Decentralized Dynamic Optimization for Power Network Voltage Control

Hao Jan Liu, *Student Member, IEEE*, Wei Shi, *Member, IEEE*, and Hao Zhu, *Member, IEEE*

Abstract—**Voltage control in power distribution networks has been greatly challenged by the increasing penetration of volatile and intermittent devices. These devices can also provide limited reactive power resources that can be used to regulate the network-wide voltage. A decentralized voltage control strategy can be designed by minimizing a quadratic voltage mismatch error objective using gradient-projection (GP) updates. Coupled with the power network flow, the local voltage can provide the instantaneous gradient information. This paper aims to analyze the performance of this decentralized GP-based voltage control design under two dynamic scenarios: First, the nodes perform the decentralized update in an asynchronous fashion. Second, the network operating condition is time varying. For the asynchronous voltage control, we improve the existing convergence condition by recognizing that the voltage-based gradient is always up-to-date. By modeling the network dynamics using an autoregressive process and considering time-varying resource constraints, we provide an error bound in tracking the instantaneous optimal solution to the quadratic error objective. This result can be extended to more general *constrained dynamic optimization* problems with smooth strongly convex objective functions under stochastic processes that have bounded iterative changes. Extensive numerical tests have been performed to demonstrate and validate our analytical results for realistic power networks.**

Index Terms—**Constrained dynamic optimization, decentralized control, droop control, reactive power support, voltage regulation.**

I. INTRODUCTION

THE smart grid vision has led to continued proliferation of distributed energy resources (DERs) in the power distribution networks, including the rooftop photovoltaic (PV) panels and batteries of electric vehicles. Albeit environmentally friendly, the DERs could greatly challenge the operational goal of maintaining a satisfactory voltage level per power system reliability standards. High variability of PV generation and elastic loads can cause unexpected network-wide voltage fluctuations, at much faster dynamics than the time-scale of traditional voltage regulation devices; see e.g., [1]–[3] and references therein.

Thanks to advances in power electronics, DER devices can also be excellent sources of *reactive power*, a quantity that is known to have a significant impact on the network voltage level. Thus, one major task for operating future distribution networks is to design effective voltage control strategies to utilize the reactive power from DERs, at minimal hardware requirements.

By and large, the communication infrastructure that supports power distribution networks is, and will continue to be lacking in the foreseen future. Meanwhile, cost concerns for DER products inevitably limit their sensing/computation capabilities. Therefore, voltage control approaches by minimizing a centralized voltage mismatch error are not practically attractive, because their success strongly relies on high-quality communications either between a control center and remote devices [4], or among neighboring devices [3], [5], [6]. These optimization-based open-loop control designs may become unstable under communication delays or noises during online implementations. Also, it is unclear whether these approaches will be robust to asynchronous computational speeds among the networked DERs of heterogeneous hardware capabilities. To cope with the limited cyber infrastructure, a decentralized voltage control framework is more preferred for distribution network operations; see e.g., [1], [2], [7], [8]. Under this framework, each node only needs to measure its locally available voltage level as the controller input. Our earlier work [8] has offered an overarching framework that generalizes a variety of decentralized control designs, along with convergence analysis for static system scenarios. Interestingly, the local voltage measurement naturally provides the instantaneous gradient information for a centralized error objective by weighting the voltage mismatch. Hence, the decentralized voltage control approach using this measurement boils down to the classical gradient-projection (GP) method accounting for the limits on reactive power resources. This voltage-measurement-based decentralized design does not require any real-time communications, and can be implemented with minimal upgrades in sensing hardware.

The goal of this paper is to analyze the performance of this decentralized GP-based voltage control design under two dynamic scenarios: i) the nodes perform the decentralized update in an asynchronous fashion; and ii) the network operating condition is dynamically changing. The scenario of asynchronous updates arises from heterogeneous hardware capabilities among different DERs. It is also motivated to allow the "plug-and-play" functionality for flexible DER integration to distribution networks. The classical asynchronous optimization framework always accounts for the case of outdated information from other nodes, and thus the choice of step-size has to be more

conservative due to the information delay; see e.g., [9] and [10]. Different from this, the voltage measurement for our decentralized control design always provides the up-to-date gradient information and does not suffer any information delay. Thanks to this physical power network coupling, we can show that the choice of the step-size for the asynchronous decentralized updates is the same to the synchronous case. Hence, its convergence condition is robust to potential discrepancy in the control update rates among different DERs.

In power networks, voltage control designs under time-varying operating conditions have been implemented as the static optimal power flow solutions to dynamic settings in a heuristic fashion [11], [12]. The scope of these efforts is more centered around dynamic voltage control implementations rather than providing control performance guarantees. The latter is of high interest when accounting for the variability of networked generations and loads in practice. With a time-varying objective function, this problem becomes a *stochastic optimization* one; see e.g., [13]. Stochastic approximation algorithms such as stochastic (sub-)gradient decent have been developed in e.g., [14], [15] and have been adopted by [16] for this voltage control problem. Nonetheless, performance analysis for stochastic optimization algorithms has focused on the convergence to the optimal solution that minimizes the expected objective function [13]. Aiming at the error bound in tracking the instantaneous optimal solution, our analysis is more closely related to the body of work on *dynamic convex optimization*; see e.g., [17]–[19]. This type of problems typically arises from applications in autonomous teams and wireless sensor networks such as target tracking [20] and estimation of the stochastic path [21]. Some of these dynamic optimization algorithms follow a gradient descent update, but none of them has considered the formulation of constrained optimization. This is the key difference from our voltage control problem since the control input has to be feasible under the dynamic reactive power limits. Hence, the main contribution of our work lies in the fact that it explicitly accounts for the time-varying projection operation of *constrained* dynamic optimization. Our tracking error performance bounds will be derived for a quadratic objective function and an autoregressive dynamic model, motivated by this specific voltage control problem. Nonetheless, our analytical results can be extended to more general constrained dynamic optimization problems with smooth strongly convex objective functions under stochastic processes that have bounded iterative changes.

The remainder of this paper is organized as follows. Section II presents the modeling of power distribution networks as the basis of our analysis, while Section III designs a decentralized voltage control strategy using the GP updates. Performance analysis under asynchronous updates or time-varying network operating condition is offered in Section IV and Section V, respectively. Section VI presents numerical results to demonstrate and validate our analytical results, and the paper is wrapped up in Section VII.

II. SYSTEM MODELING

A power distribution network can be modeled using a tree-topology graph $(\mathcal{N}, \mathcal{E})$ with the set of buses (nodes) $\mathcal{N} :=$
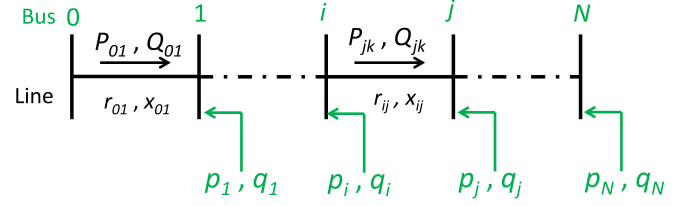

Fig. 1.   A radial distribution network with bus and line associated variables.

