Distributed Optimal Frequency Control under Communication Packet Loss in Multi-Agent Electric Energy Systems

Siyu Xie [a], Masoud H. Nazari [b], Le Yi Wang [b], George Yin [c], Xinyu Zhang [a,d]

[a] *School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China*

[b] *Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA*

[c] *Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA*

[d] *School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China*

Abstract

This paper investigates the impact of communication packet loss on Distributed Optimal Frequency Control (DOFC) in Alternating Current (AC) electric energy systems, populated with multiple clusters of hybrid producer-consumer (prosumer) agents. The paper first establishes rigorous relationships between the communication packet delivery ratio and the convergence rate of the proposed DOFC algorithm. This provides a foundation for resource allocation on communication systems to enhance the convergence speed of distributed optimization and control algorithms, such as DOFC, under noisy and disrupted communication systems. The paper develops a systematic approach to identify the best possible convergence rate over all possible algorithms, by introducing an algorithm that can achieve asymptotically the Cramér-Rao (CR) lower bound. This fundamental result links the information contents of data to the best possible mean-square estimation error. Simulation studies on an electric energy system validate the theoretical results.

Key words: Communication uncertainty, convergence rate, distributed optimal frequency control, multi-agent electric energy systems, packet delivery ratio.

1 Introduction

Electric energy systems are moving towards a hybrid centralized-distributed architecture with a large penetration of distributed energy resources (DERs), such as distributed solar generation, energy storage, connected buildings, and electric vehicles and their supporting infrastructures. Several distributed algorithms have been proposed for solving power system problems at different time scales, such as DC Optimal Power Flow (OPF) (Kraning, Chu, Lavaei, & Boyd, 2014; Yi, Hong, & Liu, 2016; Persis, Weitenberg, & Dörfler, 2018), AC OPF (Dall'Anese, Zhu, & Giannakis, 2013; Zhang, Lam, Domínguez-García, & Tse, 2014), and the optimal frequency control problem (Nazari, Costello, Feizollahi, Grijalva, & Egerstedt, 2014; Chang & Zhang, 2016; Xi, Dubbeldam, Lin, & van Schuppen, 2018; Nazari, Wang, Grijalva, & Egerstedt, 2020; Zhang & Cortés, 2021). The state-of-the-art distributed algorithms proposed for smart power grids need to converge in the cyber network before the solutions can be implemented on the physical grid and the intermittent iterations are not satisfying power flow and other system constraints (Molzahn et al., 2017). Note that "intermediate iterations" imply the iterative steps before approximately reaching the optimal solution. In other words, this is an "transient period" in search for the optimal solution. On the other hand, fully decentralized methods lose optimality and can lead to inter-area oscillations among sub-systems (Nazari & Ilic, 2014).

Note that the impact of communications on networked

⋆ This research was supported in part by the National Science Foundation under Grants DMS-2229108 and DMS-2229109.

Email addresses: syxie@uestc.edu.cn (Siyu Xie), masoud.nazari@wayne.edu (Masoud H. Nazari), lywang@wayne.edu (Le Yi Wang), gyin@uconn.edu (George Yin), xyu_zhang@hotmail.com (Xinyu Zhang).

control systems has been widely studied in many perspectives, such as noisy communication channels (Li, Jin, & Yan, 2021; Huang, Dey, Nair, & Manton, 2010), time delays (Huang & Tian, 2018;Dong, Li, Nie, Song, & Yang, 2019), event-trigger-based strategies (Liuzza, Dimarogonas, Di Bernardo, & Johansson, 2016; Li, Tang, & Karimi, 2020), and so on. In Moreau, 2005, the communication systems were modeled as a time-varying network topology in terms of mobility and the impact on the network control quality was established. Moreover, an in-depth study of coordinated control and communication design was conducted in Xu, Wang, Yin, & Zhang, 2014a. Also, the authors in Xu, Wang, Yin, & Zhang, 2014b focused on block erasure channels. The authors in Richardson & Urbanke, 2008 established safety distances of modulation signals. Furthermore,Nguyen et al., 2018 studied the impact of communication packet delivery ratio on highway platoon performance.

In our previous work (Nazari, Xie, Wang, Yin, & Chen, 2021; Xie, Nazari, Wang, Yin, & Chen, 2021), we investigated the impact of communication packet loss and noisy environment on the performance of optimal load tracking and allocation (OLTA) in DC microgrids (MGs). This paper extends our earlier work to distributed optimal frequency control algorithms in prosumer-based AC electric energy systems. The main problem considered in this paper is related to gradient-based distributed optimization, which has been widely studied in the literature (Nedić & Ozdaglar, 2009;Duchi, Agarwal, & Wainwright, 2011; Wang et al., 2015; Yuan, Ling, & Yin, 2016; Wu et al., 2017; Tian, Sun, & Scutari, 2020). It is noted that in deterministic iterative algorithms without stochastic noises, linear convergence of several classical and more recent algorithms has been established, namely the error sequence $e_{k+1}/e_k \to \sigma < 1$. This implies that asymptotically, $e_k \leq c\sigma^k e_0$, achieving exponential convergence. This is fundamentally different from stochastic systems. The iteration algorithms (stochastic approximation algorithms) proposed in this paper involve stochastic noises. Consequently, it is impossible to achieve "linear convergence" or equivalently exponential convergence rates. Instead, the Fisher information dictates the best achievable rates, which are of certain polynomial orders. In other words, while the convergence analysis for deterministic systems is somehow related, it cannot be directly applied to prove the convergence properties of stochastic system analysis in this paper.

In data-based statistical analysis, information contained in observation data is used to estimate unknown parameters or seek unknown optimal solutions. The error variance is a measure of performance in this pursuit. The CR lower bound and Fisher information (Marzetta, 1993) provide the lower bound that the information content in data can be used in reducing the error variance, independent of searching algorithms. When an algorithm achieves this lower bound asymptotically, it becomes the optimal or best possible among all possible algorithms, implying that the information in data has been fully utilized. Our statements in this paper follow this convention in data-based science and statistical analysis.

In summary, this paper establishes a quantitative and fundamental relationship between communication packet loss ratio and the convergence rate of the DOFC algorithm in AC power grids, populated with multiple hybrid consumer-producer (prosumer) agents. The main contributions of the paper are as follows:

(1) Implementing stochastic network models to represent communication network dynamics in prosumer-based electric energy grids.
(2) Embedding packet delivery ratio and communication uncertainties into the DOFC algorithm, and laying a foundation for rigorous analysis of integrated communication and optimal control schemes.
(3) Quantitatively characterizing the fundamental relationship between packet delivery ratios and convergence rates of the DOFC algorithm to develop a practical criterion for securing reliability of optimal frequency control under communication uncertainties.
(4) Illustrating that the convergence rate of the DOFC algorithm can asymptotically achieve the CR lower bound. Note that the CR bound represents the lower bound on the variances of errors between the optimal control action and estimated control solution of the DOFC algorithm.

