Impact of Stochastic Generation/Load Variations on Distributed Optimal Energy Management in DC Microgrids for Transportation Electrification

Siyu Xie, Masoud H. Nazari, *Senior Member, IEEE*, Le Yi Wang, *Life Fellow, IEEE*, George Yin, *Life Fellow, IEEE*, and Wen Chen, *Senior Member, IEEE*

***Abstract*— This paper studies the impact of stochastic load variations on distributed optimal load tracking and allocation (OLTA) problems in cyber-physical DC microgrids (MGs) for transportation electrification. Without load variations, the distributed optimization strategies developed in our earlier work can achieve convergence to global optimal solutions in a multi-objective optimization that balances fair load allocation and power loss reduction. Under persistent stochastic load variations, this paper develops distributed optimal strategies to track time-varying loads under noisy observations and establishes their convergence properties and error bounds. The limiting behavior of the errors characterizes the fundamental impact of the step size on irreducible errors due to conflict between attenuating observation noises and tracking load changes. Optimality conditions and algorithms for selecting the optimal step size are introduced to guide step size selection in practical applications. Simulation studies on real-world systems demonstrate the effectiveness of the proposed algorithms and validate the theoretical results.**

***Index Terms*— DC microgrids, optimal load distribution, load variations, distributed algorithms, stochastic analysis.**

NOMENCLATURE

$OLTA$:	Optimal load tracking and allocation
DC/AC :	Direct/Alternating current
MG :	Microgrids
PV :	Photovoltaic
LFC :	Load flow converter
EV :	Electric vehicle
PWM :	Pulse-width modulation

I. INTRODUCTION

DC microgrids (MGs) offer unique advantages in distribution-level power grids with high penetration of renewable generators, energy storage systems, and electric vehicles. Vehicle electrification, including passenger vehicles, trolley buses, electric trains, has emerged as a critical driving force for reducing traffic congestion and air pollution [1], [2]. Since PV generators, fuel cells, battery energy storage systems, electric buses, and trolley buses can generate and utilize DC power directly, DC MGs have become a highly feasible and efficient framework deployed in many large cities in the world [3].

DC MGs avoid unnecessary AC/DC conversions and can simplify control design. To achieve high-performance control of such systems, several advanced optimization and control algorithms have been designed, e.g., primal-dual decomposition distributed optimization algorithms for power, voltage, and droop control [4], optimal power sharing control strategies [5], distributed, robust, and optimal control architecture [6], LFC methods for connecting two adjacent DC MGs and controlling their bidirectional power flows [7]. Network control strategies have been developed for DC MGs, exemplified by distributed secondary/primary control frameworks [8], and decentralized current-sharing control strategies for fast transient response [9]. Reference [10] studies a bidirectional hybrid DC-DC converter as an interface between two DC voltage buses in DC MG applications. In [11] autonomous DC voltage control is studied for a DC MG with multiple power and slack terminals. Moreover, [12] presents a decentralized output constrained control algorithm for single-bus DC MGs, and [13] proposes a new distributed control scheme that achieves current sharing and average voltage regulation in DC MGs. In addition, [14] develops an observer-based DC voltage droop and current feed-forward control for a DC MG; and [15] employs a trust-based cooperative controller to detect and mitigate adverse effects of attacks on communication links and controller hijacking by using only local and neighbor information. Furthermore, [16] introduces a distributed nonlinear controller to achieve simultaneously accurate current-sharing and voltage regulation in DC MGs. These previous results do not deal with stochastic load variations from vehicle-grid integration, which will be the focus of this paper.

Optimal load tracking and allocation (OLTA) of DC MGs for supporting charging stations, in-road wireless charging networks, trolleybuses, subways, and regional trains is highly challenging [17], [18]. For example, trolleybuses offer unique advantages in fuel economy, cost reduction, and passenger capacity, but their mobility introduces highly volatile load variations [19], [20]. In our early work [19]–[21], distributed optimization strategies for such DC MGs were developed for

Manuscript received April 19, 2020; revised August 14, 2020; accepted March 11, 2021. This work was supported in part by the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-18-1-0268. The Associate Editor for this article was R. Arghandeh. *(Corresponding author: Siyu Xie.)*

Siyu Xie, Masoud H. Nazari, and Le Yi Wang are with the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202 USA (e-mail: syxie@wayne.edu; masoud.nazari@wayne.edu; lywang@wayne.edu).

George Yin is with the Department of Mathematics, University of Connecticut, Storrs, CT 06269 USA (e-mail: gyin@uconn.edu).

Wen Chen is with the Division of Engineering Technology, Wayne State University, Detroit, MI 48202 USA (e-mail: wchenc@wayne.edu).

Digital Object Identifier 10.1109/TITS.2021.3067196

achieving optimal tradeoff among load allocation, reduction of power loss, and voltage management. These methods deal with constant loads and study OLTA in a deterministic framework of optimization.

However, it is well known that the defining features of MGs with distributed renewable generators and diversified load types include randomly intermittent power outputs from solar and wind generators and large load variations from vehicles of diversified types. Consequently, it is critically important to understand performance, optimality, and limitations of distributed algorithms under persistent and stochastic generation and load variations. This paper aims to present a comprehensive and rigorous study of the impact of stochastic generation and load variations on optimality and performance of distributed OLTA algorithms. While the methodology of this paper is applicable to a broad class of distributed algorithms in power systems, for concreteness and conciseness this paper will employ the DC MG platform and distributed algorithms in [20], [21] to derive models, perform system analysis, and understand optimality and tradeoff under persistent stochastic load variations and noisy observations.

In DC MGs for transportation electrification, both generation and load variations can be viewed and aggregated as stochastic load disturbances [22]–[24]. Under such stochastic load disturbances, the optimal solutions become stochastic processes, whose bounds, convergence, and stochastic features become more difficult to analyze. These aspects are not covered in [20], [21]. To the best of our knowledge, this is the first comprehensive and rigorous study of such issues in distributed OLTA problems in DC MGs for transportation electrification.

Research on smart grids and related power management strategies has been quite extensive; see [25], [26] for distribution power systems with renewable and distributed generators. For optimization performance measures, it is noted that power losses are a traditional performance criterion and have been considered in many MGs [27]–[30]. Within centralized strategies, global optimality has been analyzed in [31]. It is well known that traditional frequency and voltage droop control methods for standing-alone MGs are decentralized [32]–[35], which have fundamental limitations in achieving optimality. Distributed strategies have been pursued [36]–[39] with promising features [33], [40]. In our early work, a consensus method was introduced to power system load allocation in [41]. These results do not include stochastic load variations.

This paper extends the models of DC MGs in [20], [21] by including stochastic uncertainties. The main contributions of the paper are in the following aspects: 1) We have formulated and studied impact of persistent and stochastic load variations on distributed optimization algorithms in DC MGs. It is shown that interaction between load variations and noisy observations introduces fundamental tradeoff in reducing errors on optimal solutions. 2) We have developed distributed optimal algorithms to track time-varying loads, attenuate observation noises. Error bounds and convergence properties have been established rigorously and comprehensively. 3) Relationship among algorithm step size selection, error bound, and tracking ability have been derived quantitatively. 4) The limiting behavior of the optimality errors has been derived that characterizes the fundamental impact of the step size selection on attenuating observation noises and tracking load changes. 5) Optimality conditions and algorithms for the optimal step size have been introduced to guide step size selection in practical applications.

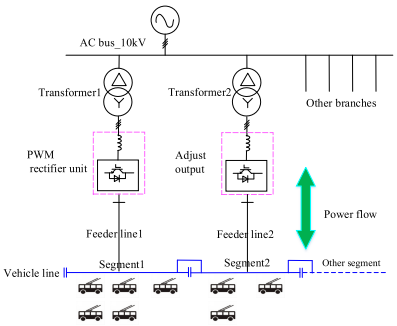


Fig. 1. Diagram of a new trolleybus power supply network.

The remainder of the paper is organized as follows. In Section II, OLTA problems in DC MGs are formulated with motivations from DC networks that support electric trolleybus systems. Global optimality conditions and distributed optimization algorithms are presented in Section III. Convergence properties of the algorithms are established in Section IV under different types of observation noise and load variations. Fundamental tradeoff and optimal selection of step size are discussed. Theoretical results are evaluated by simulation case studies in Section V. Finally, Section VI provides concluding remarks on our findings and potential future directions. For clarity of presentation, the proofs of most technical results are relegated to the Appendix.