$\{0, ..., N\}$ and the set of line segments (edges) $\mathcal{E} := \{(i, j)\}$; see Fig. 1 for a radial network illustration. At every bus j, let v_j denote its voltage magnitude and p_j (q_j) represent the real (reactive) power injection, respectively. Bus 0 corresponds to the point of common coupling, assumed to have unity reference voltage; i.e., $v_0 = 1$. For each line (i, j), let r_{ij} and x_{ij} denote its resistance and reactance and P_{ij} and Q_{ij} the power flow from i to j, respectively.

To tackle the nonlinearity of power flow models, one can assume negligible line losses and almost flat voltage, i.e., $v_j \cong 1$, $\forall j$. Under these assumptions, the so-termed LinDistFlow model has been developed in [22], and its accuracy can be numerically corroborated by several recent work [7], [8], [23]–[25]. For each (i, j), the LinDistFlow model asserts the bus power balance and line voltage drop, as given by

$$P_{ij} - \sum_{k \in \mathcal{N}_j^+} P_{jk} = -p_j, \tag{1a}$$

$$Q_{ij} - \sum_{k \in \mathcal{N}_j^+} Q_{jk} = -q_j, \tag{1b}$$

$$v_i - v_j = r_{ij}P_{ij} + x_{ij}Q_{ij} \tag{1c}$$

where the neighboring bus set $\mathcal{N}_j^+ := \{k | (j, k) \in \mathcal{E}$, and k is downstream from $j\}$. For example, we have $\mathcal{N}_i^+ = \{j\}$ in Fig. 1.

To construct the matrix form of (1), denote the graph incidence matrix using the $(N + 1) \times N$ matrix $\mathbf{M}^o$. Each one of its ℓ-th column corresponds to a line (i, j), with all zero entries except for the i-th and j-th (see e.g., [26, pg. 6]). Let us set $M_{i\ell}^o = 1$ and $M_{j\ell}^o = -1$ if $j \in \mathcal{N}_i^+$. Let $\mathbf{m}_0^\top$ represent the first row of $\mathbf{M}^o$ corresponding to bus 0, with the rest of the rows in the $N \times N$ submatrix $\mathbf{M}$. Accordingly, $\mathbf{M}$ is full-rank and invertible because the network is connected under the tree-topology assumption [26]. Upon concatenating all scalar variables into vector form, one can represent (1) as

$$-\mathbf{MP} = -\mathbf{p}, \tag{2a}$$

$$-\mathbf{MQ} = -\mathbf{q}, \tag{2b}$$

$$\mathbf{m}_0 + \mathbf{M}^\top \mathbf{v} = \mathbf{D}_r \mathbf{P} + \mathbf{D}_x \mathbf{Q} \tag{2c}$$

where the $N \times N$ diagonal matrix $\mathbf{D}_r$ has diagonals equal to all line r_{ij}'s; and similarly for $\mathbf{D}_x$ having all x_{ij}'s. Viewing the uncontrollable $\mathbf{p}$ as a constant, one can solve for $\mathbf{P}$ and $\mathbf{Q}$ in (2) to establish the following:

$$\mathbf{v} = \mathbf{Xq} + \bar{\mathbf{v}} \tag{3}$$

where the nominal voltage vector $\bar{\mathbf{v}}$ captures the effects of $\mathbf{p}$ when $\mathbf{q} = \mathbf{0}$, while $\mathbf{X} := (\mathbf{M}^\top)^{-1} \mathbf{D}_x \mathbf{M}^{-1}$ can be viewed as the

network reactance matrix. The linear model (3) constitutes the basis for developing decentralized voltage control algorithms.

Remark 1 (Modeling Considerations): Nonlinearity of the power flow model could be tackled by the formulation of semidefinite programming (SDP) [27]. Generally, a rank relaxation approach is adopted in order to obtain a SDP convex problem formulation. Nonetheless, the SDP-based power flow formulation would face several challenges when applied to power (distribution) networks in practice. First, the resultant power flow solution could be non-exact [28] under this setting when power networks are unbalanced multi-phase. Furthermore, performance guarantees for the SDP approach fail to hold for general network topology such as meshed systems [29]. Last, the SDP solution increases the computational burden significantly as the size of the network grows. Compared to the SDP modeling approach, the LinDistFlow model (3) holds for more general scenarios including meshed topology and unbalanced three-phase systems while the resultant algorithms enjoy minimal computational complexity [8], [30]. Hence, we have adopted the LinDistFlow model for the ensuing analysis while the numerical tests in Sec. VI will be performed using exact power flow solvers on practical power networks.

III. DECENTRALIZED VOLTAGE CONTROL

The goal is to control the reactive power $\mathbf{q}$, such that $\mathbf{v}$ approaches a given desirable voltage profile $\boldsymbol{\mu}$. The flat voltage profile is typically chosen; i.e., $\boldsymbol{\mu} := \mathbf{1}$. At every time instance k, let $\overline{\mathbf{v}}_k$ denote the instantaneous nominal voltage profile. To allow for a decentralized control design, it turns out that one can minimize a weighted voltage mismatch error using $\mathbf{B} := \mathbf{X}^{-1}$, as given by

$$\mathbf{q}_k^* = \arg \min_{\mathbf{q} \in \mathcal{Q}_k} f_k(\mathbf{q}) := \frac{1}{2}(\mathbf{X}\mathbf{q} + \overline{\mathbf{v}}_k - \boldsymbol{\mu})^\top \mathbf{B}(\mathbf{X}\mathbf{q} + \overline{\mathbf{v}}_k - \boldsymbol{\mu})$$
$$(4)$$

where the constraint set $\mathcal{Q}_k := \{\mathbf{q} | \mathbf{q} \in [\underline{\mathbf{q}}_k, \overline{\mathbf{q}}_k]\}$ accounts for the time-varying limits of local reactive power resources at every bus [2]. Interestingly, matrix $\mathbf{B} = \mathbf{M}\mathbf{D}_x^{-1}\mathbf{M}^\top$, by definition, is a weighted, reduced graph Laplacian for $(\mathcal{N}, \mathcal{E})$. Since all the reactance values are positive, $\mathbf{B}$ is symmetric and positive definite [8]. Accordingly, the weighted voltage mismatch error objective of (4) is convex, and in fact quadratic, in the variable $\mathbf{q}$. Ideally, the unweighted error norm $\|\mathbf{v} - \boldsymbol{\mu}\|$ is the best objective in order to achieve the flat voltage profile. Compared to the traditional paradigm of maintaining the voltage within limits, this unweighted objective can improve the system-wide voltage profile by coordinating network-wide VAR resources. This is attractive for energy saving programs such as a conservation voltage reduction implementation [31]. Albeit the problem (4) minimizes a surrogate objective, it has been shown in [8] that $\mathbf{q}_k^*$ can closely approximate the optimal solution to the ideal unweighted error norm, especially if there are abundant reactive power resources. Last, it is possible to use other convex error penalty functions such as the Huber's loss function [32] instead of the squared error norm objective in (4). This approach generalizes the current design to allow for some tolerance in the voltage mismatch error, which may be more attractive under the scenarios of limited VAR resources.