The rest of the paper is organized as follows. Section 2 gives an overview of distributed optimal frequency control problems in AC electric energy systems. Section 3 presents the global optimality conditions, develops distributed control algorithms with embedded communication uncertainty, and introduces stochastic models for erasure channels in communication systems. The main results are established in Section 4, where error bounds, strong convergence, and asymptotic optimality are derived. The technical findings are illustrated on two realistic electric energy systems in Sections 5 to show the impact of erasure channels on electric energy system reliability. The paper concludes with discussions of the overall findings in Section 6.

2 Overview of Distributed Optimal Frequency Control

DOFC involves bringing the system-wide frequency to 60 Hz or 50 Hz after a disturbance in an economically optimal way. When the power system is clustered into multiple prosumers, DOFC will be performed at the prosumer level. Prosumers can be as small as a microgrid or smart building, or as large as a utility grid sub-system.

The prosumer-based electric energy grid forms a multi-agent network, which can be represented by a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, \cdots, n\}$ is the set of all prosumers and $\mathcal{E}$ is the set of edges or energy system tie-lines. The presence of a tie-line (i, j) indicates that prosumer i has direct electrical connection to prosumer j. We assume that the network control topology follows the prosumer grid tie-line topology. The set of neighbors of prosumer i is denoted as $\mathcal{N}_i = \{\ell \in \mathcal{V} | (\ell, i) \in \mathcal{E} \text{ or } (i, \ell) \in \mathcal{E}\}$. In a distributed architecture, prosumer i shares information with its neighboring prosumers in $\mathcal{N}_i$ to achieve system-level performance, such as optimal frequency control.

The dynamic modeling of power systems in Nazari, Costello, Feizollahi, Grijalva, & Egerstedt, 2014; Nazari, Wang, Grijalva, & Egerstedt, 2020 leads to dynamic relationships among prosumers. Typical dynamic models of power systems involve complicated, nonlinear, and high-order dynamics. However, for certain power system control problems, such as frequency regulations, approximate lower order models can be derived by using the ideas of singular perturbations (Chow, Winkelman, Pai,& Sauer, 1990). For clarity and conciseness, we use the following first-order, discretized, linearized, pseudo-stationary dynamic relationships to demonstrate our methodologies and algorithms. Consider $p_i \in \mathbb{R}$, the real power deviation from the scheduled value, associated with each node $i \in \mathcal{V}$. After collecting all p_i in $\mathcal{V}$, we obtain the ensemble state given by $p = [p_1, p_2, \ldots, p_n]^\top \in \mathbb{R}^n$. By the standard discretization with a sampling interval T, the evolution of p at discrete times kT, $k = 0, 1, \ldots$, can be simplified as the discrete-time dynamical system,

$$p(k+1) = Ap(k) + Bu(k), \tag{1}$$

where $u = [u_1, u_2, \ldots, u_n]^\top \in \mathbb{R}^n$ is the vector of prosumers' frequency control variables, $A = [a_{i,j}] \in \mathbb{R}^{n\times n}$ and $B = [b_{i,j}] \in \mathbb{R}^{n\times n}$ are system matrix and control matrix, respectively. These matrices represent the underlying electrical topology of the prosumer electric energy grid $\mathcal{G}$, i.e., $a_{i,j} \neq 0, b_{i,j} \neq 0$ if $j = i$ or $j \in \mathcal{N}_i$, and $a_{i,j} = b_{i,j} = 0$ otherwise. The matrices A and B are full rank if the network topology is connected.

The real power deviations are directly correlated with frequency deviations through power flow equations and dynamic behavior of generators and loads (Siyu et al., 2022; Nazari, Costello, Feizollahi, Grijalva, & Egerstedt, 2014). DOFC ensures that frequency stability is achieved by minimizing overall control costs. This paper is concerned with a one-step predictive optimization in which $p(k)$ is given, and a linear quadratic performance of $u(k)$ and $p(k+1)$ is to be minimized. As a result, the problem is independent of k. By defining $p(k) = p, u(k) = u, p(k+1) = x$, the cost function can be written as

$$\begin{aligned}&\min_{u,x} \ \mathcal{J}(u,x) = \min_{u,x} \frac{1}{2}(x^\top Qx + u^\top Ru)\\ &\text{s.t.} \quad x = Ap + Bu,\end{aligned} \tag{2}$$

where $Q = \operatorname{diag}\{q_1, q_2, \ldots, q_n\} \in \mathbb{R}^{n\times n}$ and $R = \operatorname{diag}\{r_1, r_2, \ldots, r_n\} \in \mathbb{R}^{n\times n}$ are positive definite diagonal cost coefficient matrices. In the next section, we will propose a gradient-based distributed algorithm to solve (2) and establish the convergence results under communication packet loss.

3 Distributed Algorithm and Communication Uncertainty

3.1 Distributed Gradient Algorithms with Embedded Communication Uncertainty

Theoretically, the global optimal solution of (2) with the equality constraint $x = Ap + Bu$ can be obtained by the Lagrange Multiplier method: For $\lambda \in \mathbb{R}^n$, $L(u, x, p, \lambda) = \frac{1}{2}(u^\top Ru + x^\top Qx) + \lambda^\top[x - (Ap + Bu)]$. Thus, the optimal solution is $u^* = -G^{-1}B^\top QAp$, where $G = R + B^\top QB \in \mathbb{R}^{n\times n}$. Since this solution involves the inverse of matrix $R + B^\top QB$, it requires global information and is not feasible in a distributed framework.

In order to obtain a distributed method to solve this optimization problem, we first define the following performance index:

$$\min_u \mathcal{J}(u) = \min_u \frac{1}{2}[u^\top Ru + (Ap + Bu)^\top Q(Ap + Bu)],$$

whose gradient is $\nabla_u \mathcal{J}(u) = Ru + B^\top Q(Bu + Ap) = Gu + B^\top QAp$. For each prosumer i ($i \in \{1, \ldots, n\}$), we denote

$$\mathcal{J}_i(u) = \frac{1}{2}\left\{ r_i u_i^2 + q_i \left[\sum_{j\in\mathcal{N}_i\cup\{i\}} (a_{i,j}p_j + b_{i,j}u_j) \right]^2 \right\}.$$

Then, we have $\mathcal{J}(u) = \sum_{i=1}^n \mathcal{J}_i(u)$. Note that

$$\nabla_u \mathcal{J}(u) = \begin{pmatrix} \nabla_{u_1}\mathcal{J}(u) \\ \vdots \\ \nabla_{u_n}\mathcal{J}(u) \end{pmatrix} \in \mathbb{R}^n, \tag{3}$$

where

$$\nabla_{u_i}\mathcal{J}(u) = \sum_{\ell=1}^n \nabla_{u_i}\mathcal{J}_\ell(u),$$

and

$$
\nabla_{u_i}\mathcal{J}_\ell(u) = \begin{cases} r_i u_i + q_i b_{i,i} \sum\limits_{j\in\mathcal{N}_i\cup\{i\}} (a_{i,j}p_j + b_{i,j}u_j), & \text{if } \ell = i, \\ q_\ell b_{\ell,i} \sum\limits_{j\in\mathcal{N}_\ell\cup\{\ell\}} (a_{\ell,j}p_j + b_{\ell,j}u_j), & \text{if } \ell \in \mathcal{N}_i, \\ 0, & \text{otherwise.} \end{cases}
$$

