we formally introduce the definition of differential privacy (Nozari et al., 2017).

Definition 1 (*Differential Privacy*)**.** Given $\sigma>0$, the initial states $P'(0)$ and $P''(0)$ are σ-adjacent if and only if, there exists $i_0\in\mathcal{V}$, $\left|P_i'(0)-P_i''(0)\right|\le\begin{cases}\sigma & i=i_0,\\ 0 & i\ne i_0,\end{cases}$ where $i\in\mathcal{V}$. Given $\{\sigma,\epsilon\}\subseteq\mathbb{R}^+$, we say the dynamics (33) preserves ϵ-differential privacy if, for any pair of σ-adjacent initial states $P'(0)$ and $P''(0)$ and any arbitrary set S, $\mathbb{P}\left\{\omega\in\Omega\mid\mathcal{I}^{out}_{P'(0)}(\omega)\in S\right\}\le\exp(\epsilon)\mathbb{P}\left\{\omega\in\Omega\mid\mathcal{I}^{out}_{P''(0)}(\omega)\in S\right\}$.

Remark 1. In (32) and (33), $\omega_i(K)$ can be replaced with the delayed version, e.g., $\omega_i(K-\tau_i)$, which does not affect the validity of the main results, since the random variables $\omega_i(K)$ are independently identically distributed.

Note that the initial states of the generators contain the important power demand information. That is, $P_d=\sum_{i=1}^N P_i(0)$. Additionally, it is worth pointing out that the optimal outputs also depend on the initial states, since $P_i^*=\frac{\lambda^*-\beta_i}{2\gamma_i}$ with $\lambda^*=\left(\sum_{i=1}^N P_i(0)+\sum_{i=1}^n\frac{\beta_i}{2\gamma_i}\right)\Big/\left(\sum_{i=1}^n\frac{1}{2\gamma_i}\right)$.

Definition 2 (*Accuracy*)**.** For $r\ge0$ and $p\in(0,1)$, Algorithm 1 is said to be (p,r)-accurate if, for any feasible initial state $P(0)$, the agents' state $P(K)$ converges to $P(\infty)$ as $K\to\infty$ with $\mathbb{E}\big(P(\infty)\big)=P^\star$, and $\mathbb{P}\{\|P(\infty)-P^\star\|_2\le r\}\ge1-p$.

It is ready to state the main results of this section.

Theorem 2. *If Assumption 1 holds,*

$$\zeta<\min\left(\frac{4\lambda_2(\mathcal{L})}{\max_{i=1,\ldots,N}\overline{M}_i\lambda_N^2(\mathcal{L})},\frac{4}{\min_{i=1,\ldots,N}\alpha_i\lambda_N(\mathcal{L})}\right), \tag{34}$$

$0<k_f<1$, and $1-k_f<\varphi_i<1$, then Algorithm 1 guarantees that $\mathbb{E}\big(f\big(P(K)\big)\big)\to f(P^\star)$ as $K\to\infty$, while preserving ϵ-differential privacy with $\epsilon=\max_{i=1}^N\frac{\overline{M}_i\sigma\varphi_i}{\varphi_i-(1-k_f)}$. Additionally, for any $p\in(0,1)$, Algorithm 1 guarantees (p,r)-accuracy, where $r=\sum_{i=1}^N\sum_{j=1}^N\frac{\sqrt{2\zeta^2 l_{ij}^2}}{\sqrt{p(1-\varphi_j^2)}}$.