Thanks to the separable structure of the box constraint $\mathcal{Q}_k$, the gradient-projection (GP) method [33, Sec. 2.3] can be invoked to solve (4). Upon forming its instantaneous gradient $\nabla f_k(\mathbf{q}_k) := \mathbf{X}\mathbf{q}_k + \overline{\mathbf{v}}_k - \boldsymbol{\mu}$, the GP iteration for a given positive step-size $\epsilon > 0$ becomes

$$\mathbf{q}_{k+1} = \mathbb{P}_k \left[\mathbf{q}_k - \epsilon \mathbf{D} \nabla f_k(\mathbf{q}_k) \right] \qquad (5)$$

where the projection operator $\mathbb{P}_k$ thresholds any input to be within $\mathcal{Q}_k$, and $\mathbf{D}$ is a diagonal scaling matrix that can be designed. As a first-order method, the GP method has a linear convergence rate, while the convergence speed depends on the condition number of the corresponding Hessian matrix [33, Sec. 3.3]. Motivated by this fact, the scaling matrix can be chosen according to the inverse of the diagonals of Hessian matrix by setting $\mathbf{D} := [\mathrm{diag}(\mathbf{X})]^{-1}$ to approximate the Newton gradient. Note that a positive diagonal matrix $\mathbf{D}$ affects neither the separability of operator $\mathbb{P}$, nor the optimality of the update (5).

By setting the GP iterate $\mathbf{q}_k \in \mathcal{Q}_k$ to be the control input at any time k, the instantaneous voltage becomes $\mathbf{v}_k = \mathbf{X}\mathbf{q}_k + \overline{\mathbf{v}}_k$ based on (3). Thanks to the physical power network coupling, $\mathbf{v}_k$ always provides the up-to-date gradient information as $\nabla f_k(\mathbf{q}_k) = \mathbf{v}_k - \boldsymbol{\mu}$. Accordingly, the GP update in (5) can be implemented by directly measuring the instantaneous voltage as

$$\mathbf{q}_{k+1} = \mathbb{P}_k \left[\mathbf{q}_k - \epsilon \mathbf{D}(\mathbf{v}_k - \boldsymbol{\mu}) \right], \qquad (6)$$

which can be completely decoupled into decentralized updates at each bus because $\mathbb{P}_k$ is separable. This decentralized voltage control is very attractive with minimal hardware requirements as each bus only needs to measure its local voltage and requires no communication. The optimality and convergence conditions for (6) have been investigated in [8], which are summarized by the following proposition.

Proposition 1: When $\overline{\mathbf{v}}_k = \overline{\mathbf{v}}$ and $\mathcal{Q}_k = \mathcal{Q}$ (time-invariant case), the decentralized update (6) approaches the unique time-invariant optimizer $\mathbf{q}^*$ of problem (4) if the step-size $\epsilon \in (0, 2/M)$ where

$$M := \lambda_{\max}\{\widetilde{\mathbf{X}}\} \qquad (7)$$

is the largest eigenvalue of matrix

$$\widetilde{\mathbf{X}} := \mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}. \qquad (8)$$

The decentralized voltage control design has to account for a variety of uncertainties in practical system implementations. First, due to heterogeneity of various DERs, it is difficult to perfectly synchronize the decentralized update at different buses. This is especially important to facilitate the "plug-and-play" functionality for flexible DER integration. Second, the volatility and intermittence of electric loads and renewable-based generations challenge the static setting of time-invariant $\overline{\mathbf{v}}_k$. It is of considerable interest to quantify the performance of the decentralized control (6) in terms of tracking the time-varying optimizer $\mathbf{q}_k^*$ to the dynamic objective f_k.

IV. ASYNCHRONOUS DECENTRALIZED VOLTAGE CONTROL

When the DERs are heterogeneous, it is logical that the decentralized voltage update should be performed in an asynchronous

fashion. This way, the buses with better computation and sensing capabilities do not need to wait for the slowest one. Accordingly, these buses can execute more updates for a same time interval and hence respond more quickly to localized voltage deviations.

To this end, let the set $\mathcal{K}_j$ collects all the time instances when bus j executes its decentralized update. The asynchronous counterpart of (6) can be modeled by

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{s}_k, \ \forall k \tag{9}$$

with the difference at bus j given by

$$s_{j,k} := \begin{cases} \mathbb{P}_{j,k}\left[q_{j,k} - \epsilon D_j(v_{j,k} - \mu_j)\right] - q_{j,k}, & k \in \mathcal{K}_j, \\ 0, & k \notin \mathcal{K}_j, \end{cases} \tag{10}$$

where D_j is the (j,j)−th entry of $\mathbf{D}$, while $\mathbb{P}_{j,k}$ projects the input to $[\underline{q}_{j,k}, \overline{q}_{j,k}]$.

To establish the convergence condition, every bus needs to update sufficiently often. Similar to the classical asynchronous algorithm analysis in [9, Ch. 7], we assume the following *bounded update delay* condition.

Assumption 1 (Bounded Update Delay): For every bus j and time instance $k \geq 0$, there exists a positive integer K such that at least one element in the set $\{k, k+1, \ldots, k+K-1\}$ belongs to $\mathcal{K}_j$. Equivalently, every bus must update at least once every K iterations.

In addition to Assumption 1, the analysis of classical asynchronous algorithms also assumes the *bounded information delay* condition [9, Ch. 7]. Due to potential communication delays among peer processors, the updates at some buses may not be executed based on the most up-to-date system-wide information. By assuming that the local information used for computing the gradient is potentially obsoleted by at most K iterations, it has been established in [9, Sec. 7.5] that the asynchronous GP algorithm in (9)-(10) converges to the optimal solution with more conservative step-size choice given by

$$0 < \epsilon < 1/[M(1 + K + NK)]. \tag{11}$$

Due to information delay, the choice of step-size would depend on the slowest processor in the network. Generally, this bound on ϵ can be much smaller than the $2/M$ bound in Proposition 1, resulting in a much slower convergence compared with the synchronous case.

For our decentralized voltage control, the gradient $\nabla f_k(\mathbf{q}_k) = \mathbf{v}_k - \boldsymbol{\mu}$ always holds. Thanks to the physical power network coupling, the local voltage $v_{j,k}$ always provides the up-to-date gradient information at every iteration k. Hence, whenever a node is active, the difference $s_{j,k}$ computed in (10) does not suffer from any information delay. This is different from most parallel and distributed algorithms where the updates at every processor require information sent by peer processors. Therefore, convergence of (9)-(10) no longer requires the more conservative choice of ϵ in [9, Sec. 7.5].

Theorem 1: Under Assumption 1, when $\overline{\mathbf{v}}_k = \overline{\mathbf{v}}$ and $\mathcal{Q}_k = \mathcal{Q}$ (time-invariant case), the asynchronous update illustrated in (9)-(10) converges to the time-invariant optimizer $\mathbf{q}^*$ if the step-size $\epsilon \in (0, 2/M)$.