For each prosumer i at step t, we can adopt the following gradient algorithm without communication uncertainty to track the optimal solution:

$$
u_i^{t+1} = u_i^t - \mu^t \nabla_{u_i}\mathcal{J}(u^t) = u_i^t - \mu^t \sum_{\ell\in\mathcal{N}_i\cup\{i\}} \nabla_{u_i}\mathcal{J}_\ell(u^t),
$$

where the step size μ^t is designed to achieve convergence, and $u^t = [u_1^t, \ldots, u_n^t]^\top$, which is the t^{th} computed value for $u = u(k)$. This algorithm is strictly distributed since for each prosumer i, it only needs the gradient information from its neighbors ($\nabla_{u_i}\mathcal{J}_\ell(u^t)$ where $\ell \in \mathcal{N}_i \cup \{i\}$) to update the solution. Thus, the DOFC algorithm can be written in the vector form $u^{t+1} = u^t - \mu^t \nabla_u \mathcal{J}(u^t)$, i.e., $u^{t+1} = u^t - \mu^t(Gu^t + B^\top QAp)$, which relies on communication between prosumers.

In practical applications, packet loss and channel interruptions may cause the cyber link to be randomly disconnected. Thus, communication packet loss can pose limitations for the convergence rate of the DOFC algorithm. The packet loss can be represented by an indicator function

$$
\gamma_{i,\ell}^t = \begin{cases} 1, \text{ if } i = \ell, \\ 1, \text{ if } i \neq \ell \text{ and the link } (i,\ell) \text{ is connected at } t, \\ 0, \text{ otherwise,} \end{cases}
$$

which is a random variable. Denote $\gamma^t = \{\gamma_{i,\ell}^t\} \in \mathbb{R}^{n\times n}$, and $\beta_i^t = \prod_{\ell\in\mathcal{N}_i\cup\{i\}} \gamma_{i,\ell}^t$, $\beta^t = \text{diag}\{\beta_1^t, \beta_2^t, \ldots, \beta_n^t\} \in \mathbb{R}^{n\times n}$. Note that $\beta_i^t = 0$ means that at least one of the links connected to prosumer i drops the packet at time instant t. When $\gamma_{i,\ell_0}^t = 0$, prosumer i lost the information from prosumer ℓ_0, i.e., $\nabla_{u_i}\mathcal{J}_{\ell_0}(u^t)$ may not be used for calculating the gradient value. Since ℓ_0 is lost, $\sum_{\ell=1}^n \gamma_{i,\ell}^t \nabla_{u_i} J_\ell(u^t)$ is not the correct gradient information of prosumer i at step t, which cannot be used for updating u_i^t directly. In this case, we just keep u_i^t unchanged until we get the correct gradient value. This may waste some information when one line is lost. It would be an interesting problem to design more suitable partial information gradients to update u_i^t in a proper way. Due to the page limitation, we cannot have an in-depth discussion in this paper, but this will be the research direction of our next work.

Thus, for each prosumer i, the updating algorithm to find the optimal control strategy with packet loss becomes

$$
u_i^{t+1} = u_i^t - \mu^t \beta_i^t \sum_{\ell\in\mathcal{N}_i\cup\{i\}} \nabla_{u_i}\mathcal{J}_\ell(u^t). \tag{4}
$$

The algorithm with embedded communication uncertainty is shown in Algorithm 1. Note that $u^{t+1} = u^t$ will only happen when $\beta^t = 0$, and the algorithm should not terminate since the value does not update in this case.

Algorithm 1 DOFC Gradient Algorithm with Embedded Communication Uncertainty and Random Noise

(1) Initial condition: Select e^0 as the threshold error, given the initial value for u^0, and let $t = 0$.
(2) Update: From u^t at each step $t \geq 0$, the control law is updated by

$$
u^{t+1} = u^t - \mu^t \beta^t (Gu^t + B^\top QAp + d^t), \tag{5}
$$

where $d^t \in \mathbb{R}^n$ is the random gradient noise.
(3) Termination condition: If $u^{t+1} \neq u^t$ and $\|u^{t+1} - u^t\| \leq e^0$, end the loop. Otherwise, let $t = t+1$ and go to Step (2).

Note that u^* satisfies $\mu^t\beta^t G u^* + \mu^t\beta^t B^\top QAp = 0$. Define the optimality error $\widetilde{u}^t = u^t - u^*$, and by (5), we can show that

$$
\widetilde{u}^{t+1} = (I_n - \mu^t\beta^t G)\widetilde{u}^t - \mu^t\beta^t d^t. \tag{6}
$$

Then, we will analyze the optimization error $\widetilde{u}^t$.

3.2 Communication Uncertainty Modeling

Information exchange among agents relies on communication channels, thus the reality of the communication network and its reliability in delivering critical data packets is essential for the convergence of the DOFC algorithm. In general, data packet can be lost due to erasure of one or multiple bits within the packet during transmission (i Fabregas & Caire, 2006; Nguyen et al., 2018). A data block is generated and coded by the source, and then transmitted to the receiver during information transmission in a given time interval. The received codeword is subject to possible erasure of bits because of channel uncertainties. After decoding and error correction, the receiver either acknowledges receipt of the data, or indicates a packet loss. If packet loss occurs during the transmission, data re-sending is permitted only within the given decision time interval. Probability of successful packet delivery is defined as the *packet delivery ratio.*

In this paper, we assume that packet losses for all channels are mutually independent, and each channel's packet loss is an independent and identically distributed (i.i.d) sequence of random variables. By applying this assumption to all channels, the communication uncertainty can be modeled by randomly switching network topologies such that the probability for each topology is generated from individual link connection probabilities. As a result, all related matrices in the distributed gradient algorithm will be random.

By the above assumption, $\gamma_{i,\ell}^t$ is an independent random variable. Suppose $\gamma_{i,\ell}^t$ is stationary and its packet delivery probability is

$$Pr\{\gamma_{i,\ell}^t = 1\} = \begin{cases} 1, \text{ if } i = \ell, \\ Pr_{i,\ell}, \text{ if link } (i,\ell) \text{ is connected at } t, \\ 0, \text{ otherwise,} \end{cases}$$

Denote $Pr = [Pr_{i,\ell}] \in \mathbb{R}^{n\times n}$, and $\bar{q}_i = \prod_{\ell\in\mathcal{N}_i} Pr_{i,\ell}$, where $\bar{Q} = \mathrm{diag}\{\bar{q}_1, \bar{q}_2, \ldots, \bar{q}_n\} \in \mathbb{R}^{n\times n}$ has the same order as u. Note that $\bar{q}_i$ denotes the packet delivery ratio for prosumer i, which is the product of ratios of all the links connected to i.