Proof. We first show that $\mathbb{E}\big(f\big(P(K)\big)\big)\to f(P^\star)$ as $K\to\infty$. Let $\triangle P_i(K)=P_i(K+1)-P_i(K)$. It follows from (33) that $\triangle P(K)=P(K+1)-P(K)=-\zeta\mathcal{L}\nabla\widetilde{f}(K)$, where $\triangle P(K)=[\triangle P_1(K),\ldots,\triangle P_N(K)]^T$, $\nabla\widetilde{f}(K)=\nabla f\big(\hat{P}_\tau(K)\big)+\omega(K)$. Similar to the proof of Theorem 1, we can write the objective function as $f\big(P(K+1)\big)=f\big(P(K)\big)-\nabla\widetilde{f}(K)^T\zeta\mathcal{L}\nabla\widetilde{f}(K)+\frac{1}{2}\nabla\widetilde{f}(K)^T\mathcal{L}^T\zeta\nabla^2 f\big(z(K)\big)\zeta\mathcal{L}\nabla\widetilde{f}(K)-\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T\zeta\mathcal{L}\nabla\widetilde{f}(K)+\omega(K)^T\zeta\mathcal{L}\nabla\widetilde{f}(K)$, where $\widetilde{\varepsilon}(K)=\mathrm{diag}\{\widetilde{\varepsilon}_1(K),\cdots,\widetilde{\varepsilon}_N(K)\}$, $2\alpha_i\le\widetilde{\varepsilon}_i(K)=\nabla^2 f_i(\widetilde{y}_i(K))\le\overline{M}_i$, and $\widetilde{y}_i(K)$ is some value between $P_i(K)$ and $\hat{P}_i(K|K-\tau_i)$. Due to the strong convexity of f, we have

$$\begin{aligned} f\big(P(K+1)\big)\le& f\big(P(K)\big)-\frac{1}{2}\nabla\widetilde{f}(K)^T W\nabla\widetilde{f}(K) \\ &-\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T\zeta\mathcal{L}\nabla\widetilde{f}(K) \\ &+\omega(K)^T\zeta\mathcal{L}\nabla\widetilde{f}(K), \end{aligned} \tag{35}$$

where $W=\zeta(2\mathcal{L}-\zeta\mathcal{L}^T\Psi\mathcal{L})$, and Ψ is defined in (5). The inequality $0<(\min_{i=1,\ldots,N}\overline{M}_i)\lambda_i^2(\mathcal{L})\le\lambda_i(\mathcal{L}^T\Psi\mathcal{L})\le(\max_{i=1,\ldots,N}\overline{M}_i)\lambda_i^2(\mathcal{L})$ yields $2\zeta\lambda_i(\mathcal{L})-\zeta^2(\max_{i=1,\ldots,N}\overline{M}_i)\lambda_N^2\le\lambda_i(W)\le2\zeta\lambda_i(\mathcal{L})-\zeta^2\times(\min_{i=1,\ldots,N}\overline{M}_i)\lambda_2^2(\mathcal{L})$, for $i=2,\ldots,N$. If ζ is chosen to satisfy $\zeta<\frac{2\lambda_2(\mathcal{L})}{(\max_{i=1,\ldots,N}\overline{M}_i)\lambda_N^2(\mathcal{L})}$, we have $2\zeta\lambda_i(\mathcal{L})-\zeta^2\left(\max_{i=1,\ldots,N}\overline{M}_i\right)\lambda_N^2(\mathcal{L})>0$. Moreover, if ζ is chosen to satisfy $\zeta<\frac{4}{\min_{i=1,\ldots,N}\alpha_i\lambda_N(\mathcal{L})}$, we have $\lambda_2(W)<\frac{4}{\min_{i=1,\ldots,N}\alpha_i}$. Hence by choosing ζ as in (34), the matrix W can be made positive semi-definite and $\lambda_2(W)<\frac{4}{\min_{i=1,\ldots,N}\alpha_i}$.

Eq. (35) yields $f\big(P(K+1)\big) \le f\big(P(K)\big) - \frac{1}{2}\lambda_2(W)\|\nabla\widetilde{f}(K)\|_2^2 - \big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T\zeta\mathcal{L}\nabla\widetilde{f}(K) + \omega(K)^T\zeta\mathcal{L}\nabla\widetilde{f}(K)$. Applying the Taylor expansion of $f_i(\cdot)$ at $P_i(K)$ leads to

$$\begin{aligned} f(P^\star) \ge & f\big(P(K)\big) + \nabla f\big(P(K)\big)^T\big(P^\star - P(K)\big) \\ & + \frac{1}{2}\big(P^\star - P(K)\big)^T \Pi \big(P^\star - P(K)\big), \end{aligned} \tag{36}$$

where Π is defined in (5). It follows from (20) and (32) that $\nabla f\big(P(K)\big) = \nabla\widetilde{f}(K) - \omega(K) + \widetilde{\varepsilon}(K)\delta_\tau(K)$, substituting which into (36) leads to