Proof 1: First, it is easy to show that the fixed-point of (9)-(10) is the same to (6) using contradiction. As for the convergence, by projecting any scalar q to $[\underline{q}_j, \overline{q}_j]$, it holds that $[\mathbb{P}_j(q) - q_{j,k}][\mathbb{P}_j(q) - q] \leq 0$ for the iterate $q_{j,k} \in [\underline{q}_{j,k}, \overline{q}_{j,k}]$ where $\mathbb{P}_j = \mathbb{P}_{j,k}$; see e.g., [9, Sec. 3.3.1]. This implies that at every iteration $k \in \mathcal{K}_j$ (use $f_k = f$ for time-invariant case)

$$s_{j,k}[s_{j,k} + \epsilon D_j \nabla_j f(\mathbf{q}_k)] = \epsilon D_j s_{j,k} \nabla_j f(\mathbf{q}_k) + (s_{j,k})^2 \leq 0. \tag{12}$$

The descent lemma in [9, Sec. 3.3.2] together with the Lipschitz continuity of $f(\cdot)$ entails for every k

$$f(\mathbf{q}_{k+1}) = f(\mathbf{q}_k + \mathbf{s}_k)$$
$$\leq f(\mathbf{q}_k) + \mathbf{s}_k^\top \nabla f(\mathbf{q}_k) + (M/2)\|\mathbf{s}_k\|^2$$
$$\leq f(\mathbf{q}_k) - \left(\frac{1}{\epsilon} - \frac{M}{2}\right)\|\mathbf{s}_k\|^2. \ [\text{cf. (12)}]$$

Summing up the inequality over all iterations yields

$$\sum_{k=0}^{\infty} \|\mathbf{s}_k\|^2 \leq \left(\frac{1}{\epsilon} - \frac{M}{2}\right)^{-1} f(\mathbf{q}_0) < \infty,$$

which holds as long as $\left(\frac{1}{\epsilon} - \frac{M}{2}\right)$ is positive. Thus, if $0 < \epsilon < 2/M$, $\|\mathbf{s}_k\|^2$ is summable and the convergence $\lim_{k\to\infty} s_{j,k} = 0$ holds for every j. And this completes the asymptotic convergence claim for q_k to its fixed point $\mathbf{q}^*$. ∎

Convergence analysis for the asynchronous voltage control updates ensures that the heterogeneity of DERs does not affect the choice of ϵ. As for online implementation, this result can provide guaranteed stability for the proposed control design.

V. DYNAMIC DECENTRALIZED VOLTAGE CONTROL

In addition to the asynchronous voltage control updates, the uncertainty in the nominal voltage $\overline{\mathbf{v}}_k$ further challenges the performance of decentralized voltage control. The volatility and intermittence of loads and generations lead to temporal variations in the network operating condition, i.e., a dynamic $\overline{\mathbf{v}}_k$. Thus, it is imperative to analyze the performance of the decentralized voltage control under a dynamic setting.

To this end, the first order autoregressive (AR(1)) process is adapted to model the short-term dynamics.

Assumption 2 (Dynamic Voltage Profile): For a given constant vector $\overline{\mathbf{c}}$, the nominal voltage $\overline{\mathbf{v}}_k$ follows a wide-sense stationary AR(1) process, as given by

$$\overline{\mathbf{v}}_{k+1} = \mathbf{A}\overline{\mathbf{v}}_k + \boldsymbol{\eta}_{k+1} + \overline{\mathbf{c}} \tag{13}$$

where $\mathbf{A}$ is a time-invariant transition matrix with its spectral radius less than 1, while $\boldsymbol{\eta}_{k+1}$ represents a zero-mean white noise process with covariance matrix $\boldsymbol{\Sigma}_\eta$.

The AR(1) model (13) can capture both a short-term temporal and spatial correlation of the nominal voltage profile. Its validity has been corroborated by [34] from real data-based tests. Under Assumption 2, for every time k, the nominal voltage $\overline{\mathbf{v}}_k$ has constant mean $\mathbb{E}\overline{\mathbf{v}}_k = (\mathbf{I} - \mathbf{A})^{-1}\overline{\mathbf{c}}$ with a bounded covariance matrix $\boldsymbol{\Sigma}_{\overline{v}}$ satisfying

$$\boldsymbol{\Sigma}_{\overline{v}} = \mathbf{A}\boldsymbol{\Sigma}_{\overline{v}}\mathbf{A}^\top + \boldsymbol{\Sigma}_\eta.$$

For ease of exposition, the spatial correlation for power networks is not considered as it is often time negligible [35], corroborated by the structure of inverse of reduced graph Laplacian matrix $\mathbf{X}$. Since $\mathbf{X}$ is in fact diagonally dominant, the variations in loads and generations tend to have very localized impacts. In addition, we assume an equal variation level in the temporal dynamics across the network; i.e., $\mathbf{A} = \alpha \mathbf{I}$. However, the ensuing analysis holds for the original AR modeling in (13) or even higher-order AR modeling as long as it has bounded successive differences. These assumptions simplify the AR(1) model as follows

$$\overline{\mathbf{v}}_{k+1} = \alpha \overline{\mathbf{v}}_k + \boldsymbol{\eta}_{k+1} + \overline{\mathbf{c}} \tag{14}$$

with $\boldsymbol{\Sigma}_\eta = \sigma^2 \mathbf{I}$. Accordingly, the stability condition boils down to $|\alpha| < 1$, while $\overline{\mathbf{v}}_k$ the mean $\mathbb{E}\overline{\mathbf{v}}_k = \overline{\mathbf{c}}/(1 - \alpha)$ and the covariance $\boldsymbol{\Sigma}_{\overline{v}} = \sigma^2/(1 - \alpha^2)\mathbf{I}$. The smaller the value of $|\alpha|$ is, the faster that the nominal voltage $\overline{\mathbf{v}}_k$ evolves.

Proposition 2 (Lemma 1 in [36]): Under Assumption 2, the expectation of the weighted norm of consecutive difference is bounded, i.e., there exists a bounded constant B_1 such that

$$\mathbb{E}\|\overline{\mathbf{v}}_{k+1} - \overline{\mathbf{v}}_k\|_{\mathbf{D}}^2 = \frac{2\sigma^2 \text{Tr}\mathbf{D}}{1 + \alpha} \le B_1 \text{ for all } k = 0, 1, \dots . \tag{15}$$

where the weighted norm $\|\mathbf{v}\|_{\mathbf{D}}^2 := \mathbf{v}^\top \mathbf{D} \mathbf{v}$ for any $\mathbf{v}$.

Under the settings of both dynamic objective and constraint, we formally state the assumption we make for the performance analysis of the gradient projection approach (6).

Assumption 3 (Bounded Drift of Optimizer): The successive difference of the transient optimal solution is bounded, i.e., there exists some bounded constant B_2 such that (see (4) for the definition of $\mathbf{q}_k^*$)

$$\mathbb{E}\|\mathbf{q}_k^* - \mathbf{q}_{k+1}^*\|_{\mathbf{D}^{-1}}^2 \le B_2 \text{ for all } k = 0, 1, \dots .$$

This assumption is related to the boundedness of voltage drift (see Prop. 2) and the compactness of box constraints. For instance, when the reactive power is unlimited, i.e., $\mathcal{Q}_k = \mathbb{R}^N$, one can easily verify that $\mathbb{E}\|\mathbf{q}_k^* - \mathbf{q}_{k+1}^*\|_{\mathbf{D}^{-1}}^2$ is bounded. When the reactive power is uniformly limited, i.e., $\mathcal{Q}_k$ is always some compact set (double-sided box constraint suffices) for all k, we still have the bounded optimizer drift due to the fact that $\mathbf{q}_k^* \in \mathcal{Q}_k$. Albeit the error bound (stability) we are going to construct will depend on B_2, intuitively, a smaller voltage drift bound B_1 tends to decrease the drift of the optimizer bound B_2 in power networks.