Notation: For any deterministic matrix $X \in \mathbb{R}^{m\times n}$, its norm is denoted by $\|X\| = (\lambda_{\max}(XX^\top))^{\frac{1}{2}}$, where $\lambda_{\max}(\cdot)$ is the largest eigenvalue of a matrix; and for any stochastic vector $x \in \mathbb{R}^n$, its norm is defined as $\|x\| = (E[x^\top x])^{\frac{1}{2}}$, where $E[\cdot]$ is the mathematical expectation operator. We also use $\mathrm{Tr}(\cdot)$ to denote the trace of the corresponding matrix. For two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times m}$, $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ holds.

II. Problem Formulation

A. Motivating Scenarios

The mathematical formulation of power management in DC MGs is motivated by the Beijing Dual-Source Trolleybus Power Supply Network that was used in [20]. While the methodologies, algorithms, convergence properties can be readily extended to a broad class of DC MGs, for concreteness and conciseness this paper focuses on OLTA problems in DC MGs for transportation electrification.

In the DC power supply MG, segments can be interconnected as shown in Fig. 1. The current output of each converter can be controlled. Distributed OLTA strategies aim at using own and the neighboring feeders' information so that loads on feeders can be distributed more evenly to avoid overloads.

As explained in [20], in early 2012, 180 new dual-source trolleybuses were put into operation by the Beijing Public Transport Group. Then, 16 dual-source trolleybus lines (with onboard battery systems) were developed with more than

588 trolleybuses in the next three years. The total length of the trolleybus routes was about 214 km, which were supplied by 15 power stations. These DC power stations form a MG whose OLTA is critically important for suitable load distribution to enhance safety, reliable operation, and energy efficiency, under highly volatile random generation/load variations. The work in [19], [20] does not consider stochastic load variations and their impact on optimality, efficiency, and safety. This paper will focus on these aspects.

B. Mathematical Formulation and Assumptions

Consider a grid of n subsystems. Each subsystem may be a feeder of a street section, a renewable generating site, a charge station, a community, etc. For $i \in \{1, \ldots, n\}$, u_i is the local power or current input, which represents all controllable assets, and ℓ_i is the local power consumption which represents all uncontrollable assets and is viewed as a disturbance. The power flow from node i to node j is denoted by x_{ij}. The vectors containing the control, load, and network state will be denoted by u, ℓ, and x, respectively. The network topology of the MG will be represented by an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E}$ is the set of edges.

In a market-oriented grid structure, power exchange between subsystems introduces a power loss or cost that can be included in the performance index. Mathematically, we will use the following separable and convex performance index:

$$\min J = f(u, x) = \sum_{i \in \mathcal{V}} f_i(u_i) + \sum_{(i,j) \in \mathcal{E}} f_{ij}(x_{ij}). \qquad (1)$$

In particular, we will consider the following quadratic performance measures:

$$\min J = f(u, x) = \frac{1}{2}(x^\top R x + u^\top Q u) \qquad (2)$$

with $R \in \mathbb{R}^{(n-1)\times(n-1)}$ and $Q \in \mathbb{R}^{n \times n}$, and the matrices R, Q be diagonal positive definite matrices. Here, the first term reflects the cost of current production or penalty on overload, and the second term represents the line power loss or trading cost between users. We assume that the variations of voltage are negligible, thus voltage regulation is not pursued in this paper.[1]

The physical network imposes relations among the variables, which is expressed as equality constraints $h(u, x, \ell) = 0$. For the typical DC MGs exemplified in [20], the physical system constraints entail a linear structure which will be adopted here.

Assumption 2.1: The constraint equation is linear $u = Mx + \ell$, and M is full column rank.

In trolleybus systems, the currents of different feeder lines vary rapidly and substantially; see Fig. 2 for some recorded load variations during a three-hour operation. In general, for smart grids with high penetration of renewable generators, the generation/load variations are substantial and stochastic.

[1]The scope of this paper is on load tracking and allocation in DC MGs. The reader is referred to [21] for an integrated load distribution and voltage regulation.

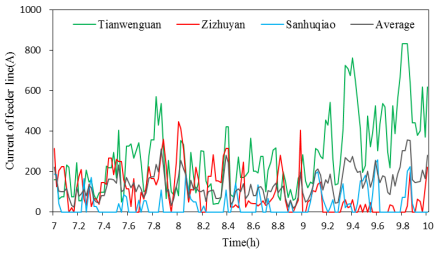


Fig. 2. Substantial and rapid load variations on different feeders.

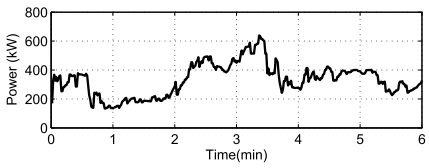


Fig. 3. Load variations on a feeder.

The generation/load variation ℓ is a stochastic process. The load variation $\delta(k+1) = \ell(k+1) - \ell(k)$ represents the "rate of variation". Furthermore, measurement of variables in a network encounters measurement and communication errors. Such errors are denoted by a stochastic process $d(k)$ which corrupts measured variables; see (6).

Since the buses in the dual-source trolleybus have on-board battery systems, very fast load variations can be absorbed (smoothed) by the battery. After smoothing, the resulting load variations remain stochastic, but can be reasonably modeled as "slowly time variations" ($\delta(k)$ is bounded and small). Due to traffic changes during day and night, load variations are non-zero mean; see a typical load profile on a feeder in Fig. 3. Consequently, optimal solutions must track "load trends". These features are put in an abstract form in our assumptions on observation noise and load variations.

Although it is common to assume independent and zero mean observation noises or variations, to accommodate practical scenarios of dependent noises and drifting load variations overtime (such as from daytime peak demand to nighttime low usage), this paper includes a much broader class of noise and load variations. In particular, we consider the following three classes of noise and load stochastic processes that will include as special cases the martingale difference, zero-mean noise, ϕ-mixing and α-mixing sequences, and processes driven by white noises via linear systems.

1) **Bounded Load Variations and Observation Noise**
   In DC MGs that support electric vehicle networks, load variations are often substantial, not independent, and not Gaussian. Here we consider bounded noise and load variations.
   Assumption 2.2: $\sigma \stackrel{\triangle}{=} \sup_k \|d(k)\| < \infty$, and $\delta \stackrel{\triangle}{=} \sup_k \|\delta(k)\| < \infty$.
2) **Bounded Load Variations and Observation Noise with Decaying Averages**
   A random sequence $x = \{x(k)\}$ is called an element of the weakly dependent set $\mathcal{D}$, if there exists a constant C_x depending only on the distribution of $\{x(k)\}$ such that for any $k \geq 0$ and $h \geq 1$, $\left\| \sum_{i=k+1}^{k+h} x(i) \right\| \leq C_x h^{\frac{1}{2}}$.

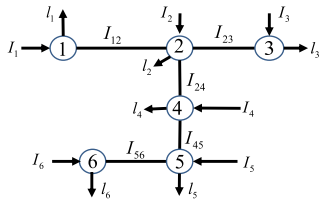


Fig. 4. A DC power supply network for the electric trolley bus system. The network is a state tree.

It is known that the martingale difference, zero mean ϕ-mixing and α-mixing sequences, and the linear process driven by white noises are all in $\mathcal{D}$ [42]. The stochastic processes in $\mathcal{D}$ are uniformly bounded (by using $h = 1$), but its average $\frac{1}{h}\left\|\sum_{i=k+1}^{k+h} x(i)\right\| \leq C_x/h^{\frac{1}{2}}$. As a result, the average decays to zero with rate $1/h^{\frac{1}{2}}$. Thus, load variations are volatile over a small period of time (such as trolleybus loads in rush hours), but on average they demonstrate certain sustained loads (such as the daily average electricity consumption).

Assumption 2.3: $\{d(k), k \geq 0\} \in \mathcal{D}$, and $\{\delta(k), k \geq 0\} \in \mathcal{D}$.

3) **Bounded Load Variations and Zero-Mean Observation Noise**

Typical scenarios in practical systems consist of observation noises that are zero mean and load variations that are slowly time varying but can change substantially over a long horizon. This is reflected in the following assumptions.