$$\begin{aligned} f(P^\star) \ge & f\big(P(K)\big) + \nabla\widetilde{f}(K)^T\big(P^\star - P(K)\big) \\ & + \Big(\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T - \omega(K)^T\Big)\big(P^\star - P(K)\big) \\ & + \frac{1}{2}\big(P^\star - P(K)\big)^T \Pi \big(P^\star - P(K)\big) \\ \ge & f\big(P(K)\big) + \nabla\widetilde{f}(K)^T\big(P^\star - P(K)\big) \\ & + \Big(\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T - \omega(K)^T\Big)\big(P^\star - P(K)\big) \\ & + \min_{i=1,\ldots,N}\alpha_i\|\big(P^\star - P(K)\big)\|_2^2. \end{aligned} \tag{37}$$

Minimizing the RHS of (37) over $\big(P^\star - P(K)\big)$ gives

$$\begin{aligned} f(P^\star) \ge & f\big(P(K)\big) - \frac{2}{\min_{i=1,\ldots,N}\alpha_i}\|\nabla\widetilde{f}(K)\|_2^2 \\ & - \frac{2}{\min_{i=1,\ldots,N}\alpha_i}\|\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big) - \omega(K)\|_2^2. \end{aligned} \tag{38}$$

It follows that $f\big(P(K)\big) - f\big(P(K+1)\big) \ge \frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)\Big(f\big(P(K)\big) - f(P^\star)\Big) - \frac{1}{2}\lambda_2(W)\|\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big) - \omega(K)\|_2^2 + \big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T\zeta\mathcal{L}\nabla\widetilde{f}(K) - \omega(K)^T\zeta\mathcal{L}\nabla\widetilde{f}(K)$, leading to $f\big(P(K+1)\big) + \frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)f\big(P(K)\big) \le f\big(P(K)\big) + \frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)f(P^\star) + \frac{1}{2}\lambda_2(W)\|\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)\|_2^2 + \frac{1}{2}\lambda_2(W)\|\omega(K)\|_2^2 + \omega(K)^T\zeta\mathcal{L}\nabla\widetilde{f}(K) - \big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T\zeta\mathcal{L}\nabla\widetilde{f}(K)$. Adding $-\frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)\,f\big(P(K)\big) - f(P^\star)$ on both sides yields

$$\begin{aligned} & f\big(P(K+1)\big) - f(P^\star) \\ \le & \Big(1 - \frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)\Big)\Big(f\big(P(K)\big) - f(P^\star)\Big) \\ & + \frac{1}{2}\lambda_2(W)\|\omega(K)\|_2^2 + \frac{1}{2}\lambda_2(W)\,\Big\|\big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)\Big\|_2^2 \\ & + \omega(K)^T\zeta\mathcal{L}\omega(K) + \omega(K)^T\zeta\mathcal{L}\nabla f\big(\hat{P}_\tau(K)\big) \\ & - \big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T\zeta\mathcal{L}\nabla f\big(\hat{P}_\tau(K)\big) \\ & - \big(\widetilde{\varepsilon}(K)\delta_\tau(K)\big)^T\zeta\mathcal{L}\omega(K). \end{aligned} \tag{39}$$

By (11), (39) can be written as $f\big(P(K+1)\big) - f(P^\star) \le \Big(1 - \frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)\Big)\Big(f\big(P(K)\big) - f(P^\star)\Big) + \frac{1}{2}\lambda_2(W)\|\omega(K)\|_2^2 + \frac{1}{2}(1-k_f)^2\lambda_2(W)\,\Big\|\big(\widetilde{\varepsilon}(K)\delta_\tau(K-1)\big)\Big\|_2^2 + \zeta\lambda_N(\mathcal{L})\|\omega(K)\|_2^2 + \omega(K)^T\zeta\mathcal{L}\nabla f\big(\hat{P}_\tau(K)\big) - (1-k_f)\big(\widetilde{\varepsilon}(K)\delta_\tau(K-1)\big)^T\zeta\mathcal{L}\nabla f\big(\hat{P}_\tau(K)\big) - (1-k_f)\big(\widetilde{\varepsilon}(K)\delta_\tau(K-1)\big)^T\zeta\mathcal{L}\omega(K)$. Taking the expectation on both sides gives