We first introduce a few quantities to simplify the presentation:

$$\mathbf{y}_k := \mathbf{D}^{-\frac{1}{2}}\mathbf{q}_k, \ \mathbf{y}_k^* := \mathbf{D}^{-\frac{1}{2}}\mathbf{q}_k^*,$$

$$\mathbf{u}_k := \mathbf{D}^{-\frac{1}{2}}(\overline{\mathbf{v}}_k - \boldsymbol{\mu}),$$

and $\widetilde{\mathbb{P}}_k[\cdot]$ is an operator that projects its input onto the set

$$\widetilde{\mathcal{Q}}_k := \{\mathbf{q} | \mathbf{q} \in [\mathbf{D}^{-\frac{1}{2}}\underline{\mathbf{q}}_k, \mathbf{D}^{-\frac{1}{2}}\overline{\mathbf{q}}_k]\}.$$

This way, the original iterative update in (6) is equivalent to

$$\mathbf{y}_{k+1} = \widetilde{\mathbb{P}}_k[\mathbf{y}_k - \epsilon(\widetilde{\mathbf{X}}\mathbf{y}_k + \mathbf{u}_k)], \tag{16}$$

which can be viewed as the standard gradient projection update for the following dynamic constrained optimization problem

$$\min_{\mathbf{y} \in \widetilde{\mathcal{Q}}_k} \widetilde{f}_k(\mathbf{y}) := \frac{1}{2}\|\mathbf{P}\mathbf{y} + (\mathbf{P}^\top)^{-1}\mathbf{u}_k\|^2 \tag{17}$$

where $\mathbf{P}$ is obtained by the Cholesky factorization for the symmetric PD matrix $\widetilde{\mathbf{X}} = \mathbf{P}^\top \mathbf{P}$. Correspondingly, $\mathbf{y}_k^*$ is the optimizer of (17). Also, we denote

$$C := \min_{k \ge 0; \mathbf{y} \in \mathbb{R}^N} \left\{ \lambda_{\min}\{\nabla^2 \widetilde{f}_k(\mathbf{y})\} \right\} = \lambda_{\min}\{\widetilde{\mathbf{X}}\}$$

which is the smallest eigenvalue of $\widetilde{\mathbf{X}}$. We say a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex with some positive constant c if for any x and y, we have $f(y) \ge f(x) + \langle \nabla f(x), y - x \rangle + \frac{c}{2}\|x\|^2$ where $\langle \cdot, \cdot \rangle$ is the inner product of two vectors. A differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is gradient Lipschitz continuous with some positive constant m if for any x and y, we have $f(y) \le f(x) + \langle \nabla f(x), y - x \rangle + \frac{m}{2}\|x\|^2$. Note that, under this definition, C also serves as the least strong convexity constant of $\widetilde{f}_k(\mathbf{y})$, $\forall k$, while M also serves as the greatest gradient Lipschitz continuity constant of $\widetilde{f}_k(\mathbf{y})$, $\forall k$. A vector v is called a subgradient of a convex function $f : \mathcal{X} \to \mathbb{R} \bigcup \{+\infty\}$ at point $x \in \mathcal{X}$ if $f(y) \ge f(x) + \langle v, y - x \rangle$ for any $y \in \mathcal{X}$. The set of all subgradients at x is called the subdifferential at x. We use $\widetilde{\nabla} \widetilde{f}_k(\mathbf{y})$ and $\partial \widetilde{f}_k(\mathbf{y})$ to denote a subgradient and the subdifferential of the function $\widetilde{f}_k$ at $\mathbf{y}$, respectively. These notations were also used in [37]. The subgradient used in the algorithm or analysis will be specified in the context, and our analysis will be based on the equivalent update and optimization problem in (16) and (17), respectively. The following lemma gives the first-order optimality condition of (17) and an equivalent recursive relation of (16).

Lemma 1 (First-Order Optimality Condition and Recursive Relation): The instantaneous optimizer $\mathbf{y}_k^*$ to the dynamic optimization problem (17) and iterates $\mathbf{y}_k$ satisfy the following conditions, for $k = 0, 1, \dots,$

$$\widetilde{\mathbf{X}}\mathbf{y}_k^* + \mathbf{u}_k + \widetilde{\nabla} g_k(\mathbf{y}_k^*) = 0 \tag{18}$$

and

$$\mathbf{y}_{k+1} - \mathbf{y}_k = -\epsilon \left[\widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) + \widetilde{\nabla} g_k(\mathbf{y}_{k+1}) - \widetilde{\nabla} g_k(\mathbf{y}_k^*) \right] \tag{19}$$

where

$$g_k(\mathbf{y}) = \begin{cases} 0, & \text{if } \mathbf{y} \in \widetilde{\mathcal{Q}}_k, \\ +\infty, & \text{if } \mathbf{y} \notin \widetilde{\mathcal{Q}}_k \end{cases}$$

is the indicator function of the set $\widetilde{\mathcal{Q}}_k$ at time k.

Proof 2: We first replace the projection operation in (16) by a subgradient step featured by the indicator function $g_k(\cdot)$. By definition, the projection of any $\boldsymbol{\omega}$ to $\widetilde{\mathcal{Q}}_k$ equals to

$$\widetilde{\mathbb{P}}_k(\boldsymbol{\omega}) = \arg\min_{\mathbf{x}} \epsilon g_k(\mathbf{x}) + \frac{1}{2}\|\mathbf{x} - \boldsymbol{\omega}\|^2. \tag{20}$$

The first-order optimality condition leads to $\epsilon \widetilde{\nabla} g_k(\widetilde{\mathbb{P}}_k(\boldsymbol{\omega})) + \widetilde{\mathbb{P}}_k(\boldsymbol{\omega}) - \boldsymbol{\omega} = 0$. Thus by letting $\boldsymbol{\omega} = \mathbf{y}_k - \epsilon(\widetilde{\mathbf{X}}\mathbf{y}_k + \mathbf{u}_k)$ and

$\widetilde{\mathbb{P}}_k(\boldsymbol{\omega}) = \mathbf{y}_{k+1}$, we obtain

$$\mathbf{y}_{k+1} = \mathbf{y}_k - \epsilon[\widetilde{\mathbf{X}}\mathbf{y}_k + \mathbf{u}_k + \widetilde{\nabla} g_k(\mathbf{y}_{k+1})]. \qquad (21)$$

Furthermore, by using the indicator function, (17) is equivalent to

$$\mathbf{y}_k^* = \arg\min_{\mathbf{y}} \frac{1}{2}\|\mathbf{P}\mathbf{y} + (\mathbf{P}^\top)^{-1}\mathbf{u}_k\|^2 + g_k(\mathbf{y}). \qquad (22)$$