3.3 Communication Packet Loss and Electric Energy System Reliability

Timing is critical for electric energy systems operating tasks, particularly the optimal solution to DOFC needs to be computed in less than a few seconds. The expected communication times for different operating tasks have very strict time constraints (Wang, Xu, & Khanna, 2011). However, the communication networks are not always able to meet the standard requirements in many practical situations. Decreasing the packet delivery ratio between prosumers can slow down the convergence rate of the DOFC algorithm, which will increase the risk of violating the reliability criteria since the algorithm will take more iterations, namely more time to converge to the optimal solution and the intermittent iterations are not satisfying system constraints. Note that communication systems' bandwidth allocation and transmission power are commonly used to ensure a required packet delivery ratio, and the main task of this paper is to quantify how the packet delivery ratio should be controlled to meet the grid reliability criteria.

4 Main Results

We make the following basic assumptions for the theoretical analysis of Algorithm 1.

Assumption 1 *(1) $\mathcal{G}$ (power graph) is connected.*

(2) The noise $\{d^t \in \mathbb{R}^{n\times 1}\}$ is a sequence of i.i.d. random variables such that $\mathbb{E}[d^t] = \mathbf{0}_n \in \mathbb{R}^n$ and $\mathbb{E}[d^t \cdot (d^t)^\top] = \Sigma_d \in \mathbb{R}^{n\times n}$, where $\mathbb{E}[\cdot]$ is the mathematical expectation operator and Σ_d is symmetric positive definite.

(3) $\{\beta^t \in \mathbb{R}^{n\times n}\}$ is a sequence of i.i.d. random variables such that $\mathbb{E}[\beta^t] = \bar{\beta} \in \mathbb{R}^{n\times n}$ and $\mathbb{E}[\beta^t \cdot (\beta^t)^\top] = \Sigma_\beta \in \mathbb{R}^{n\times n}$. Both $\bar{\beta}$ and Σ_β are positive definite.

Remark 1 *Since β^t satisfies Bernoulli distribution, we know that $\bar{\beta} = \Sigma_\beta = \bar{Q}$. Also, the matrix $\bar{\beta}G$ is positive definite under* Assumption 1.

4.1 Identifying Optimization Error Bounds

By the above assumption, we can obtain the following exponential convergence result for the optimization error mean of the DOFC algorithm. Note that for two real symmetric matrices $X \in \mathbb{R}^{n\times n}$ and $Y \in \mathbb{R}^{n\times n}$, $X \geq Y (X > Y, X \leq Y, X < Y)$ means that $X - Y$ is a semi-positive (positive, semi-negative, negative) definite matrix. Let $\lambda_{\min}\{\cdot\}$ and $\lambda_{\max}\{\cdot\}$ denote the smallest and largest eigenvalue of a matrix, respectively. For any deterministic matrix $X \in \mathbb{R}^{s\times t}$, the Euclidean norm is defined as $\|X\| = (\lambda_{\max}\{XX^\top\})^{\frac{1}{2}}$; and for any random matrix Y, its norm is defined as $\|Y\| = \{\mathbb{E}[\|Y\|^2]\}^{\frac{1}{2}}$. Here we first establish the mean convergence in the following result.

Theorem 1 *Suppose that $\mu^t = \mu$. Under Assumption 1, there exist constants $\mu^* > 0$ and $c > 0$ such that for any $\mu \in (0, \mu^*)$, $0 \leq 1 - \mu c < 1$ and*

$$\|\mathbb{E}[\widetilde{u}^t]\| \leq (1 - \mu c)^t \|\mathbb{E}[\widetilde{u}^0]\|.$$

PROOF. Since β^t and d^t are independent, and $\mathbb{E}[d^t] = \mathbf{0}_n$, then we have $\mathbb{E}[\widetilde{u}^{t+1}] = (I_n - \mu\bar{\beta}G)\mathbb{E}[\widetilde{u}^t]$. In addition, since $\bar{\beta}G$ is positive definite by Remark 1, there exists two constants $c_1 = \lambda_{\min}\{\bar{\beta}G\} > 0$ and $c_2 = \lambda_{\max}\{\bar{\beta}G\} > 0$ such that $c_1 I_n \leq \bar{\beta}G \leq c_2 I_n$. Thus, we have $(1 - \mu c_2)I_n \leq I_n - \mu\bar{\beta}G \leq (1 - \mu c_1)I_n$.

If the step size μ is selected to satisfy $1 - \mu c_1 < 1$ and $1 - \mu c_2 > -1$, then $\|(I_n - \mu\bar{\beta}G)^t\| \leq (1 - \mu c)^t$. Thus, we can choose $\mu^* = 2/c_2 > 0$ and c such that $1 - \mu c = \max\{|1 - \mu c_1|, |1 - \mu c_2|\} \in [0, 1)$, which implies that $\|\mathbb{E}[\widetilde{u}^t]\| \leq (1 - \mu c)^t \|\mathbb{E}[\widetilde{u}^0]\|$, which converges to 0 exponentially. This completes the proof.

Then we can obtain the following optimization error bound of the DOFC algorithm.

Theorem 2 *Suppose that $\mu^t = \mu$. Under Assumption 1, for all $t > 0$, there exists a constant $\mu^* > 0$, for any $\mu \in (0, \mu^*)$,*

$$\|\widetilde{u}^t\| \leq (1 - \mu c_1)^t \|\widetilde{u}^0\| + c_2\sqrt{\mu},$$

where $c_1 \in (0, 1)$ and $c_2 > 0$ are two constants.

PROOF. Note that for any $\mu \in (0, \mu^*)$ where μ^* is defined in Theorem 1, $0 \leq \mu\beta^t G < I_n$ holds. Since $\bar{\beta}G$ is positive definite, we know by Theorem 2.1 in Guo, 1994 that for any $t \geq s \geq 0$, $\| \prod_{j=s+1}^{t}(I_n - \mu\beta^j G)\| \leq (1 - \mu c_1)^{t-s}$, where $c_1 = 1-(1-\mu^*\lambda_{\min}\{\bar{\beta}G\})^{1/64\mu^*} \in (0,1)$. Using (6), and Assumption 1, by Guo & Ljung, 1995, we can obtain that $\|\widetilde{u}^{t+1}\| \leq (1-\mu c_1)^{t+1}\|\widetilde{u}^0\| + c_2\sqrt{\mu}$, where $c_2 > 0$ is a constant. This completes the proof.

Denote the mean-square error of $\widetilde{u}^t$ as $\Sigma_{\tilde{u}}^t = \mathbb{E}[\widetilde{u}^t \cdot (\widetilde{u}^t)^\top]$, and $\Sigma_{\tilde{u}} = \lim_{t\to\infty} \Sigma_{\tilde{u}}^t$, when the limits exist.