$$\begin{aligned} & \mathbb{E}\Big(f\big(P(K+1)\big) - f(P^\star)\Big) \\ \le & \Big(1 - \frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)\Big)\mathbb{E}\Big(f\big(P(K)\big) - f(P^\star)\Big) \\ & + \Big(\frac{1}{2}\lambda_2(W) + \zeta\lambda_N(\mathcal{L})\Big)\mathbb{E}\big(\|\omega(K)\|_2^2\big) \\ & + \frac{1}{2}(1-k_f)^{2K}\lambda_2(W)\mathbb{E}\left(\|\widetilde{\varepsilon}(K)\delta_\tau(0)\|_2^2\right) \\ & - (1-k_f)^K\mathbb{E}\Big(\big(\widetilde{\varepsilon}(K)\delta_\tau(0)\big)^T\zeta\mathcal{L}\nabla f\big(\hat{P}(K)\big)\Big), \end{aligned} \tag{40}$$

where we vanished the terms $-(1-k_f)\big(\widetilde{\varepsilon}(K)\delta_\tau(K-1)\big)^T\zeta\mathcal{L}\omega(K)$ and $\omega(K)^T\zeta\mathcal{L}\nabla f\big(\hat{P}(K)\big)$, because $\omega(K)$ is independent of $\delta_\tau(K)$ and $\nabla f\big(\hat{P}(K)\big)$, and $\omega(K)$ has a zero mean.

For the second term in (40), using the fact that $\omega_i(K)$ is independently identically distributed (i.i.d.), $\omega_i(K) \sim \mathrm{Lap}(\varphi_i^K)$, $\mathbb{E}\big(\omega_i(K)\omega_j(K)\big) = \mathbb{E}\big(\omega_i(K)\big)\mathbb{E}\big(\omega_j(K)\big) = 0$, for $i \ne j$, and $\mathbb{E}\big(\omega_i(K)^2\big) = \mathrm{Var}\big(\omega_i(K)\big) = 2\varphi_i^{2K} \to 0$ as $K \to \infty$, it follows that $\mathbb{E}\big(\|\omega(K)\|_2^2\big) \to 0$ as $K \to \infty$. If $0 < k_f < 1$, we have $-(1-k_f)^K\mathbb{E}\Big(\big(\widetilde{\varepsilon}(K)\delta_\tau(0)\big)^T\zeta\mathcal{L}\nabla f\big(\hat{P}(K)\big)\Big) \to 0$ and $\frac{1}{2}(1-k_f)^{2K}\lambda_2(W)\,\mathbb{E}\,\|\widetilde{\varepsilon}(K)\delta_\tau(0)\|_2^2 \to 0$ as $K \to \infty$. Hence, as $K \to \infty$, (40) boils down to $\mathbb{E}\Big(f\big(P(K+1)\big) - f(P^\star)\Big) \le \big(1 - \frac{\min_{i=1,\ldots,N}\alpha_i}{4}\lambda_2(W)\big)\mathbb{E}\Big(f\big(P(K)\big) - f(P^\star)\Big)$. We have shown previously that $\lambda_2(W) < \frac{4}{\min_{i=1,\ldots,N}\alpha_i\lambda_2(\mathcal{L})}$, indicating that $\mathbb{E}\big(f\big(P(K)\big) - f(P^\star)\big) \to 0$ as $K \to \infty$.

Next, we show that the proposed scheme guarantees ϵ-differential privacy. For any K, let

$$\begin{cases} R'(K) = \{\omega(K) \in \Omega_K \mid \mathcal{I}^{out}_{P'(0)}\big(\omega(K)\big) \in S_K\}, \\ R''(K) = \{\omega(K) \in \Omega_K \mid \mathcal{I}^{out}_{P''(0)}\big(\omega(K)\big) \in S_K\}, \end{cases}$$

where Ω_K is the sample space up to time K, and S_K is the set by truncating the elements of S to the subsequence of length $K+1$. According to the continuity of probability (Durrett, 2010), we have $\mathbb{P}\{\omega \in \Omega \mid \mathcal{I}^{out}_{P'(0)}(\omega) \in S\} = \lim_{K\to\infty}\int_{R'(K)} F\big(\omega'(K)\big)d\omega'(K)$ and