Assumption 2.4: 1) $\{d(k)\}$ is a sequence of independent and identically distributed random variables with $E[d(k)] = 0$ and $E[d(k)d^\top(k)] = P_d$. 2) $\{\delta(k)\}$ is a sequence of independent and identically distributed with $E[\delta(k)] = \eta_\delta$ and $E[\delta(k)\delta^\top(k)] = P_\delta$. 3) $x(0)$, $d(k)$, and $\delta(k)$ are mutually independent.

C. A DC MG Example

Before we further detail algorithms and their convergence properties, we will use an example to illustrate the main structures of the typical DC MGs that will be used in our subsequent development.

Example 2.1: Consider the feeder system for the "Junbaose Trolleybus Network" in Beijing. Its network consists of 6 feeders and its network topology is shown in Fig. 4. The supply currents $(I_i, i \in \{1, 2, \ldots, 6\})$ and the load currents $(\ell_i, i \in \{1, 2, \ldots, 6\})$ are local variables on the feeders, and the transmission line currents $(I_{ij}, (i, j) \in \{(1, 2), (2, 3), (2, 4), (4, 5), (5, 6)\})$ are network variables. Here we assume that the load currents are given. The local supply currents are the actual local control variables u and the transmission line currents are the network states x. Thus, u and x are dependent variables.

These physical variables are related by Kirchhoff's current law equality constraints. If we fix the loads, the bus supply currents (u) are the only control variables that can be physically manipulated. All line currents will follow as physical dependent variables. The network graph contains 6 nodes, i.e., $\mathcal{V} = \{1, 2, 3, 4, 5, 6\}$, and 5 edges, i.e., $\mathcal{E} = \{(1, 2), (2, 3), (2, 4), (4, 5), (5, 6)\}$. Define $u = [I_1, \ldots, I_6]^\top$, $x = [I_{12}, I_{23}, I_{24}, I_{45}, I_{56}]^\top$, and $\ell = [\ell_1, \ldots, \ell_6]^\top$, then the equality constraint is $u = Mx + \ell$, where

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \in \mathbb{R}^{6\times 5}. \tag{3}$$

We note that these constraints are consistent with the physical network topology. Given x and ℓ, obtaining u involves only local computation.

It is easy to verify that $\text{rank}(M) = 5$. Note that u is the actual physical control variables, and $x = (M^\top M)^{-1}M^\top(u - \ell)$ is the physical dependent variable, which is global, namely, computing x_{ij} needs global information on u.

In general, we may write the equality constraint as $u = Mx+\ell$, where $u \in \mathbb{R}^n$, $x \in \mathbb{R}^{n-1}$, $\ell \in \mathbb{R}^n$, $M \in \mathbb{R}^{n\times(n-1)}$. It is easy to verify that each column of M has only 1 and -1 with other elements being 0, which means that at each link, power flows from one subsystem to another. Moreover, M is full column rank, indicating $x = (M^\top M)^{-1}M^\top(u - \ell)$.

III. Global Optimality and Distributed Algorithms

A. Global Optimality Conditions

Theoretically, the global optimal solution of (2) with the equality constraint $u = Mx + \ell$ can be obtained by the Lagrange Multiplier method: For $\lambda \in \mathbb{R}^{n\times 1}$,

$$L = \frac{1}{2}(x^\top Rx + u^\top Qu) + \lambda^\top(u - Mx - \ell), \tag{4}$$

with the following optimality conditions:

$$\begin{cases} \nabla_x L = Rx - M^\top\lambda = 0, \\ \nabla_u L = Qu + \lambda = 0, \\ \nabla_\lambda L = u - Mx - \ell = 0. \end{cases}$$

Thus, the optimal solution is

$$x^* = -G^{-1}M^\top Q\ell, \quad u^* = Mx^* + \ell, \quad \lambda^* = -Qu^*,$$

where $G = R + M^\top QM$.

However, this global optimal solution is not feasible in distributed strategies since we have to calculate the inverse of the matrix $G = R+M^\top QM$. In our methodology, we seek (strictly) distributed optimization on x and then update u accordingly. To be concrete, we assume that each subsystem only knows its own state and the powers and loads of the nodes that are connected to it. We first define the following performance index:

$$J(x) = \frac{1}{2}(x^\top Rx + (Mx + \ell)^\top Q(Mx + \ell)), \tag{5}$$

whose gradient is $\nabla_x J = Rx+M^\top Q(Mx+\ell) = Gx+M^\top Q\ell$.

Let $y = -M^\top Q\ell$. Then, the optimal solution is reduced to seek x^* such that $y = Gx^*$. This can be viewed as an adaptive filtering problem since the solution x^* depends on

the time-varying load ℓ, and hence it is an unknown and time-varying process.

B. Distributed Optimization Algorithms

In algorithm development, we note that y is locally measured with errors at each time instant $k \geq 0$, and hence can be expressed as

$$y(k) = Gx^*(k) + d(k) \tag{6}$$

where $x^*(k)$ is the true time-varying parameter (the optimal power dispatch) which needs to be tracked, and $d(k)$ is a stochastic measurement noise process. The variation of ℓ is $\delta(k+1) = \ell(k+1) - \ell(k)$. Thus,

$$\begin{aligned} x^*(k+1) - x^*(k) &= -G^{-1}M^\top Q[\ell(k+1) - \ell(k)] \\ &= -G^{-1}M^\top Q\delta(k+1). \end{aligned} \tag{7}$$

Here we can adopt the following algorithm to track the unknown parameter $x^*(k)$:

$$x(k+1) = x(k) - \mu\nabla_x J(k) = x(k) + \mu(y(k) - Gx(k)) \tag{8}$$

where $\mu > 0$ is the step size. The above algorithm is strictly distributed, since the update of $x_{ij}(k)$ only involves the information from its own and neighbors, which can be seen from the example in Section VI.

Algorithm (8) is a stochastic approximation whose step size μ is a design variable. In other words, there is a fundamental tradeoff in selecting μ: 1) To reduce the noise effect, the step size must be reduced gradually to zero so that the algorithm converges to the true optimal solution. 2) However, when the step size approaches zero, $x(k)$ will approach a fixed value as k approaches infinity, losing its ability to track time varying loads.

In a stochastic environment, the gradient-based and distributed algorithm (8) employed in this paper has been shown to achieve the Cramér-Rao (CR) lower bound[2] asymptotically, and it is "optimal" in the sense of convergence rate. This allows us to study the fundamental tradeoff that is pursued in this paper. Other commonly used searching algorithms such as primal-dual, alternating direction method of multipliers (ADMM), and dual-sub-gradient algorithms do not have stochastic convergence results. Thus, we employ (8) for the performance analysis. In Fig. 5, the main ideas, cyber-physical interaction, and algorithm structures are illustrated.

Before analyzing the property of the above algorithm, by substituting (6) into (8) and using the notation (7), we can first establish the following tracking error equation for $\widetilde{x}(k) = x(k) - x^*(k)$:

$$\begin{aligned} \widetilde{x}(k+1) &= x(k) + \mu(Gx^*(k) + d(k) - Gx(k)) \\ &\quad -x^*(k) + G^{-1}M^\top Q\delta(k+1) \\ &= (I - \mu G)\widetilde{x}(k) + \mu d(k) \\ &\quad +G^{-1}M^\top Q\delta(k+1), \end{aligned} \tag{9}$$

[2]This is the lower bound on the variance of unbiased estimators of a fixed and unknown parameter given the observation data points [43], which is the reciprocal of the Fisher information. The Fisher information is calculated from the variance of the natural logarithm of the likelihood function. Any algorithm that can achieve the CR bound is the "fastest" algorithm.

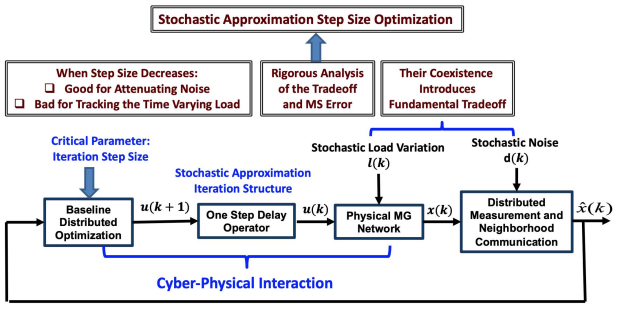


Fig. 5. The cyber-physical structure and the main logic of the algorithm.

which can be written as

$$\widetilde{x}(k+1) = F_\mu\widetilde{x}(k) + \mu d(k) + H\delta(k+1) \tag{10}$$

where $F_\mu = I - \mu G = I - \mu(R + M^\top QM)$, and $H = G^{-1}M^\top Q = (R + M^\top QM)^{-1}M^\top Q$.