Theorem 3 *Under Assumption 1, for all $t \geq 1$, there exists a constant $\mu^* > 0$ such that for any $\mu \in (0, \mu^*)$,*

(1) The error variance

$$\Sigma_{\tilde{u}}^{t+1} =(I - \mu\bar{\beta}G)^{t+1}\Sigma_{\tilde{u}}^0(I - \mu\bar{\beta}G)^{t+1} + \mu^2 \sum_{\ell=0}^{t}(I - \mu\bar{\beta}G)^\ell \Sigma_\beta \Sigma_d (I - \mu\bar{\beta}G)^\ell. \quad (7)$$

(2) $\Sigma_{\tilde{u}}$ is the solution to the Lyapunov equation $(I - \mu\bar{\beta}G)\Sigma_{\tilde{u}}(I - \mu\bar{\beta}G) - \Sigma_{\tilde{u}} = -\mu^2\Sigma_\beta\Sigma_d$, or explicitly

$$\Sigma_{\tilde{u}} = \mu^2 \sum_{\ell=0}^{\infty}(I_n - \mu\bar{\beta}G)^\ell \Sigma_\beta \Sigma_d (I_n - \mu\bar{\beta}G)^\ell. \quad (8)$$

PROOF. 1) By (6) and Assumption 1, because $\Sigma_{\tilde{u}}^{t+1} = \mathbb{E}[\widetilde{u}^{t+1} \cdot (\widetilde{u}^{t+1})^\top]$, we have

$$\begin{aligned}\Sigma_{\tilde{u}}^{t+1} =&(I_n - \mu\bar{\beta}G)\Sigma_{\tilde{u}}^t(I_n - \mu\bar{\beta}G) + \mu^2\mathbb{E}[\beta^t\Sigma_d(\beta^t)^\top] \\ =&(I_n - \mu\bar{\beta}G)^{t+1}\Sigma_{\tilde{u}}^0(I_n - \mu\bar{\beta}G)^{t+1} \\ &+ \mu^2 \sum_{\ell=0}^{t}(I_n - \mu\bar{\beta}G)^\ell \Sigma_\beta \Sigma_d (I_n - \mu\bar{\beta}G)^\ell.\end{aligned}$$

2) By (7) and Assumption 1, $I - \mu\bar{\beta}G$ is stable, and $\Sigma_{\tilde{u}} = \lim_{t\to\infty} \Sigma_{\tilde{u}}^t = \mu^2 \sum_{\ell=0}^{\infty}(I_n - \mu\bar{\beta}G)^\ell \Sigma_\beta \Sigma_d (I_n - \mu\bar{\beta}G)^\ell$. Then

$$\begin{aligned}&(I - \mu\bar{\beta}G)\Sigma_{\tilde{u}}(I_n - \mu\bar{\beta}G) \\ =&\mu^2 \sum_{\ell=1}^{\infty}(I_n - \mu\bar{\beta}G)^\ell \Sigma_\beta \Sigma_d (I_n - \mu\bar{\beta}G)^\ell \\ =&\Sigma_{\tilde{u}} - \mu^2\Sigma_\beta\Sigma_d.\end{aligned}$$

The desired result thus follows.

Remark 2 *Theorem 3 establishes the mean square error on the optimal solution of the DOFC algorithm, which represents the fundamental impact of packet delivery ratio and step size selection on obtaining the optimal solution. Since for any $\mu \in (0, \mu^*)$, $\|I - \mu\bar{\beta}G\| < 1$ and $0 \leq \mu\beta^t G < I_n$ holds, the summation in (8) with ℓ from 0 to infinity is convergent to a finite limit, which is the solution to the corresponding continuous-time Lyapunov equation*[1] .

4.2 Convergence Results of Gradient-based Distributed Optimal Frequency Control

To achieve strong convergence of the DOFC algorithm, the step size should satisfy the following condition.

Assumption 2 *The step size satisfies the following properties: $\mu^t \geq 0$, $\mu^t \to 0$ as $t \to \infty$, and $\sum_t \mu^t = \infty$.*

The limit ODE (ordinary differential equation) of (5) is $\dot{u} = -\bar{\beta}Gu - \bar{\beta}B^\top QAp$, whose equilibrium point is precisely the optimal solution $u^* = -G^{-1}B^\top QAp$. The "limit ODE" method is a well established and comprehensive methodology for studying convergence properties of a large class of iterative stochastic approximation algorithms. It relates the algorithms' convergence to the stability of a related ODE in continuous time. Our algorithms are stochastic approximation. As a result, their convergence analysis can benefit from the limit ODE method. Using the ODE method in stochastic approximation (Kushner & Yin, 2003), the following strong convergence result can be obtained for the DOFC algorithm.

Theorem 4 *Under Assumptions 1 and 2, the control action $\{u^t\}$ generated by (5) converges to the optimal solution $u^t \to u^*$ with probability one (w.p.1) as $t \to \infty$.*

For simplicity, we omit the detailed proof and refer the reader to Chapters 5 and 6 of Kushner & Yin, 2003. While the actual proof of Theorem 4 will be skipped, the main ideas can be summarized as follows. Define $\xi^t = \sum_{j=0}^{t-1}\mu^j$, $\varpi(\xi) = \max\{t : \xi^t \leq \xi\}$, the piecewise constant interpolation $u_0(\xi) = u^t$ for $\xi \in [\xi^t, \xi^{t+1})$, and the shift sequence $u_t(\xi) = u_0(\xi + \xi^t)$. Under Assumption 2, the interpolated sequence $\{u_t(\cdot)\}$ is uniformly bounded and equicontinuous. By Ascoli-Arzéla's theorem (see Rudin, 1976), we can extract a subsequence $\{u_{t_\ell}(\cdot)\}$, which converges to $u(\cdot)$ on any compact intervals w.p.1 such that $u(\cdot)$ is a solution. The ODE has a unique equilibrium point, which is the optimal solution. Now, by using the Lyapunov method, the equilibrium point u^* is an asymptotically stable point, since $-\bar{\beta}G$ is stable by Remark 1. This theoretical result leads to the desired property for Algorithm 1.

[1] If $S = \sum_{k=0}^{\infty} H^k\Sigma H^k$, where H is symmetric and a contraction, then $HSH - S = -\Sigma$ is the Lyapunov equation. This is a set of linear equations and can be solved exactly.