$$\begin{aligned} & \mathbb{P}\{\omega \in \Omega \mid \mathcal{I}^{out}_{P''(0)}(\omega) \in S\} \\ = & \lim_{K\to\infty}\int_{R''(K)} F\big(\omega''(K)\big)d\omega''(K), \end{aligned} \tag{41}$$

where $F(\cdot)$ is the $N(K+1)$-dimensional joint Laplace probability density function given by $F\big(\omega(K)\big) = \prod_{i=1}^N\prod_{t=0}^K \mathrm{PDF}\big(\omega_i(t)\big)$. Without loss of generality, we assume that for $i_0 \in \mathcal{V}$, $P''_{i_0}(0) = P'_{i_0}(0) + \sigma$ and $P''_i(0) = P'_i(0)$ for all $i \ne i_0$. For $\omega'(K) \in R'(K)$, define

$$\omega''_i(K) = \begin{cases} \omega'_i(K) - (1-k_f)^K\kappa_i(K)\sigma & i = i_0, \\ \omega'_i(K) & i \ne i_0, \end{cases}$$

where $2\alpha_{i_0} \le \kappa_{i_0}(K) = \nabla^2 f_{i_0}(y_{i_0}(K)) \le \overline{M}_{i_0}$, and $y_{i_0}(K)$ is some value between $P'_{i_0}(K)$ and $\hat{P}'_{i_0}(K|K-\tau)$. It is straightforward to see that $\mathcal{I}^{out}_{P'(0)}\big(\omega'(K)\big) = \mathcal{I}^{out}_{P''(0)}\big(\omega''(K)\big)$, yielding $\omega''(K) \in R''(K)$. Therefore, there exists a unique $(\omega'(K), \triangle\omega_K)$ such that $\omega''(K) = \omega'(K) + \triangle\omega_K$, where $\triangle\omega_K = (\triangle\omega_{1,K}, \ldots, \triangle\omega_{N,K})^T$. It is clear that $\triangle\omega_K$ is fixed and independent of $\omega''(K)$. Hence, we can rewrite (41) as $\mathbb{P}\{\omega \in \Omega \mid \mathcal{I}^{out}_{P''(0)}(\omega) \in S\} = \lim_{K\to\infty}\int_{R'(K)} F\big(\omega'(K) + \triangle\omega_K\big)d\omega'(K)$. It follows that $\frac{F\left(\omega'(K)\right)}{F\left(\omega'(K)+\triangle\omega_K\right)} \le \exp\Big(\sum_{t=0}^K \frac{(1-k_f)^t\overline{M}_{i_0}\sigma}{\varphi_{i_0}^t}\Big)$. If $1 - k_f < \varphi_{i_0} < 1$, we have $\frac{F\left(\omega'(K)\right)}{F\left(\omega'(K)+\triangle\omega_K\right)} \le \exp\Big(\frac{\overline{M}_{i_0}\sigma\varphi_{i_0}}{\varphi_{i_0}-(1-k_f)}\Big)$, leading to

$$F\big(\omega'(K)\big) \le \exp\left(\frac{\overline{M}_{i_0}\sigma\varphi_{i_0}}{\varphi_{i_0} - (1-k_f)}\right)F\big(\omega'(K) + \triangle\omega_K\big). \tag{42}$$

Integrating both sides of (42) over $R(K)$ and letting $K \to \infty$, we have

$$\begin{aligned} & \mathbb{P}\{\omega \in \Omega \mid \mathcal{I}^{out}_{P'(0)}(\omega) \in S\} \\ \le & \exp\left(\frac{\overline{M}_{i_0}\sigma\varphi_{i_0}}{\varphi_{i_0} - (1-k_f)}\right)\mathbb{P}\{\omega \in \Omega \mid \mathcal{I}^{out}_{P''(0)}(\omega) \in S\}, \end{aligned} \tag{43}$$

which establishes the ϵ_{i_0}-differential privacy of agent i_0, where $\epsilon_{i_0} = \frac{\overline{M}_{i_0}\sigma\varphi_{i_0}}{\varphi_{i_0}-(1-k_f)}$. Since i_0 can be any agent, ϵ-differential privacy preservation is established.