From the first-order optimality condition of (22), we have $\widetilde{\mathbf{X}}\mathbf{y}_k^* + \mathbf{u}_k + \widetilde{\nabla} g_k(\mathbf{y}_k^*) = 0$. This along with (21) proves the recursive relation (19). ∎

Note that the subgradient $\widetilde{\nabla} g_k(\mathbf{y}_{k+1})$ used in (21) is well-defined because (i) $\mathbf{y}^{k+1} \in \widetilde{\mathcal{Q}}_k$, (ii) $g_k(\cdot)$ is continuous over $\widetilde{\mathcal{Q}}_k$, and (iii) the minimum in (20) is uniquely attained since $\|\mathbf{x} - \boldsymbol{\omega}\|^2$ is real-valued, strictly convex, and coercive. Using the aforementioned notation, our analysis coincides with those earlier results on nonsmooth optimization; see e.g., similar notions and analysis schemes have appeared in [37]–[39] and references therein. Our main result is as follows:

Theorem 2 (BIBO Stability with Geometric Decaying): Under Assumption 3, for any step-size choice

$$\epsilon \in \left(0, \frac{2}{C+M}\right],$$

the expectation of the weighted tracking error between the decentralized control update $\mathbf{q}_k$ of (6) and the instantaneous optimal solution $\mathbf{q}_k^*$ can be bounded by

$$\mathbb{E}\|\mathbf{q}_k - \mathbf{q}_k^*\|_{\mathbf{D}^{-1}}^2 \leq \rho^k \mathbb{E}\|\mathbf{q}_0 - \mathbf{q}_0^*\|_{\mathbf{D}^{-1}}^2 + \frac{1-\rho^k}{1-\rho}\Theta, \ \forall k \quad (23)$$

where the geometric rate $\rho \in (0,1)$ and Θ is a bounded positive constant gap.

Proof 3: By the smoothness and convexity of $\widetilde{f}_k$, it follows that [40] (see (7) for the definition of M)

$$\frac{CM}{C+M}\|\mathbf{y}_k - \mathbf{y}_k^*\|^2 + \frac{1}{C+M}\|\widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\|^2$$
$$\leq \langle \mathbf{y}_k - \mathbf{y}_k^*, \widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\rangle. \qquad (24)$$

By applying the basic inequality

$$2\langle\sqrt{\beta}\mathbf{a}, \frac{1}{\sqrt{\beta}}\mathbf{b}\rangle \leq a\|\mathbf{a}\|^2 + \frac{1}{a}\|\mathbf{b}\|^2$$

which holds for any $\beta > 0$ and any real vectors $\mathbf{a}$ and $\mathbf{b}$ of the same dimension, the right-hand side of (24) can be upper bounded by

$$\langle \mathbf{y}_k - \mathbf{y}_{k+1} + \mathbf{y}_{k+1} - \mathbf{y}_k^*, \widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\rangle$$
$$\leq \frac{C+M}{4}\|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2 + \frac{1}{C+M}\|\widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\|^2$$
$$+ \langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\rangle. \qquad (25)$$

Substituting (25) into (24) leads to

$$\frac{CM}{C+M}\|\mathbf{y}_k - \mathbf{y}_k^*\|^2 \leq \frac{C+M}{4}\|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2$$
$$+ \langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\rangle. \qquad (26)$$

Since the indicator function $g_k(\cdot)$ is convex due to the fact that $\widetilde{\mathcal{Q}}_k$ is a convex set in our settings, its subgradient $\widetilde{\nabla} g_k(\cdot)$ (subdifferential $\partial g_k(\cdot)$) is a (set-valued) monotone mapping (this can also be obtained from the subgradient inequality [39]), i.e.,

$$\langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \widetilde{\nabla} g_k(\mathbf{y}_{k+1}) - \widetilde{\nabla} g_k(\mathbf{y}_k^*)\rangle \geq 0 \qquad (27)$$

Combining (26) and (27) we have

$$\frac{C+M}{4}\|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2 + \langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)$$
$$+ \widetilde{\nabla} g_k(\mathbf{y}_{k+1}) - \widetilde{\nabla} g_k(\mathbf{y}_k^*)\rangle \geq \frac{CM}{C+M}\|\mathbf{y}_k - \mathbf{y}_k^*\|^2. \quad (28)$$

Substituting (19) into (28) for $\widetilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) + \widetilde{\nabla} g_k(\mathbf{y}_{k+1}) - \widetilde{\nabla} g_k(\mathbf{y}_k^*)$ leads to

$$\frac{C+M}{4}\|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2 + \frac{1}{\epsilon}\langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \mathbf{y}_k - \mathbf{y}_{k+1}\rangle$$
$$\geq \frac{CM}{C+M}\|\mathbf{y}_k - \mathbf{y}_k^*\|^2.$$

Using the equality $\langle \mathbf{y}_k^* - \mathbf{y}_{k+1}, \mathbf{y}_{k+1} - \mathbf{y}_k\rangle = \|\mathbf{y}_k^* - \mathbf{y}_k\|^2 - \|\mathbf{y}_k^* - \mathbf{y}_{k+1}\|^2 - \|\mathbf{y}_{k+1} - \mathbf{y}_k\|^2$ to expand the inner product, we have

$$\|\mathbf{y}_{k+1} - \mathbf{y}_k^*\|^2 \leq \left(1 - \frac{2\epsilon CM}{C+M}\right)\|\mathbf{y}_k - \mathbf{y}_k^*\|^2$$
$$+ \left(\frac{\epsilon M + \epsilon L}{2} - 1\right)\|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2. \qquad (29)$$

By choosing $\epsilon \leq \frac{2}{C+M}$ to ensure the second term on the right-hand-side of (29) being nonnegative, the inequality (29) can be further relaxed to

$$\|\mathbf{y}_{k+1} - \mathbf{y}_k^*\|^2 \leq \left(1 - \frac{2\epsilon CM}{C+M}\right)\|\mathbf{y}_k - \mathbf{y}_k^*\|^2. \qquad (30)$$

By applying another basic inequality

$$\|\mathbf{a}' + \mathbf{b}'\|^2 \leq (1+\beta')\|\mathbf{a}'\|^2 + (1+\frac{1}{\beta'})\|\mathbf{b}'\|^2$$

which holds for any $\beta' > 0$ and any real vectors $\mathbf{a}'$ and $\mathbf{b}'$ of the same dimension, we have

$$\|\mathbf{y}_{k+1} - \mathbf{y}_k^* + \mathbf{y}_k^* - \mathbf{y}_{k+1}^*\|^2 \leq (1+\beta')\|\mathbf{y}_{k+1} - \mathbf{y}_k^*\|^2$$
$$+ \left(1+\frac{1}{\beta'}\right)\|\mathbf{y}_k^* - \mathbf{y}_{k+1}^*\|^2 \leq \rho\|\mathbf{y}_k - \mathbf{y}_k^*\|^2$$
$$+ \left(1+\frac{1}{\beta'}\right)\|\mathbf{y}_k^* - \mathbf{y}_{k+1}^*\|^2 \qquad (31)$$

where $\rho := (1+\beta')(1 - \frac{2\epsilon CM}{C+M})$ while the second inequality comes from (30). Let us denote $\Theta := (1+\frac{1}{\beta'})B_2$, and thus taking expectation on both sides of (28) gives us [cf. Assumption 3]

$$\mathbb{E}\|\mathbf{y}_{k+1} - \mathbf{y}_{k+1}^*\|^2 \leq \rho\mathbb{E}\|\mathbf{y}_k - \mathbf{y}_k^*\|^2 + \Theta. \qquad (32)$$