Next, we establish the convergence rate of Algorithm 1, which is the property of how fast the intermittent DOFC iterations, u^t move toward the optimal solution. Note that the scaling factor $\sqrt{\mu^t}$ together with the asymptotic covariance gives the desired rate of convergence. The standard central limit theorem argument yields that $\frac{1}{\sqrt{t}}\sum_{j=k}^{k+t-1}\beta^j d^j$ converges weakly to $N(\mathbf{0}_n,\Sigma)$, where $N(\mathbf{0}_n,\Sigma)$ is a normal random variable whose mean is $\mathbf{0}_n$ and the variance is given by

$$\Sigma = \mathbb{E}[\beta^1 d^1 (d^1)^\top (\beta^1)^\top] = \mathbb{E}[\beta^1 \Sigma_d (\beta^1)^\top] \in \mathbb{R}^{n\times n}. \quad (9)$$

Next, we will show how the packet delivery ratio will affect convergence rate of the DOFC algorithm. In fact, there is a fundamental lower bound on the achievable convergence rate of the DOFC algorithm, called the CR bound (Anderson, 1984). Any algorithm that can asymptotically achieve the CR bound is the "fastest" algorithm. To achieve the CR bound, we derive an improved algorithm that provides the "optimal" convergence rate in the sense of the CR bound. This is done by using iterate averaging, i.e., (11). The idea of iterate averaging has a long history dated back in Chung, 1954, then in the 1970th by Polyak, in the 1990th by Polyak again, then Kushner & Yin, 2003, etc. By using a large step size, it approaches the design parameter faster initially. By using the averaging, "smaller" variance (covariance) is obtained.

For simplicity, we take the step size $\mu^t = 1/t^\gamma$, where $1/2 < \gamma < 1$. Then,

$$u^{t+1} = u^t - \frac{1}{t^\gamma}\beta^t (Gu^t + B^\top QAp + d^t), \quad (10)$$

and

$$\bar{u}^t = \frac{\sum_{j=0}^{t-1} u^j}{t} = \frac{(t-1)\bar{u}^{t-1} + u^{t-1}}{t}. \quad (11)$$

Thus, we can obtain the following result.

Theorem 5 $\sqrt{t}(\bar{u}^t - u^*)$ *converges weakly to a normal random variable with mean* $\mathbf{0}_n$ *and asymptotic covariance*

$$\Sigma^* = (\bar{\beta}G)^{-1}\Sigma(G\bar{\beta})^{-1} \in \mathbb{R}^{n\times n},$$

where Σ *is defined by (9).*

Remark 3 *For a coverage on the CR lower bound, we refer the reader to Rice, 2007 (pp. 300-302). Since the proof of* Theorem 5 *is similar to Chapter 11 of Kushner & Yin, 2003, it is omitted here. Note that* $\bar{u}^t - u^*$ *is asymptotically normal (Gaussian distributed) with zero mean and covariance* Σ^*/t. *For (10),* $\bar{u}^t$ *will converge to its limit at a convergence rate that approaches asymptotically the corresponding CR lower bound (Yin, 1991; Polyak & Juditsky, 1992). In this sense,* Σ^* *is the "smallest" covariance possible, which is a common error measure and a main performance indicator for the convergence rate. Hence,* Σ^* *is the measure of reliability of the DOFC algorithm against communication uncertainties. The focus of this paper is to use the CR lower bound and asymptotic optimality in the convergence rate of the algorithms to study certain fundamental impact of communication packet delivery rate.*

Remark 4 *Here we use the error covariance matrix* Σ^* *to evaluate how fast convergence to the optimal solution can be achieved and how the convergence rate depends on the packet delivery ratio. Note that*

$$\begin{aligned}\det(\Sigma^*) &= \det((\bar{\beta}G)^{-1}\Sigma(G\bar{\beta})^{-1})\\ &= \det((\bar{\beta}G)^{-1}\mathbb{E}[\beta^1\Sigma_d(\beta^1)^\top](G\bar{\beta})^{-1})\\ &\quad \det(\bar{\beta}^{-1}G^{-1}\bar{\beta}\Sigma_d G^{-1}\bar{\beta}^{-1})\\ &= \det(\bar{\beta}^{-1})\det(G^{-1}\Sigma_d G^{-1})\\ &= \left(\prod_{i=1}^{n}\prod_{\ell\in\mathcal{N}_i} Pr_{i,\ell}^{-1}\right)\det(G^{-1}\Sigma_d G^{-1}), \end{aligned} \quad (12)$$

where $\det(\cdot)$ *denotes the determinant operator. It can be observed that the lower the packet delivery ratio* $Pr_{i,\ell}$ *is, the "larger"* Σ^* *is and in turn the DOFC algorithm weakly converges slower to the optimal solution. This relationship will become a foundation for resource allocation (on communication systems) and reliability assessment (on power systems).*

5 Case Studies and Discussions

In this part, we use realistic prosumer electric energy systems to illustrate the relationship between the packet delivery ratio and the convergence rate of the DOFC algorithm. The IEEE 24-bus system represents a bulk electric energy system, which has 38 power lines and 32 generators. The average demand of the system is 2577 MW. The detailed description of the physical layer can be found in Grigg et al., 1999. In this study, the system is clustered into 10 prosumers with 25 tie-line connections between prosumers. Each prosumer represents a balancing authority area for frequency control. Fig. 1 illustrates the cyber layer of the prosumer-based IEEE 24-bus system. Note that the cyber-layer follows the sparsity structure of the prosumer-based grid to allow peer-to-peer communication among prosumers.

For simulation settings, we assume that the gradient noises are i.i.d Gaussian random variables with zero mean and variance 1. For a chosen packet delivery ratio, we repeat the simulation for $m = 500$ times with the same initial states for Algorithm 1. Then we can get m sequences $\{\|u^{t,j} - u^*\|,\ \ t \geq 1\}, j = 1,\ldots,m$, where $u^* = [u_1^*,\ldots,u_{10}^*] \in \mathbb{R}^{10}$ is the optimal solution

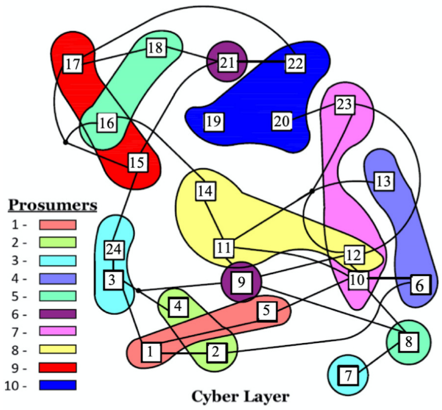

Fig. 1. Schematics of the cyber network of the IEEE 24 bus system.

and the superscript j denotes the j-th simulation result. We calculate the following mean error norm trajectories: $Mu^t = \frac{1}{m}\sum_{j=1}^{m}\|u^{t,j} - u^*\|, t \geq 1$, and plot $\log(Mu^t)$ under different packet delivery ratios in Fig. 2, which shows that by decreasing the packet delivery ratio, the convergence rate of the DOFC algorithm for the IEEE 24-bus system decreases.