Finally, we show the accuracy. Repeating (33) gives $P_i(K) = P_i(0)-\zeta\sum_{t=0}^{K-1}\sum_{j=1}^{N} l_{ij}\nabla f_j(\hat{P}_j(K|K-\tau_j))-\zeta\sum_{t=0}^{K-1}\sum_{j=1}^{N} l_{ij}w_j(K)$. Because the noise is independent over time and among agents, it follows that, for any time step $K \geq 0$, $\mathrm{Var}\big(P_i(K)\big) = 2\zeta^2\sum_{t=0}^{K-1}\sum_{j=1}^{N} l_{ij}^2\varphi_j^{2t}$. As $K \to \infty$, we have $\mathbb{E}\big(P_i(\infty)\big) = P_i^\star$ and $\mathrm{Var}\big(P_i(\infty)\big) = 2\zeta^2\sum_{j=1}^{N}\frac{l_{ij}^2}{1-\varphi_j^2}$. Let $\mathrm{Var}\big(P(\infty)\big) = \sum_{i=1}^{N}\mathrm{Var}\big(P_i(\infty)\big)$. It follows that $\mathrm{Var}\big(P(\infty)\big) = 2\zeta^2\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{l_{ij}^2}{1-\varphi_j^2}$. By Chebyshev's inequality, we have $\mathbb{P}\{\|P(\infty)-P^\star\|_2 \leq r\} = 1-\mathbb{P}\{\|P(\infty)-P^\star\|_2 > r\} \geq 1-\frac{\mathrm{Var}(P(\infty))}{r^2}$. Choosing $r = \frac{\sqrt{\mathrm{Var}(P(\infty))}}{\sqrt{p}} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sqrt{2\zeta^2 l_{ij}^2}}{\sqrt{p(1-\varphi_j^2)}}$ yields $\mathbb{P}\{\|P(\infty)-P^\star\|_2 \leq r\} \geq 1-p$. □

Theorem 2 has established some key convergence properties of Algorithm 1. In particular, Eq. (40) shows that if ζ satisfies (34) and $0 < k_f < 1$, then $\mathbb{E}\big(f\big(P(K)\big)\big)$ converges to $f(P^\star)$ as $K \to \infty$; Eq. (43) establishes the ϵ-differential privacy preservation property, which shows that a more dispersive noise distribution guarantees a higher privacy.

Remark 2. Theorem 2 incorporates the techniques of predictive control and injecting Laplace noise together to solve a distributed economic dispatch problem with heterogeneous time-delays under the requirement of preserving generators' privacy. Compared with traditional privacy preservation schemes in distributed optimization, due to the additional dynamics of the predictor, we have to handle a higher-dimensional system which is coupled with the optimization system through both time-delays and the added noise. This yields significant challenges in analyzing the system convergence, e.g., deriving the privacy preservation constant ϵ and showing the (p, r) accuracy, since we have to tackle three different kinds of control errors, i.e., the optimization error, the prediction error, and the error caused by the noise, simultaneously.

6. Conclusion

This paper has investigated a distributed predictive scheme for the ED problem with heterogeneous time-delays. We employed the θ-logarithmic barrier function to relax the box constraint of the ED problem. We developed a consensus-based optimization algorithm to solve the ED problem, where the supply demand constraint can be efficiently satisfied over the whole time horizon. A predictive scheme has been proposed to compensate for the effect of time-delays. We have provided the theoretical analysis on the convergence of the proposed algorithms. Among other things, a privacy-preserved predictive scheme has been proposed to meet the requirement of privacy preservation, for which we have carefully characterized its convergence, differential privacy properties, as well as accuracy.

References

Abedini, Reza, Pinto, Tiago, Morais, Hugo, & Vale, Zita (2013). Multi-agent approach for power system in a smart grid protection context. In *PowerTech (POWERTECH), 2013 IEEE grenoble* (pp. 1–6). IEEE.

Binetti, Giulio, Davoudi, Ali, Lewis, Frank L, Naso, David, & Turchiano, Biagio (2014). Distributed consensus-based economic dispatch with transmission losses. *IEEE Transactions on Power Systems*, *29*(4), 1711–1720.

Chen, Pohung, & Chang, Hongchan (1995). Large-scale economic dispatch by genetic algorithm. *IEEE Transactions on Power Systems*, *10*(4), 1919–1926.