Applying recursive induction on (32), we eventually obtain

$$\mathbb{E}\|\mathbf{y}_{k+1} - \mathbf{y}_{k+1}^*\|^2 \leq \rho^{k+1}\mathbb{E}\|\mathbf{y}_0 - \mathbf{y}_0^*\|^2 + \frac{1-\rho^{k+1}}{1-\rho}\Theta,$$

which recovers (23) by the definition of $\mathbf{y}_{k+1}$ and $\mathbf{y}_{k+1}^*$. It shows that as long as $\rho \in (0,1)$ and $\Theta \in [0,\infty)$, $\mathbb{E}\|\mathbf{y}_{k+1} - \mathbf{y}_{k+1}^*\|^2$ is bounded for all k. Note that the choice of β' can be arbitrary close to 0. Hence ρ can always achieve a value that is less than 1

as long as $\epsilon > 0$ and $C > 0$ (a simple choice to demonstrate this is $\beta' = \epsilon CM/(C + M - 2\epsilon CM)$). Finally, we conclude that the step-size condition is $0 < \epsilon \leq \frac{2}{C+M}$. ∎

Theorem 2 establishes that the tracking error of the decentralized control update (6) under dynamical settings exponentially decreases until a constant error bound is reached. Moreover, the AR(1) process assumed to model the $\overline{\mathbf{v}}_k$ series can be potentially extended to a general stochastic process that has bounded iterative changes. This is because the constant Θ in (23) is bounded as long as the condition in (15) holds. To extract a more specific result, let us choose $\beta' = \epsilon CM/(C + M - 2\epsilon CM)$. In this case, the steady-state ($k \to \infty$) error is explicitly bounded by

$$\lim_{k \to \infty} \frac{1 - \rho^k}{1 - \rho} \Theta = \frac{1}{1 - (1 + \beta')\left(1 - \frac{2\epsilon CM}{C+M}\right)} \left(1 + \frac{1}{\beta'}\right) B_2$$

$$= \frac{(C + M)(C + M - \epsilon CM)}{(\epsilon CM)^2} B_2. \qquad (33)$$

This error depends on system parameters C, M, step-size ϵ, and the constant B_2 which bounds the successive difference of the instantaneous optimal solutions. It can be seen that the larger the step-size is, the smaller the steady-state error bound is. Letting the step-size be $\epsilon = \frac{2}{C+M}$ (best achievable) further yields that the steady-state error does not exceed $\frac{(C+M)^2(C^2+M^2)B_2}{4C^2M^2}$. To sum up, under the settings of both dynamic objective and dynamic constraint, the stable step-size is slightly smaller than the one in the static case (no optimality drift) but under both situations, the achievable step-sizes are on the same order $O\left(\frac{1}{M}\right)$ because $\frac{2}{M} \geq \frac{2}{C+M} \geq \frac{2}{M+M} = \frac{1}{M}$.

Remark 2 (Time-Invariant Box Constraints): For the special case that the box constraints are time-invariant (only objective is time-varying), we can show that the stepsize choice to achieve stability is the same to the static case of $\epsilon \in (0, 2/M)$ [36]. This way, the same choice holds for static, dynamic, or asynchronous scenarios. Constant limits on reactive power are the case for photovoltaic inverters if the solar irradiance stays the same during e.g., night time and no-cloud scenarios.

VI. NUMERICAL TESTS

We investigate the performance of the decentralized voltage control scheme under the settings of asynchronous update and dynamically time-varying network operating conditions. The desired voltage magnitude μ_j is chosen to be 1 at every bus j. Each bus is assumed to have a certain number of PV panels installed, and thus it is able to control its reactive power via advanced inverter design. All numerical tests are performed in MathWorks © MATLAB 2014a software.

A single-phase radial power distribution network consisting of 21 buses is first used to test the algorithm. This network is equivalent to the system in Fig. 1 for $N = 20$. The impedance of each line segment is set to be $(0.233 + j0.366)\Omega$. Hence, the linearized flow equations in (2) is only an approximate model. The limits of reactive power resources at every bus is chosen to be $[-100, 100]$ kVA. More realistic test using a 123-bus multi-phase network will be presented later on.
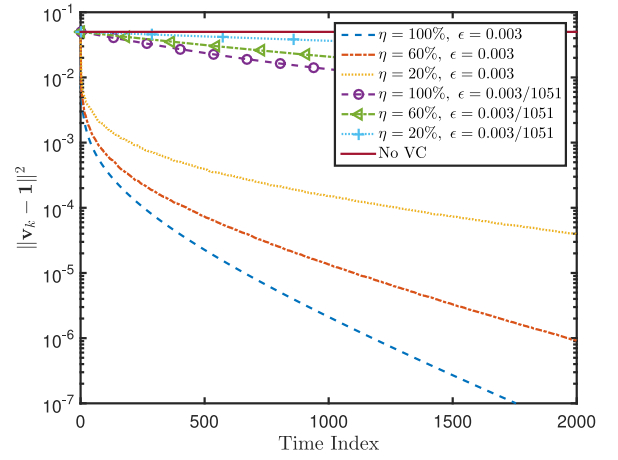

Fig. 2. Iterative voltage mismatch error performance for the asynchronous decentralized voltage control scheme under various choices of duty cycle η and step-size ϵ.

Test Case 1: The impact of asynchronous updates across different buses is first considered under a constant nominal voltage $\overline{\mathbf{v}}_k$. The maximum update delay is set to be $K = 50$. For the 21-bus network, the theoretical upper bound on the step-size is $\epsilon < 2/M = 0.0062$ following Theorem 1. Hence, we set the step-size to be $\epsilon = 1/M = 0.0031$. To model the level of asynchronous updates among multiple buses, we introduce a duty cycle parameter $\eta \in (0, 100\%]$. For a cycle of total $\frac{K}{2}$ time slots, we randomly pick $\lceil \eta \times \frac{K}{2} \rceil$ number of slots for bus j to implement its voltage control update. Hence, the maximum update delay among any two nodes is no more than K. In addition, the larger η is, the more frequently every bus performs an update, and the smaller the effective update delay would be. In particular, the setting of $\eta = 100\%$ provides the benchmark performance of the synchronous scenario where each bus updates at every time slot. Fig. 2 plots the iterative voltage mismatch error performance for the decentralized voltage control design under different η values and choices of step-size. The case of no voltage control is also plotted with the corresponding error staying constant. Using the classical convergence conditions for asynchronous GP updates in (11), the step-size should be chosen as $\epsilon = 0.0031/1051$ with $N = 20$ and $K = 50$. As shown in Fig. 2, this choice of step-size is too conservative. Thus, the resultant convergence speed is much slower than that of the choice $\epsilon = 0.0031$ following Theorem 1. This demonstrates that our theoretical results for asynchronous GP updates are more competitive for the specific decentralized voltage control application here.