Lemma 3.1: Under *Assumption 2.1*, there exists constants $\mu^* > 0$ and $c > 0$ such that for any $\mu \in (0, \mu^*)$, $0 \leq 1 - \mu c < 1$ and $\|F_\mu^{k-i}\| \leq (1 - \mu c)^{k-i}$.

For clarification, all proofs are presented in the Appendix.

IV. Convergence Properties

When loads are persistently time varying, global optimal solutions must track the loads and are subject to certain tracking errors. In this section, the main convergence properties and error bounds of the algorithms will be established.

Before detailed convergence analysis under different types of noise and load variations, it is noted that if $\{d_k\}$ and $\{\delta_k\}$ are uncorrelated sequences with zero mean and that $\{d_k\}$ and $\{\delta_k\}$ are independent, then the Lyapunov approach may be used to establish exponential convergence as follows: From (9) and (10), one may define a Lyapunov candidate $V(x) = E\|x\|^2$. Then, it is readily seen that there exists a constant $\mu^* > 0$ such that for any $\mu \in (0, \mu^*)$ and any initial condition, $E\|\widetilde{x}(k)\|^2 \to 0$ exponentially.

In power system applications, load variations are typically non-zero mean. One of the main goals of this paper is to establish errors on optimality from combined load variations and observation noise. To accommodate different application scenarios, we will consider three types of noise and load variations, derive corresponding error bounds on optimal solutions, and discuss design principles.

A. Bounded Load Variations and Observation Noises

When both noises and load variations are uniformly bounded, we can obtain the following tracking error bound.

Theorem 4.1: Under *Assumptions 2.1* and *2.2*, for all $k > 0$, there exists a constant $\mu^* > 0$ such that for any $\mu \in (0, \mu^*)$,

$$\|\widetilde{x}(k)\| \leq \frac{1}{c}\left(\sigma + \frac{\delta b}{\mu}\right) + (1 - \mu c)^{k+1}\|\widetilde{x}(0)\|, \tag{11}$$

where $c \in (0, 1)$, and it is a constant defined in *Lemma 3.1*, and $b = \|H\|$.

Remark 4.1: Note that when both the measurement noise $d(k)$ and load variation $\delta(k)$ are small, the tracking error will be small. We further remark that since the measurement noise $d(k)$ in *Theorem 4.1* is not assumed to be zero mean, it is not

expected that the upper bound of the tracking error will tend to zero as μ tends to zero when $\delta(k) = 0$. The meaningfulness of *Theorem 4.1* lies in the fact that when the magnitudes of both the measurement noise $d(k)$ and the load variation $\delta(k)$ are small, the upper bound of the tracking error $\widetilde{x}(k)$ will be small. Moreover, if we further assume that the measurement noise $d(k)$ satisfies more statistic properties including zero mean, then by *Theorem 4.2*, it is possible to see that when $\delta(k) = 0$, the upper bound of the estimation error will tend to zero as μ tends to zero.

B. Bounded Load Variations and Observation Noises With Decaying Averages

With additional assumptions, we can further establish a better upper bound for the error equation. This assumption simply implies that both the measurement noise and load variation are weakly coupled.

Theorem 4.2: Under *Assumptions 2.1* and *2.3*, for all $k > 0$, there exists a constant $\mu^* > 0$ such that for any $\mu \in (0, \mu^*)$,

$$\|\widetilde{x}(k)\| \leq C_c C_d \sqrt{\mu} + \frac{bC_c C_\delta}{\sqrt{\mu}} + (1 - \mu c)^{k+1} \|\widetilde{x}(0)\|, \quad (12)$$

where $c \in (0, 1)$ is a constant, and it is defined in *Lemma 3.1*, C_c is a constant only depends on c, $b = \|H\|$, and C_d, C_δ are two constants defined in *Assumption 2.3*.

Remark 4.2: Note that the upper bound in *Theorem 4.2* shows that if the step size μ decreases, the tracking error caused by the measurement noise will decrease, but the tracking error caused by the load variation will increase, and vice versa. Therefore, there is a fundamental tradeoff between noise sensitivity and tracking ability for (9).

Remark 4.3: Compare with *Theorem 4* in [20], if the noise $d(k)$ is a ϕ-mixing sequence such that $E[d(k)] = 0$, we know that $\{d(k), k \geq 0\} \in \mathcal{D}$ holds [42]. If we further assume that $\delta \stackrel{\triangle}{=} \sup_k \|\delta(k)\| < \infty$, then by *Theorems 4.1* and *4.2*, we know that for all $k > 0$, there exists a constant $\mu^* > 0$ such that for any $\mu \in (0, \mu^*)$,

$$\|\widetilde{x}(k)\| \leq C_c C_d \sqrt{\mu} + \frac{b\delta}{c\mu} + (1 - \mu c)^{k+1} \|\widetilde{x}(0)\|.$$

Thus, the upper bound of the tracking error will tend to zero as μ tends to zero when $\delta(k) = 0$.

C. Bounded Load Variations and Zero-Mean Observation Noises

Denote $E[\widetilde{x}(k)] = \eta_{\widetilde{x}}(k)$, $\eta_{\widetilde{x}} = \lim_{k\to\infty} \eta_{\widetilde{x}}(k)$, $\Sigma_{\widetilde{x}}(k) = E[(\widetilde{x}(k) - \eta_{\widetilde{x}})(\widetilde{x}(k) - \eta_{\widetilde{x}})^\top]$, and $\Sigma_{\widetilde{x}} = \lim_{k\to\infty} \Sigma_{\widetilde{x}}(k)$, when the limits exist.

Theorem 4.3: Under *Assumptions 2.1* and *2.4*, for all $k \geq 1$, there exists a constant $\mu^* > 0$ such that for any $\mu \in (0, \mu^*)$,

1) $\eta_{\widetilde{x}}(k) = F_\mu^k \eta_{\widetilde{x}}(0) + \sum_{\ell=0}^{k-1} F_\mu^\ell H \eta_\delta$.
2) $\eta_{\widetilde{x}} = \frac{1}{\mu} G^{-1} H \eta_\delta$.
3) The error variance

$$\Sigma_{\widetilde{x}}(k) = F_\mu^k \Sigma_{\widetilde{x}}(0) F_\mu^k + \sum_{\ell=0}^{k-1} F_\mu^\ell (\mu^2 P_d + H P_\delta H^\top) F_\mu^\ell. \quad (13)$$

4) $\Sigma_{\widetilde{x}}$ is the solution to the (discrete-time) Lyapunov equation $F_\mu \Sigma_{\widetilde{x}} F_\mu - \Sigma_{\widetilde{x}} = -\mu^2 P_d - H P_\delta H^\top$, or explicitly

$$\Sigma_{\widetilde{x}} = \sum_{\ell=0}^{\infty} F_\mu^\ell (\mu^2 P_d + H P_\delta H^\top) F_\mu^\ell. \quad (14)$$

The total mean-square optimality error is a function of the step size μ

$$V(\mu) = E[\|\widetilde{x}\|^2] = \mathrm{Tr}(\Sigma_{\widetilde{x}}) + \eta_{\widetilde{x}}^\top \eta_{\widetilde{x}}.$$

To reduce the optimality error, it is natural to seek $\min_\mu V(\mu)$. The optimality condition for the step size selection is given below.