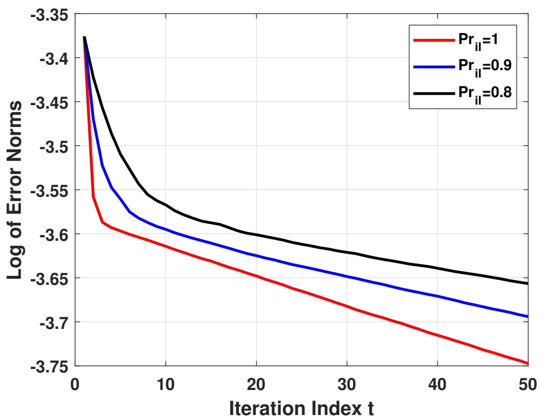

Fig. 2. Log of Error norms with different packet delivery ratios for IEEE 24 bus system.

Also, Fig. 3 shows the frequency dynamics of the center of inertia for different packet delivery ratios. Note that when the convergence rate increases, the DOFC algorithm may take less steps to obtain the optimal solution, which also implies that it takes less time to recover from a disturbance, i.e., a faster convergence rate ensures greater reliability of the prosumer energy gird. It takes 340 iterations to achieve 0.005 per unit (p.u.) error threshold for $P_{r_{i,\ell}} = 1$, and 529 iterations for $P_{r_{i,\ell}} = 0.9$, 843 iterations for $P_{r_{i,\ell}} = 0.8$. Thus, improving packet delivery ratio can directly increase the power system reliability. It implies a desirable approach of control-communication co-design in which distributed optimization algorithms and communication resource allocation must be coordinated.

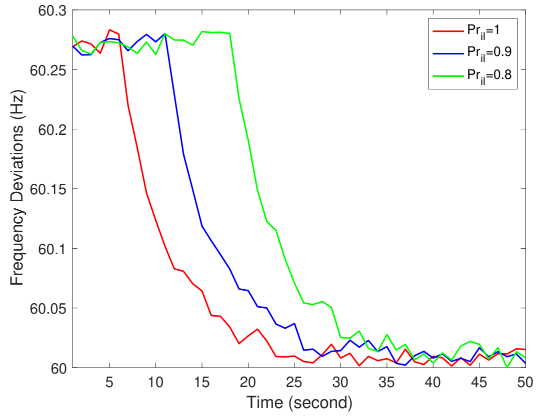

Fig. 3. Frequency dynamics of the center of inertia with different packet delivery ratios for IEEE 24 bus system.

Moreover, let the packet delivery ratio be 0.9, then we know that $\mathrm{Tr}(\Sigma^*) = 0.8946$. We apply the iterate averaging algorithm (10), (11) and repeat the simulation for $m = 500$ times with the same initial states. We can calculate the following mean covariance matrix: $Cu^t = \frac{1}{m}\sum_{j=1}^{m}(\bar{u}^{t,j} - u^*)(\bar{u}^{t,j} - u^*)^\top \in \mathbb{R}^{10\times 10}$. Then, we plot the trace of the mean covariance: $covu^t = \mathrm{Tr}(Cu^t), t \geq 1$, in Fig. 4, which approaches $\mathrm{Tr}(\Sigma^*)/t$ as t increases.

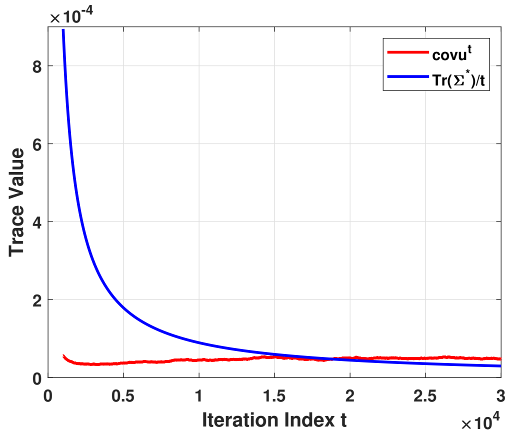

Fig. 4. Trace of the covariance matrix for IEEE 24 bus system.

6 Concluding Remarks

This paper developed rigorous analysis of the impact of data packet loss on the convergence of DOFC in AC electric energy systems, populated with many prosumer agents. The paper first modeled the communication uncertainties using probabilistic descriptions of packet delivery ratio. The proposed channel model was congregated to derive a stochastic description of the cyber network topology in AC electric energy systems. Stochastic approximation methods were used for the convergence analysis of the DOFC algorithm in prosumer-based energy systems. By embedding the information network switching model into the DOFC algorithm and performing stochastic analysis, we established a rigorous relationship between the packet delivery ratio of communication erasure channels, algorithm convergence rate, and accuracy of DOFC solutions. The theoretical findings provided guidance on the choice of control parameters for guaranteed convergence of the DOFC algorithm. The future research endeavor is to develop fault-tolerant distributed algorithms to mitigate the impact of communication failures, and to design more suitable partial information gradients to update u in a proper way.

References

Anderson, T. (1984). *An introduction to multivariate statistical analysis.* John Wiley & sons, New York, USA.

Chang, C. Y., & Zhang, W. (2016). Distributed control of inverter-based lossy microgrids for power sharing and frequency regulation under voltage constraints. *Automatica*, 66, 85–95.

Chow, J. H., Winkelman, J.R., Pai, M. A., & Sauer, P.W. (1990). Singular perturbation analysis of large-scale power systems. *International Journal of Electrical Power & Energy Systems*, 12(2), 117–126.

Chung, K. L. (1954). On a stochastic approximation method. *The Annals of Mathematical Statistics*, 25(3): 463-483.

Dall'Anese, E., Zhu, H., & Giannakis, G. B. (2013). Distributed optimal power flow for smart microgrids. *IEEE Transactions on Smart Grid*, 4(3), 1464–1475.

Dong, M., Li, L., Nie, Y., Song, D., & Yang, J. (2019). Stability analysis of a novel distributed secondary control considering communication delay in DC microgrids. *IEEE Transactions on Smart Grid*, 10(6), 6690–6700.

Duchi, J. C., Agarwal, A., & Wainwright, M. J. (2011). Dual averaging for distributed optimization: convergence analysis and network scaling. *IEEE Transactions on Automatic control*, 57(3), 592–606.

Grigg, C., Wong, P., Albrecht, P., Allan, R., Bhavaraju, M., Billinton, R., Chen, Q., Fong, C., Haddad, S., Kuruganty, S., & Li, W. (1999). The IEEE reliability test system-1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee. *IEEE Transactions on Power Systems*, 14(3), 1010–1020.

Guo, L. (1994). Stability of recursive stochastic tracking al gorithms. *SIAM Journal on Control and Optimization*, 32(5), 1195–1225.

Guo, L., & Ljung, L. (1995). Performance analysis of general tracking algorithms. *IEEE Transactions on Automatic Control*, 40(8), 1388– 1402.

Huang, M., Dey, S., Nair, G. N., & Manton, J. H. (2010). Stochastic consensus over noisy networks with Markovian and arbitrary switches. *Automatica*, 46(10), 1571–1583.

Huang, X. Z., & Tian, Y. P. (2018). Asynchronous distributed localization in networks with communication delays and packet losses. *Automatica*, 96, 134-140.

i Fabregas, A. G., & Caire, G. (2006). Coded modulation in the block-fading channel: Coding theorems and code construction. *IEEE Transactions on Information Theory*, 52(1), 91–114.