Chen, Xiaozheng, & Chen, Fei (2017). A distributed predictive scheme for economic dispatch with heterogeneous time delays in smart grids. In *Industrial electronics society, IECON 2017-43rd annual conference of the IEEE* (pp. 5107–5112). IEEE.

Durrett, Rick (2010). *Probability: theory and examples*. Cambridge university press.

Fiacco, Anthony V., & McCormick, Garth P. (1990). *Nonlinear programming: sequential unconstrained minimization techniques, Vol. 4*. Siam.

Gu, Keqin, Chen, Jie, & Kharitonov, Vladimir L. (2003). *Stability of time-delay systems*. Springer Science & Business Media.

He, Xing, Ho, Daniel WC, Huang, Tingwen, Yu, Junzhi, Abu-Rub, Haitham, & Li, Chaojie (2018). Second-order continuous-time algorithms for economic power dispatch in smart grids. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *48*(9), 1482–1492.

Ji, Meng, Ferrari-Trecate, Giancarlo, Egerstedt, Magnus, & Buffa, Annalisa (2008). Containment control in mobile networks. *IEEE Transactions on Automatic Control*, *53*(8), 1972–1975.

Khalil, Hassan K. (1996). *Noninear systems, Vol. 2 (5)*. New Jersey: Prentice-Hall, 651–651.

Li, Huaqing, Lü, Qingguo, Liao, Xiaofeng, & Huang, Tingwen (2018). Accelerated convergence algorithm for distributed constrained optimization under time-varying general directed graphs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Li, Chaojie, Yu, Xinghuo, & Yu, Wenwu (2014). Optimal economic dispatch by fast distributed gradient. In *Control automation robotics & vision (ICARCV), 2014 13th international conference on* (pp. 571–576). IEEE.

Li, Chaojie, Yu, Xinghuo, Yu, Wenwu, Huang, Tingwen, & Liu, Zhiwei (2016). Distributed event-triggered scheme for economic dispatch in smart grids. *IEEE Transactions on Industrial Informatics*, *12*(5), 1775–1785.

Liu, Qingshan, Yang, Shaofu, & Hong, Yiguang (2017). Constrained consensus algorithms with fixed step size for distributed convex optimization over multi-agent networks. *IEEE Transactions on Automatic Control*.

Lü, Qingguo, Li, Huaqing, & Xia, Dawen (2017). Distributed optimization of first-order discrete-time multi-agent systems with event-triggered communication. *Neurocomputing*, *235*, 255–263.

Mandal, Avikarsha (2016). Privacy preserving consensus-based economic dispatch in smart grid systems. In *International conference on future network systems and security* (pp. 98–110). Springer.

Nagata, Takeshi, Ueda, Yuji, & Utatani, Masahiro (2012). A multi-agent approach to smart grid energy management. In *IPEC, 2012 conference on power & energy* (pp. 327–331). IEEE.

Nedić, Angelia, & Olshevsky, Alex (2015). Distributed optimization over time-varying directed graphs. *IEEE Transactions on Automatic Control*, *60*(3), 601–615.

Nozari, Erfan, Tallapragada, Pavankumar, & Cortés, Jorge (2017). Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. *Automatica*, *81*, 221–231.

Olfati-Saber, Reza, & Murray, Richard M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, *49*(9), 1520–1533.

Park, Jong Bae, Jeong, Yun Won, Shin, Joong Rin, & Lee, Kwang Y (2010). An improved particle swarm optimization for nonconvex economic dispatch problems. *IEEE Transactions on Power Systems*, *25*(1), 156–166.

Wen, Guanghui, Yu, Wenwu, Yu, Xinghuo, & Cao, Jinde (2016). Designing adaptive consensus-based scheme for economic dispatch of smart grid. In *Advanced computational intelligence (ICACI), 2016 eighth international conference on* (pp. 236–241). IEEE.

Wood, Allen J., & Wollenberg, Bruce F. (2012). *Power generation, operation, and control*. John Wiley & Sons.

Yan, Ye, Qian, Yi, Sharif, Hamid, & Tipper, David (2013). A survey on smart grid communication infrastructures: Motivations, requirements and challenges. *IEEE Communications Surveys & Tutorials*, *15*(1), 5–20.