Proposition 4.1: Under *Assumptions 2.1* and *2.4*, the optimal step size μ^o that minimizes $\min V(\mu) = \mathrm{Tr}(\Sigma_{\widetilde{x}}) + \eta_{\widetilde{x}}^\top \eta_{\widetilde{x}}$, is the solution $\mu^o \in (0, \mu^*)$ to

$$\mu\, \mathrm{Tr}(\, (2I - \mu G)^{-2} G^{-1} (H P_\delta H^\top - \mu G H P_\delta H^\top - \mu^2 P_d)) + \eta_\delta^\top H^\top G H \eta_\delta = 0. \quad (15)$$

Remark 4.4: It can be shown that $V(\mu)$ is smooth and strongly convex, see Fig. 10 in *Example 5.2*. As a result, μ^o exists and is unique. Since (15) is a scalar nonlinear convex function, its solution via line search is easily obtained by classical numerical algorithms, such as the Newton-Raphson method. [3]

To understand how the optimal step size changes with the size of the observation noise and load variation rate, consider the simplified case $P_\delta = \delta_0 I$, $\eta_\delta = \eta_0 [1, \ldots, 1]^\top$, and $P_d = d_0 I$. Then, (15) becomes

$$\mu\, \mathrm{Tr}((2I - \mu G)^{-2} G^{-1} (\frac{\delta_0}{d_0} (I - \mu G) H H^\top - \mu^2)) + \frac{\eta_0^2}{d_0} H^\top G H = 0.$$

It follows that μ^o is a function of the ratios $\frac{\delta_0}{d_0}$ and $\frac{\eta_0^2}{d_0}$. In particular, when the load variation is very small, it can be shown that $\lim_{\delta_0, \eta_0 \to 0} \mu^o = 0$ and correspondingly $V(\mu) \to 0$. In other words, the constant load case is indeed the limiting case of time-varying load situations.

Remark 4.5: The optimal value μ^o depends on the statistical properties of the observation noise and the load variation, which are impossible to be obtained in advance in many practical situations. A potentially useful approach is to adjust the step size μ with an adaptive scheme which is designed to converge to the optimal value [44]. This is a possible research problem to be considered in the future.

V. Case Study for Evaluation

A. Testing System and Performance Measures

To illustrate and evaluate the theoretical results, we employ the Beijing Dual-Source Trolleybus System in Fig. 1. The "Junbaose Trolleybus Network" is selected, whose power supply network contains 6 feeders with the topology in Fig. 4. Its communication network topology is identical to the physical

[3] Without noise, exponential convergence rates can be obtained for such line search problems for strongly convex functions.

TABLE I
LINK RESISTANCE OF JUNBAOSE TROLLEYBUS NETWORK

Link Number	(1,2)	(2,3)	(2,4)	(4,5)	(5,6)
Link Length (km)	2	1.9	1.7	1.55	1.8
Line Resistance (ohm)	0.4	0.38	0.34	0.31	0.36

network. In this simulation study, we assume that all lines are to be operated within the physical limit of 1100 A. The load conditions and system parameters are obtained from the physical system measurements.

The resistance density of the cable line is 0.2 (Ohm) per kilometer. Based on the length of each line, we can calculate the resistances between the segments, and we can calculate the line resistance values as shown in Table I. This leads to the line power loss weighting matrix $R = \text{diag}\{0.4, 0.38, 0.34, 0.31, 0.36\}$.

The network has communication lines between 1 and 2, 2 and 3, 2 and 4, 4 and 5, 5 and 6 in the network. Thus, we have $u = [I_1, I_2, I_3, I_4, I_5, I_6]^\top$, $x = [I_{12}, I_{23}, I_{24}, I_{45}, I_{56}]^\top$, and $\ell = [\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6]^\top$ with the matrix M has the form (3). Thus, $u = Mx + \ell$ holds. Here we assume that the initial currents of feeder lines are $I_1(0) = 713\ A, I_2(0) = 811\ A, I_3(0) = 960\ A, I_4(0) = 844\ A, I_5(0) = 887\ A, I_6(0) = 823\ A$ and the initial loads on the segments are $\ell_1(0) = 681\ A, \ell_2(0) = 783\ A, \ell_3(0) = 1009\ A, \ell_4(0) = 842\ A, \ell_5(0) = 921\ A, \ell_6(0) = 803\ A$.

The power dispatch has two objectives: 1) balancing the feeder currents, namely $I_i(k) \to \bar{I}, i = 1, \ldots, 6$, where $\bar{I} = \sum_{i=1}^{6} I_i(0)/6$; 2) reducing the line losses $\sum_{ij}^{5} I_{ij}^2 R_{ij} = x^\top R x$. Thus, the performance index is given as follows:

$$\min_{x,u} J = \frac{1}{2}\Big(x^\top R x + (u - \bar{I}z)^\top (u - \bar{I}z)\Big),$$
$$s.t.\ \ u = Mx + \ell, \tag{16}$$

where z is a column vector of all 1's. Since $M^\top z = 0$, it is easy to see that the gradient of

$$J(x) = \frac{1}{2}\Big(x^\top R x + (Mx + \ell - \bar{I}z)^\top (Mx + \ell - \bar{I}z)\Big)$$

is $\nabla J_x = Rx + M^\top (Mx\ell) = (R + M^\top M)x + M^\top \ell$, and the optimal solution is $x^* = -(R + M^\top M)^{-1} M^\top \ell$ and $u^* = Mx^* + \ell$.

We consider the following observation equation: $y(k) = -M^\top \ell(k) + d(k) = (R + M^\top M)x^*(k) + d(k)$, where $y(k) = [y_1(k), \ldots, y_5(k)]^\top \in \mathbb{R}^5$ and $d(k) = [d_1(k), \ldots, d_5(k)]^\top \in \mathbb{R}^5$, and we can adopt (8) to track $x^*(k)$, i.e., $x(k+1) = x(k) + \mu[y(k) - (R + M^\top M)x(k)]$. By the structure of the matrix $R + M^\top M$, the above algorithm is distributed, since the adaptation of I_{ij} only involves the information from nodes i and j. Then, the corresponding control input at time instant k should be $u(k) = Mx(k) + \ell(k)$.

B. Evaluation Results

Example 5.1: We first demonstrate convergence properties. Assume that for $i = 1, \ldots, 6$ and $k \geq 0$, the observation noise $d_i(k) \sim N(0, 10^2)$, and the variation of the load $\ell_i(k) \sim N(0, \delta^2)$ where $\delta \geq 0$ is a constant. Next,

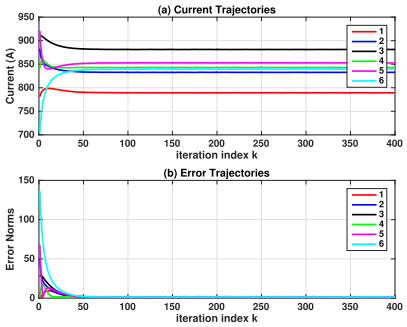


Fig. 6. Currents and error trajectories with $\delta = 0$, $\mu = 0.1$.

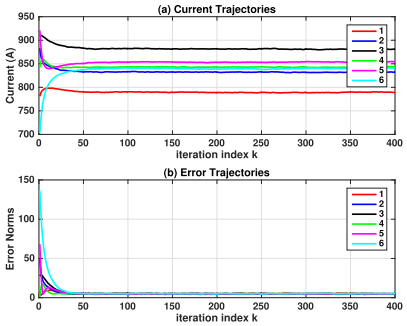


Fig. 7. Currents and error trajectories with $\delta = 5$, $\mu = 0.1$.

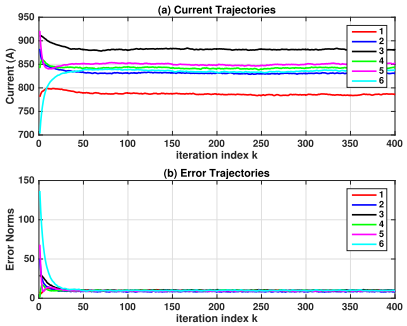


Fig. 8. Currents and error trajectories with $\delta = 10$, $\mu = 0.1$.

we repeat the simulation for $m = 500$ with the same initial states. Then, for each I_i, we can get 6 m sequences $\{I_i^s(k),\ k = 1, 2, \ldots, 200\}$, $i = 1, \ldots, 6$, $s = 1, \ldots, m$, where the superscript s denotes the s-th simulation result. We use $\frac{1}{m}\sum_{s=1}^{m} I_i^s(k)$ and $\frac{1}{m}\sum_{s=1}^{m} \|I_i^s(k) - I_i^*(k)\|$, $k = 1, 2, \ldots, 200$, $i = 1, \ldots, 6$, to approximate the mean current inputs and the mean estimation or tracking errors, respectively. Note that $I_i^*(k)$ is the time-varying optimal solution, and the step size should satisfy $0 < \mu < 0.4370$.