IEC 61850-5. (2003). Communication networks and systems in substations – Part 5: Communication requirements for functions and device models.

Kraning, M., Chu, E., Lavaei, J., & Boyd, S. P. (2014). *Dynamic network energy management via proximal message passing.* Now Publishers.

Kushner, H., & Yin, G. (2003). *Stochastic approximation and recursive algorithms and applications.* 2nd ed. New York, NY, USA: Springer-Verlag.

Li, H., Jin, B., & Yan, W. (2021). Distributed model predictive control for linear systems under communication noise: Algorithm, theory and implementation. *Automatica*, 125, 109422.

Li, X., Tang, Y., & Karimi, H. R. (2020). Consensus of multi-agent systems via fully distributed event-triggered control. *Automatica*, 116, 108898.

Liuzza, D., Dimarogonas, D. V., Di Bernardo, M., & Johansson, K. H. (2016). Distributed model based event-triggered control for synchronization of multi-agent systems. *Automatica*, 73, 1–7.

Marzetta, T. (1993). A simple derivation of the constrained multiple parameter cramer-rao bound. *IEEE Transactions on Signal Processing*, 41(6), 2247–2249.

Molzahn, D. K., Dörfler, F., Sandberg, H., Low, S. H., Chakrabarti, S., Baldick, R., & Lavaei, J. (2017). A survey of distributed optimization and control algorithms for electric power systems. *IEEE Transactions on Smart Grid*, 8(6), 2941–2962.

Moreau, L., (2005). Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(2), 169–182.

Nazari, M. H. (2013). *Small-Signal Stability Analysis of Electric Power Systems on the Azores Archipelago.* in Chapter 17 of Engineering IT-Enabled Sustainable Electricity Services. Springer, 2013.

S. Xie, M. H. Nazari, F. Nezampasandarbabi and L. Y. Wang, *Leveraging Deep Learning to Improve Performance of Distributed Optimal Frequency Control under Communication Failures.* in IEEE Transactions on Smart Grid, 2022, doi: 10.1109/TSG.2022.3194131.

Nazari, M. H., Costello, Z., Feizollahi, M. J., Grijalva, S., & Egerstedt, M. (2014). Distributed frequency control of prosumer-based electric energy systems. *IEEE Transactions on Power Systems*, 29(6), 2934–2942.

Nazari, M. H., & Ilic, M. (2014). Dynamic modelling and control of distribution energy systems: comparison with transmission power systems. *IET Generation, Transmission and Distribution*, 8(1), 26–34.

Nazari, M. H., Wang, L. Y., Grijalva, S., & Egerstedt, M. (2020). Communication-failure-resilient distributed frequency control in smart grids: part i: architecture and distributed algorithms. *IEEE Transactions on Power Systems*, 35(2), 1317–1326.

Nazari, M. H., Xie, S. Y., Wang, L. Y, Yin, G., & Chen, W. (2021). Impact of communication packet delivery ratio on reliability of optimal load tracking and allocation in DC microgrids. *IEEE Transactions on Smart Grid*, 12(4), 2812–2821.

Nedić, A., & Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1), 48–61.

Nguyen, T., Wang, L. Y., Yin, G., Zhang, H., Li, S. E., & Li, K. (2018). Impact of communication erasure channels on control performance of connected and automated vehicles. *IEEE Transactions on Vehicular Technology*, 67(1), 29–43.

Persis, D. C., Weitenberg, E. R., & Dörfler, F. (2018). A power consensus algorithm for DC microgrids. *Automatica*, 89, 364–375.

Polyak, B. T., & Juditsky, A. B. (1992). Acceleration of stochastic approximation by averaging. *SIAM Journal on Control and Optimization*, 30(4), 838–855.

Popli, N, & Ilic, M. (2013). *Modeling and Control Framework to Ensure Intra- dispatch Regulation Reserves, in Chapter 14 of Engineering IT-Enabled Sustainable Electricity Services.* Springer.

Richardson, T., & Urbanke, R. (2008). *Modern coding theory.* Cambridge university press.

Rudin, W. (1976). *Principles of Mathematical Analysis.* McGraw-hill New York.

Tian, Y., Sun, Y., & Scutari, G. (2020). Achieving linear convergence in distributed asynchronous multiagent optimization. *IEEE Transactions on Automatic Control*, 65(12), 5264–5279.

Wang, H., Liao, X., Huang, T., & Li, C. (2015). Cooperative distributed optimization in multiagent networks with delays. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(2), 363–369.

Wang, W., Xu, Y., & Khanna, M. (2011). A survey on the communication architectures in smart grid. Computer networks, 55(15), 3604–3629.

Wu, T., Yuan, K., Ling, Q., Yin, W., & Sayed, A. H. (2017). Decentralized consensus optimization with asynchrony and delays. *IEEE Transactions on Signal and Information Processing over Networks*, 4(2), 293–307.

Xi, K., Dubbeldam, J. L., Lin, H. X., & van Schuppen, J. H. (2018). Power-imbalance allocation control of power systems-secondary frequency control. *Automatica*, 92, 72–85.

Xie, S. Y., Nazari, M. H., Wang, L. Y., Yin, G., & Chen, W. (2021). Impact of stochastic generation/load variations on distributed optimal energy management in DC microgrids for transportation electrification. *IEEE Transactions on Intelligent Transportation Systems*, 1-10.

Xu, L., Wang, L. Y., Yin, G., & Zhang, H. (2014). Communication information structures and contents for enhanced safety of highway vehicle platoons. *IEEE Transactions on Vehicular Technology*, 63(9), 4206–4220.

Xu, L., Wang, L. Y., Yin, G., & Zhang, H. (2014). Impact of communication erasure channels on the safety of highway vehicle platoons. *IEEE Transactions on Intelligent Transportation Systems*, 16(3), 1456–1468.

Yi, P., Hong, Y. G., & Liu, F. (2016). Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and application to economic dispatch of power systems. *Automatica*, 74, 259–269.

Mathematical Statistics and Data Analysis, 3rd Ed., J.R. Rice, Brooks/Cole, 2007

Yin, G. (1991). On extensions of polyak's averaging approach t o stochastic approximation. *Stochastics: An International Journal of Probability and Stochastic Processes*, 36(3-4), 245–264.

Yuan, K., Ling, Q., & Yin, W. (2016). On the convergence of decentralized gradient descent. *SIAM Journal on Optimization*, 26(3), 1835–1854.

Zhang, Y., & Cortés, J. (2021). Model predictive control for transient frequency regulation of power networks. *Automatica*, 123, 109335.

Zhang, B., Lam, A.Y., Domínguez-García, A. D. and Tse, D. (2014). An optimal and distributed method for voltage regulation in power distribution systems. *IEEE Transactions on Power Systems*, 30(4), 1714–1726.