Yang, Shiping, Tan, Sicong, & Xu, Jian Xin (2013). Consensus based approach for economic dispatch problem in a smart grid. *IEEE Transactions on Power Systems*, *28*(4), 4416–4426.

Yang, Tao, Wu, Di, Sun, Yannan, & Lian, Jianming (2015). Impacts of time delays on distributed algorithms for economic dispatch. In *Power & energy society general meeting, 2015 IEEE* (pp. 1–5). IEEE.

Yorino, Naoto, Hafiz, Habibuddin M, Sasaki, Yutaka, & Zoka, Yoshifumi (2012). High-speed real-time dynamic economic load dispatch. *IEEE Transactions on Power Systems*, *27*(2), 621–630.

Yu, Wenwu, Li, Chaojie, Yu, Xinghuo, Wen, Guanghui, & Lü, Jinhu (2015). Distributed consensus strategy for economic power dispatch in a smart grid. In *Control conference (ASCC), 2015 10th Asian* (pp. 1–6). IEEE.

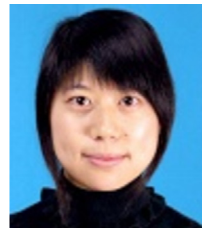
Zhang, Ziang, & Chow, Mo Yuen (2012). Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid. *IEEE Transactions on Power Systems*, *27*(4), 1761–1768.

Zhao, Chengcheng, Duan, Xiaoming, & Shi, Yang (2016). Analysis of consensus-based economic dispatch algorithm under uniform time delays. In *Decision and control (CDC), 2016 IEEE 55th conference on* (pp. 783–788). IEEE.

Zhao, Chengcheng, He, Jianping, Cheng, Peng, & Chen, Jiming (2017). Privacy-preserving consensus-based energy management in smart grid. In *Power & energy society general meeting, 2017 IEEE* (pp. 1–5). IEEE.

Fei Chen is presently a Professor with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China and School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao, China. Prior to joining Northeastern University, he was with Utah State University, Logan, Utah, from 2009 to 2010, and with Xiamen University, China from 2010 to 2018. His research interests include multiagent systems theory, nonsmooth control, cooperative control, distributed optimization, and complex networks. Dr. Chen is one author of the book "Distributed Average Tracking in Multi-agent Systems" (Springer-Verlag, 2020). He was a recipient of Natural Science Award of Hebei Province (First Class), Distinguished Young Scholars of Hebei Province, and Outstanding Reviewer for IEEE Transactions on Control of Network Systems. Dr. Chen is an Associate Editor for IEEE ACCESS and PLOS ONE. He is a Senior Member of IEEE.

Xiaozheng Chen received the B.S. degree in Automation from Fuzhou University, Fuzhou, China, in 2016 and the M.S. degree in Control Engineering from Xiamen University, Xiamen, China, in 2019. His research focuses on distributed economic dispatch for smart grids.



Linying Xiang is currently an Associate Professor with the School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao, China. She received the Ph.D. degree in Control Theory and Control Engineering from the Department of Automation at Nankai University, Tianjin, China in 2008. From November 2008 to October 2010, she was a Postdoctoral Research Fellow and from March 2013 to June 2013 she was a Senior Research Assistant, both with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. From July 2017 to July 2018, she was a Visiting Associate Researcher with the Department of Electrical and Computer Engineering, University of California, Riverside. She was an Assistant Professor (2010–2014) and an Associate Professor (2014–2018) with the Department of Automation at Xiamen University, Xiamen, China. Dr. Xiang's research interests are in the areas of synchronization, control and optimization of complex networks, and coordinated control of multiagent systems.

Wei Ren is currently a Professor with the Department of Electrical and Computer Engineering, University of California, Riverside. He received the Ph.D. degree in Electrical Engineering from Brigham Young University, Provo, UT, in 2004. Prior to joining UC Riverside in 2011, he was a faculty member at Utah State University and a postdoc at the University of Maryland. His research focuses on distributed control of multi-agent systems and autonomous control of unmanned vehicles. Dr. Ren was a recipient of the Antonio Ruberti Young Researcher Prize in 2017 and the National Science Foundation CAREER Award in 2008. He is currently an Associate Editor for Automatica and IEEE Transactions on Automatic Control. He is an IEEE Fellow and an IEEE Control Systems Society Distinguished Lecturer.