We select $\mu = 0.1$ to compare the performance of the filtering algorithm with different load variations for $\delta = 0$, $\delta = 5$ and $\delta = 10$. The results are shown in Fig. 6, Fig. 7, and Fig. 8. These simulations show that the error trajectories reduce to a neighborhood of zero, which implies that the currents tend to the global optimal solution.

Fig. 9 provides the relationship between δ and the following tracking errors (with $\mu = 0.1$) $ER(k) = \frac{1}{m}\sum_{s=1}^{m} \|I^s(k) - I^*(k)\|$, $k \geq 1$, where $I^s(k) = [I_1^s(k); \ldots; I_6^s(k)]$, and $I^*(k) = [I_1^*(k); \ldots; I_6^*(k)]$ is the optimal solution. Note that with the same step size, the larger δ leads to the larger tracking errors.

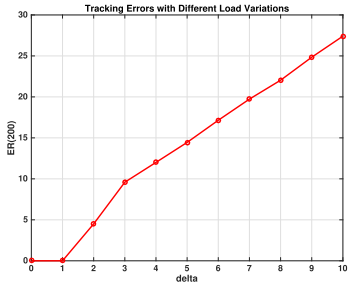


Fig. 9. Tracking errors with different load variations.

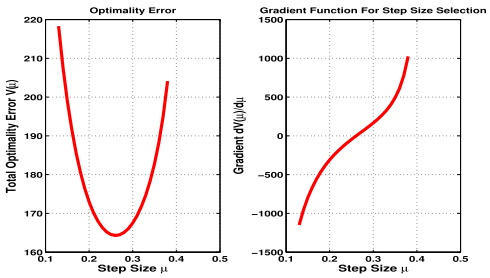


Fig. 10. The optimality error and selection of step size.

Example 5.2: The step size must be carefully selected to balance the impact from observation noise and load variations. In this example, different step sizes are used to compare the performance of the algorithm. Here, $\delta = 5$, $P_\delta = 25I_6$, $P_d = 100I_6$, $\eta_\delta = 0$. By *Proposition 4.1*, the optimal step size satisfies (15). Fig. 10 shows $V(\mu) = \mathrm{Tr}((2\mu G - \mu^2 G^2)^{-1}(\mu^2\ P_d + H P_\delta H^\top)) + \frac{1}{\mu^2}\eta_\delta^\top H^\top G H \eta_\delta$ and its gradient. Clearly, $V(\mu)$ is a smooth and convex function, and its minimum exists and is unique. The convex function solution to $V(\mu)$, namely the root to the gradient can be obtained numerically by using the Newton-Raphson algorithm, with $\mu^o = 0.2609$. This function illustrates the fundamental trade-off in attenuating noise effects and tracking time-varying loads.

VI. Concluding Remarks

This paper studies the optimal distributed power dispatch problem in cyber-physical DC MGs for transportation electrification under stochastic load variations. Compared with related existing works, the impact of persistent stochastic load variations on distributed optimization algorithms is considered, and the optimization problem is viewed as an adaptive filtering problem. Strictly distributed algorithms are developed to track the time-varying global optimal load distribution, in which only neighboring feeders are required to exchange information. Rigorous convergence analyses are obtained, and the relationship among algorithm step size, error bounds, and tracking ability are derived quantitatively. Moreover, optimal selection of the step size is established based on the limiting behavior of the optimality errors, which reflects the balance between attenuation of observation noises and tracking of the load changes. Next, simulation studies on real-world systems are provided to demonstrate the effectiveness of the proposed algorithms and validate the theoretical results.

There remains some interesting open problems in the direction of this paper. For examples, this paper does not consider communication channel interruptions or latency. Also, the methodology of this paper can be extended to more general AC power systems.

Appendix

Proof of *Lemma 3.1*:

Since M has full column rank and $R > 0, Q > 0$ are diagonal, we know that the matrix $M^\top RM+Q$ is full rank, and there exists two constants $c_1 = \lambda_{min}\{M^\top RM + Q\} > 0$ and $c_2 = \lambda_{\max}\{M^\top RM + Q\} > 0$ such that $c_1 I \le M^\top RM + Q \le c_2 I$. Thus, we have $(1-\mu c_2)I \le F_\mu \le (1-\mu c_1)I$.

If the step size μ is selected to satisfy $1-\mu c_1 < 1$ and $1-\mu c_2 > -1$, then $\|F_\mu^{k-i}\| \le (1-\mu c)^{k-i}$. Thus, it is obvious that we can choose $\mu^* = 2/c_2 > 0$ and c such that $1-\mu c = \max\{|1-\mu c_1|, |1-\mu c_2|\} \in [0,1)$. The desired result thus follows.

Proof of *Theorem 4.1*:

Using the error equation (10), it is easy to see that

$$\widetilde{x}(k+1) = F_\mu^{k+1}\widetilde{x}(0) + \sum_{i=0}^{k} F_\mu^{k-i}[\mu d(i) + H\delta(i+1)]. \tag{17}$$

Denote $b = \|H\|$. By (17), *Lemma 3.1*, and *Assumption 2.2*, we have $\|F_\mu^{k+1}\widetilde{x}(0)\|_{L_2} \le (1-\mu c)^{k+1}\|\widetilde{x}(0)\|$, and

$$\begin{aligned}
&\left\| \sum_{i=0}^{k} F_\mu^{k-i}[\mu d(i) + H\delta(i+1)] \right\| \\
&\le \mu \left\| \sum_{i=0}^{k} F_\mu^{k-i} d(i) \right\| + \left\| \sum_{i=0}^{k} F_\mu^{k-i} H\delta(i+1) \right\| \\
&\le \sigma\mu \sum_{i=0}^{k} (1-\mu c)^{k-i} + \delta b \sum_{i=0}^{k} (1-\mu c)^{k-i}.
\end{aligned}$$

It follows that

$$\|\widetilde{x}(k+1)\| \le (1-\mu c)^{k+1}\|\widetilde{x}(0)\| + \frac{1}{c}\left(\sigma + \frac{\delta b}{\mu}\right).$$

The proof is thus concluded.

Proof of *Theorem 4.2*:

By *Lemma A.2* in [42] and *Assumption 2.3*, $\left\| \sum_{i=0}^{k} F_\mu^{k-i} d(i) \right\| \le \frac{C_c C_d}{\sqrt{\mu}}$, and $\left\| \sum_{i=0}^{k} F_\mu^{k-i} H\delta(i+1) \right\| \le \frac{bC_cC_\delta}{\sqrt{\mu}}$, where C_c is a constant only depends on c, $b = \|H\|$, and C_d, C_δ are two constants defined in *Assumption 2.3*. Then, similar to the proof of *Theorem 4.1*, we know that (12) holds.

Proof of *Theorem 4.3*:

1) By (10), $\widetilde{x}(k+1) = F_\mu\widetilde{x}(k) + \mu d(k) + H\delta(k+1)$. By *Assumption 2.4*, we obtain $\eta_{\widetilde{x}}(k+1) = F_\mu\eta_{\widetilde{x}}(k) + H\eta_\delta$ which implies for $k \ge 1$

$$\eta_{\widetilde{x}}(k) = F_\mu^k \eta_{\widetilde{x}}(0) + \sum_{\ell=0}^{k-1} F_\mu^\ell H\eta_\delta. \tag{18}$$

2) From (18) and *Assumption 2.1*,

$$\begin{aligned}
\eta_{\widetilde{x}} &= \lim_{k\to\infty}\left(F_\mu^k\eta_{\widetilde{x}}(0) + \sum_{\ell=0}^{k-1} F_\mu^\ell H\eta_\delta\right) \\
&= \sum_{\ell=0}^{\infty} F_\mu^\ell H\eta_\delta
\end{aligned}$$

$$= (I - F_\mu)^{-1} H \eta_\delta$$
$$= \frac{1}{\mu}(M^\top Q M + R)^{-1} H \eta_\delta.$$

3) From (10),
$$\widetilde{x}(k+1) - \eta_{\widetilde{x}}$$
$$= F_\mu(\widetilde{x}(k) - \eta_{\widetilde{x}}) + (F_\mu - I)\eta_{\widetilde{x}} + \mu d(k) + H\delta(k+1)$$
$$= F_\mu(\widetilde{x}(k) - \eta_{\widetilde{x}}) + \mu d(k) + H(\delta(k+1) - \eta_\delta).$$
By *Assumption 2.4* and noting that F_μ is symmetric, $\Sigma_{\widetilde{x}}(k+1) = F_\mu \Sigma_{\widetilde{x}}(k) F_\mu + \mu^2\ P_d + H P_\delta H^\top$, which implies (13) holds.

4) By (13) and *Assumption 2.1*, F_μ is stable, and
$$\Sigma_{\widetilde{x}} = \lim_{k\to\infty} \Sigma_{\widetilde{x}}(k) = \sum_{\ell=0}^{\infty} F_\mu^\ell(\mu^2\ P_d + H P_\delta H^\top) F_\mu^\ell.$$
Then
$$F_\mu \Sigma_{\widetilde{x}} F_\mu = \sum_{\ell=1}^{\infty} F_\mu^\ell(\mu^2\ P_d + H P_\delta H^\top) F_\mu^\ell$$
$$= \Sigma_{\widetilde{x}} - (\mu^2\ P_d + H P_\delta H^\top),$$
or $F_\mu \Sigma_{\widetilde{x}} F_\mu - \Sigma_{\widetilde{x}} = -(\mu^2\ P_d + H P_\delta H^\top)$.

Proof of *Proposition 4.1*:

From (14) and $F_\mu = I - \mu G$, we have
$$E[\|\widetilde{x}\|^2]$$
$$= \mathrm{Tr}(\Sigma_{\widetilde{x}}) + \eta_{\widetilde{x}}^\top \eta_{\widetilde{x}}$$
$$= \sum_{\ell=0}^{\infty} \mathrm{Tr}(F_\mu^\ell(\mu^2\ P_d + H P_\delta H^\top) F_\mu^\ell) + \eta_{\widetilde{x}}^\top \eta_{\widetilde{x}}$$
$$= \sum_{\ell=0}^{\infty} \mathrm{Tr}((F_\mu^2)^\ell(\mu^2\ P_d + H P_\delta H^\top)) + \eta_{\widetilde{x}}^\top \eta_{\widetilde{x}}$$
$$= \mathrm{Tr}(\sum_{\ell=0}^{\infty} (F_\mu^2)^\ell(\mu^2\ P_d + H P_\delta H^\top)) + \eta_{\widetilde{x}}^\top \eta_{\widetilde{x}}$$
$$= \mathrm{Tr}((I - F_\mu^2)^{-1}(\mu^2\ P_d + H P_\delta H^\top)) + \eta_{\widetilde{x}}^\top \eta_{\widetilde{x}}$$
$$= \mathrm{Tr}((2\mu G - \mu^2 G^2)^{-1}(\mu^2\ P_d + H P_\delta H^\top)) + \frac{1}{\mu^2} G_1,$$
where $G_1 = \eta_\delta^\top H^\top G H \eta_\delta$. The optimal step size μ^o satisfies the stationary condition
$$0 = \frac{d E[\|\widetilde{x}\|^2]}{d\mu}$$
$$= \mathrm{Tr}\left(\frac{d(2\mu G - \mu^2\ G^2)^{-1}(\mu^2\ P_d + H P_\delta H^\top)}{d\mu}\right) - \frac{2}{\mu^3} G_1$$
$$= \mathrm{Tr}(-2(2\mu G - \mu^2\ G^2)^{-2}(G - \mu G^2)(\mu^2\ P_d + H P_\delta H^\top)$$
$$+2(2\mu G - \mu^2\ G^2)^{-1}\mu P_d) - \frac{2}{\mu^3} G_1,$$
which implies that
$$\mathrm{Tr}((2\mu G - \mu^2\ G^2)^{-2}(G - \mu G^2)(\mu^2\ P_d + H P_\delta H^\top)$$
$$-(2I - \mu G)^{-1} G^{-1} P_d) + \frac{1}{\mu^3} G_1$$
$$= \mathrm{Tr}(\frac{1}{\mu^2}(2I - \mu G)^{-2} G^{-1}(I - \mu G) H P_\delta H^\top$$
$$-(2I - \mu G)^{-2} G^{-1} P_d) + \frac{1}{\mu^3} G_1 = 0.$$
This completes the proof.

REFERENCES

[1] J. Miles and S. Potter, "Developing a viable electric bus service: The Milton Keynes demonstration project," *Res. Transp. Econ.*, vol. 48, pp. 357–363, Dec. 2014.

[2] H. He, H. Tang, and X. Wang, "Global optimal energy management strategy research for a plug-in series-parallel hybrid electric bus by using dynamic programming," *Math. Problems Eng.*, vol. 2013, pp. 1–11, Jan. 2013.

[3] S. Huang, L. He, Y. Gu, K. Wood, and S. Benjaafar, "Design of a mobile charging service for electric vehicles in an urban environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 787–798, Apr. 2015.

[4] Z. Wang, F. Liu, Y. Chen, S. H. Low, and S. Mei, "Unified distributed control of stand-alone DC microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 1013–1024, Jan. 2019.

[5] A. Khorsandi, M. Ashourloo, and H. Mokhtari, "A decentralized control method for a low-voltage DC microgrid," *IEEE Trans. Energy Convers.*, vol. 29, no. 4, pp. 793–801, Dec. 2014.

[6] M. Baranwal, A. Askarian, S. Salapaka, and M. Salapaka, "A distributed architecture for robust and optimal control of dc microgrids," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3082–3092, Jun. 2019.

[7] U. Vuyyuru, S. Maiti, and C. Chakraborty, "Active power flow control between DC microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5712–5723, Sep. 2019.

[8] V. Nasirian, S. Moayedi, A. Davoudi, and F. L. Lewis, "Distributed cooperative control of dc microgrids," *IEEE Trans. Power Electron.*, vol. 30, no. 4, pp. 2288–2303, Apr. 2015.

[9] H. Wang, M. Han, R. Han, J. M. Guerrero, and J. C. Vasquez, "A decentralized current-sharing controller endows fast transient response to parallel DC–DC converters," *IEEE Trans. Power Electron.*, vol. 33, no. 5, pp. 4362–4372, May 2018.

[10] O. Cornea, G.-D. Andreescu, N. Muntean, and D. Hulea, "Bidirectional power flow control in a DC microgrid through a switched-capacitor cell hybrid DC–DC converter," *IEEE Trans. Ind. Electron.*, vol. 64, no. 4, pp. 3012–3022, Apr. 2017.

[11] D. Chen and L. Xu, "Autonomous DC voltage control of a DC microgrid with multiple slack terminals," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1897–1905, Nov. 2012.

[12] C. Wang, J. Duan, B. Fan, Q. Yang, and W. Liu, "Decentralized high-performance control of DC microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3355–3363, May 2019.

[13] S. Trip, M. Cucuzzella, X. Cheng, and J. Scherpen, "Distributed averaging control for voltage regulation and current sharing in DC microgrids," *IEEE Control Syst. Lett.*, vol. 3, no. 1, pp. 174–179, Jan. 2019.

[14] X. Li *et al.*, "Observer-based DC voltage droop and current feed-forward control of a DC microgrid," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5207–5216, Sep. 2018.

[15] S. Abhinav, H. Modares, F. L. Lewis, and A. Davoudi, "Resilient cooperative control of DC microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 1083–1085, Jan. 2019.

[16] R. Han, L. Meng, J. M. Guerrero, and J. C. Vasquez, "Distributed nonlinear control with event-triggered communication to achieve current-sharing and voltage regulation in DC microgrids," *IEEE Trans. Power Electron.*, vol. 33, no. 7, pp. 6416–6433, Jul. 2018.

[17] M. Hosseinzadeh and F. R. Salmasi, "Power management of an isolated hybrid AC/DC micro-grid with fuzzy control of battery banks," *IET Renew. Power Gener.*, vol. 9, no. 5, pp. 484–493, Jul. 2015.

[18] T. Dragicevic, J. M. Guerrero, J. C. Vasquez, and D. Skrlec, "Supervisory control of an adaptive-droop regulated DC microgrid with battery management capability," *IEEE Trans. Power Electron.*, vol. 29, no. 2, pp. 695–706, Feb. 2014.

[19] D. Zhang, J. Jiang, L. Y. Wang, and W. Zhang, "Robust and scalable management of power networks in dual-source trolleybus systems: A consensus control framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 1029–1038, Apr. 2016.

[20] D. Zhang, L. Y. Wang, J. Jiang, and W. Zhang, "Optimal power management in DC microgrids with applications to dual-source trolleybus systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1188–1197, Apr. 2018.

[21] E. Sindi, L. Y. Wang, M. Polis, G. Yin, and L. Ding, "Distributed optimal power and voltage management in DC microgrids: Applications to dual-source trolleybus systems," *IEEE Trans. Transport. Electrific.*, vol. 4, no. 3, pp. 778–788, Sep. 2018.

[22] S. Grijalva and M. U. Tariq, "Prosumer-based smart grid architecture enables a flat, sustainable electricity industry," in *Proc. ISGT*, Jan. 2011, pp. 1–6.

[23] S. Karnouskos and P. J. Marrón, "Prosumer interactions for efficient energy management in smartgrid neighborhoods," in *Proc. 2nd Workshop eeBuildings Data Models, CIB Conf.*, Sophia Antipolis, France: European Commission, Oct. 2011, pp. 243–250.

[24] S. Grijalva, M. Costley, and N. Ainsworth, "Prosumer-based control architecture for the future electricity grid," in *Proc. IEEE Int. Conf. Control Appl. (CCA)*, Sep. 2011, pp. 43–48.

[25] C. W. Gellings, "Power to the people," *IEEE Power Energy Mag.*, vol. 9, no. 5, pp. 52–63, Aug. 2011.

[26] A. Keane and M. O'Malley, "Optimal allocation of embedded generation on distribution networks," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1640–1646, Aug. 2005.

[27] K. Clement-Nyns, E. Haesen, and J. Driesen, "The impact of charging plug-in hybrid electric vehicles on a residential distribution grid," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 371–380, Feb. 2010.

[28] P. Alluri, J. Solanki, and S. K. Solanki, "Charging coordination of plug-in electric vehicles based on the line flow limits and power losses," in *Proc. Int. Conf. Technol. Advancements Power Energy (TAP Energy)*, Jun. 2015, pp. 233–238.

[29] C. Wei, Z. M. Fadlullah, N. Kato, and I. Stojmenovic, "A novel distributed algorithm for power loss minimizing in smart grid," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 290–295.

[30] H. Lan, S. Wen, Q. Fu, D. C. Yu, and L. Zhang, "Modeling analysis and improvement of power loss in microgrid," *Math. Problems Eng.*, vol. 2015, pp. 1–8, Sep. 2015.

[31] J. A. P. Lopes, C. L. Moreira, and A. G. Madureira, "Defining control strategies for MicroGrids islanded operation," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 916–924, May 2006.

[32] R. H. Lasseter, "Smart distribution: Coupled microgrids," *Proc. IEEE*, vol. 99, no. 6, pp. 1074–1082, Jun. 2011.

[33] H. Xin, Z. Qu, J. Seuss, and A. Maknouninejad, "A self-organizing strategy for power flow control of photovoltaic generators in a distribution network," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1462–1473, Aug. 2011.

[34] A. Engler and N. Soultanis, "Droop control in LV-grids," in *Proc. Int. Conf. Future Power Syst.*, Nov. 2005, pp. 6–18.

[35] R. Tonkoski, L. A. C. Lopes, and T. H. M. El-Fouly, "Coordinated active power curtailment of grid connected PV inverters for overvoltage prevention," *IEEE Trans. Sustain. Energy*, vol. 2, no. 2, pp. 139–147, Apr. 2011.

[36] D. Pudjianto, C. Ramsay, and G. Strbac, "Microgrids and virtual power plants: Concepts to support the integration of distributed energy resources," *Proc. Inst. Mech. Eng., A, J. Power Energy*, vol. 222, no. 7, pp. 731–741, Nov. 2008.

[37] P. Tenti, H. K. M. Paredes, and P. Mattavelli, "Conservative power theory, a framework to approach control and accountability issues in smart microgrids," *IEEE Trans. Power Electron.*, vol. 26, no. 3, pp. 664–673, Mar. 2011.

[38] J. Driesen and F. Katiraei, "Design for distributed energy resources," *IEEE Power Energy Mag.*, vol. 6, no. 3, pp. 30–40, May 2008.

[39] L. Zhang, N. Gari, and L. V. Hmurcik, "Energy management in a microgrid with distributed energy resources," *Energy Convers. Manage.*, vol. 78, pp. 297–305, Feb. 2014.

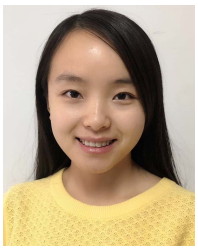
[40] A. L. Dimeas and N. D. Hatziargyriou, "Operation of a multiagent system for microgrid control," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1447–1455, Aug. 2005.

[41] L. Y. Wang, C. Wang, G. Yin, and Y. Wang, "Weighted and constrained consensus for distributed power dispatch of scalable microgrids," *Asian J. Control*, vol. 17, no. 5, pp. 1725–1741, Sep. 2015.

[42] L. Guo and L. Ljung, "Performance analysis of general tracking algorithms," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1388–1402, Aug. 1995.

[43] T. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York, NY, USA: Wiley, 1984.

[44] H. J. Kushner and J. Yang, "Analysis of adaptive step-size sa algorithms for parameter tracking," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1403–1410, Aug. 1995.



Siyu Xie received the B.S. degree in information and computing science (systems control) from the Beijing University of Aeronautics and Astronautics in 2013 and the Ph.D. degree in control theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, in 2018. She is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, Wayne State University, USA. Her research interests include networked systems, distributed optimization problems for power systems, distributed adaptive filters, machine learning, compressive sensing, and distributed control.

Masoud H. Nazari (Senior Member, IEEE) received the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA, in 2012. He was a Post-Doctoral Fellow with the School of Electrical and Computer Engineering, Georgia Institute of Technology, from 2013 to 2015. He was an Assistant Professor of electrical engineering with California State University, Long Beach, USA. He is currently an Assistant Professor of electrical and computer engineering with Wayne State University, Detroit, MI, USA. He is also the Primary Investigator of a $2.5 million California Energy Commission Project to develop an innovative building energy management system. He has more than 40 publications in the field. He is the IEEE Senior Member in the Power and Energy Society (PES). He was a recipient of the Best Paper Award in the 2017 North American Power Symposium.

Le Yi Wang (Life Fellow, IEEE) received the Ph.D. degree in electrical engineering from McGill University, Montreal, Canada, in 1990. Since 1990, he has been with Wayne State University, Detroit, MI, where he is currently a Professor with the Department of Electrical and Computer Engineering. His research interests include complexity and information, system identification, robust control, information processing and learning, as well as medical, automotive, communications, power systems, and computer applications of control methodologies. He was a plenary speaker in many international conferences. He serves on the IFAC Technical Committee on Modeling, Identification and Signal Processing. He was an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and several other journals.

George Yin (Life Fellow, IEEE) received the B.S. degree in mathematics from the University of Delaware in 1983 and the M.S. degree in electrical engineering and the Ph.D. degree in applied mathematics from Brown University in 1987. He joined Wayne State University in 1987, where he became a Professor in 1996 and the University Distinguished Professor in 2017. He moved to the University of Connecticut in 2020. His research interests include stochastic processes, stochastic systems theory, and applications. He is a fellow of IFAC and SIAM. He served as the Co-Chair for a number of conferences, was on the Board of Directors of the American Automatic Control Council. He was the Chair of the SIAM Activity Group on Control and Systems Theory. He is the Editor-in-Chief of *SIAM Journal on Control and Optimization*. He was an Associate Editor of *Automatica* from 1995 to 2011 and IEEE TRANSACTIONS ON AUTOMATIC CONTROL from 1994 to 1998 and a Senior Editor of IEEE CONTROL SYSTEMS LETTERS from 2017 to 2019.

Wen Chen (Senior Member, IEEE) received the Ph.D. degree from Simon Fraser University, BC, Canada, in 2004. He was a Post-Doctoral Researcher with the University of Louisiana at Lafayette, LA, USA, from 2005 to 2007. In 2007, he became a Control Systems Engineer at Paton Controls and was hired by Triconex, Houston, TX, USA, in 2008. He joined the Division of Engineering Technology, Wayne State University, USA, in 2009, as an Assistant Professor, where he is currently an Associate Professor. His research interest includes control and diagnosis of industrial systems.