Moreover, it is observed that the convergence accuracy would depend on the total number of updates for the whole network. Because of the asynchronous update settings, the expected number of control updates across the network for a cycle of $\frac{K}{2}$ iterations equals to $\lceil \eta \times \frac{K}{2} \rceil \times N$. Fig. 3 illustrates the voltage mismatch error performance versus the expected number of total control updates. Interestingly, the convergence speed in this plot is the same for the same step-size ϵ value, regardless of the asynchronous metric η. Hence, the average update rate across
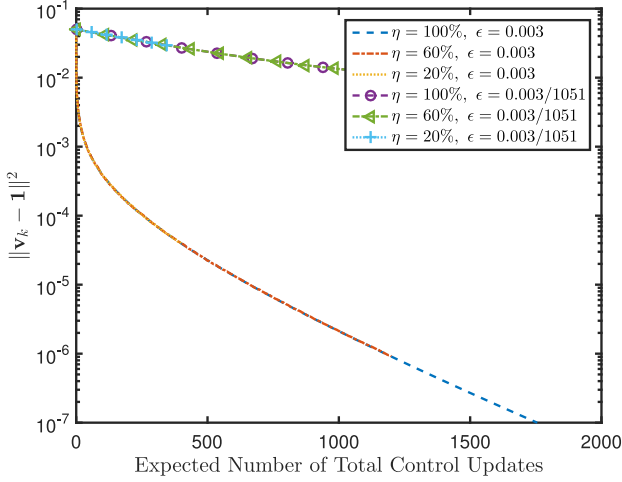
Fig. 3. Voltage mismatch error versus the total number of updates across the network for the asynchronous voltage control scheme under various choices of duty cycle η and step-size ϵ.
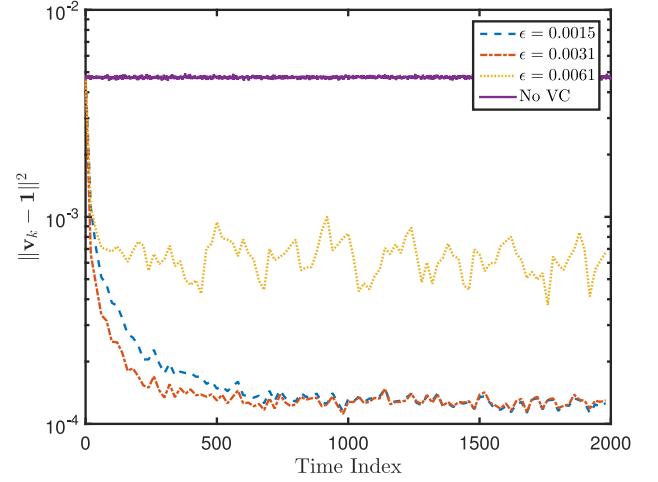

Fig. 5. Iterative voltage mismatch error performance averaged over 30 random realizations under different ϵ values.
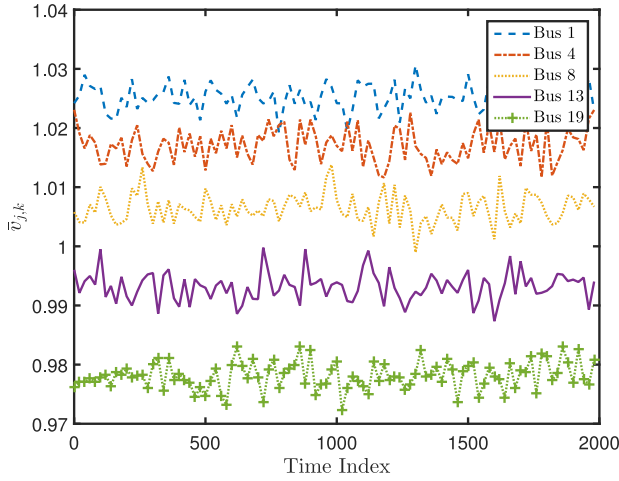

Fig. 4. An instance of the nominal voltage series $\{\overline{\mathbf{v}}_k\}$ at selected buses under the AR(1) model settings.
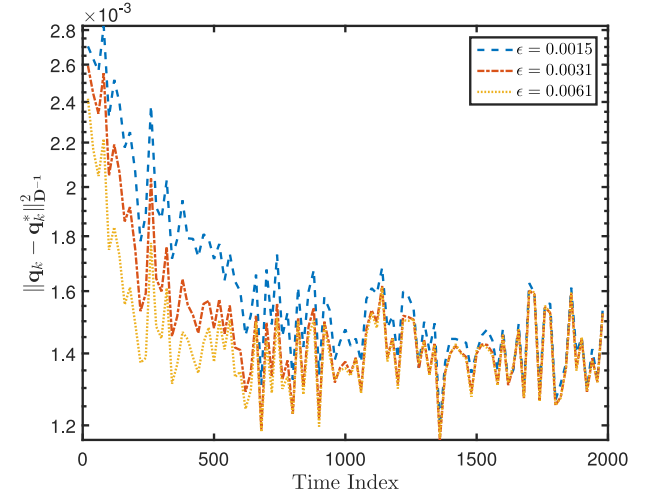

Fig. 6. Iterative tracking error averaged over 30 random realizations under different ϵ values.

all the buses determines the performance of the asynchronous decentralized voltage control scheme.

Test Case 2: To verify our results on dynamic voltage control, we generate the nominal voltage series $\{\overline{\mathbf{v}}_k\}$ using the AR(1) model in (14). Neglecting the effects of voltage regulators, it is well known that the voltage magnitude in power networks tends to decrease monotonically away from the root node, i.e., bus 0. Hence, we set the mean voltage at bus j to be $\overline{c}_j/(1-\alpha) = 1.025 - \frac{0.05}{19}(j-1)$ to follow this decreasing voltage rule. Fig. 4 plots the nominal voltage sequence at selected network locations for the choice of $\alpha = 0.1$ and noise variance $\sigma^2 = 6 \times 10^{-6}$. This choice of the forgetting factor α value leads very fast dynamics in the nominal voltage.

1) The Step-Size ϵ: Fig. 5 plots the iterative voltage mismatch error performance using different choices of ϵ, while Fig. 6 plots the weighted tracking error between the iterate $\mathbf{q}_k$ and the corresponding instantaneous optimal $\mathbf{q}_k^*$. Both plots are averaged over 30 random realizations of the nominal voltage

series to approximate the expected values. The maximum value $\epsilon = 0.0061$ is chosen according to the bound $2/(C + M)$ in Theorem 2. As shown more clearly in Fig. 6, a larger step-size ϵ leads to slightly faster convergence of the tracking error. However, the steady-state voltage mismatch error is higher for the largest ϵ as shown in Fig. 5. This observation coincides with the analytical results of Theorem 2. The convergence geometric rate ρ depends on an appropriate choice of ϵ, while the steady-state error related constant Θ tends to increase with a larger ϵ choice. Thus, the choice of ϵ would be able to trade the steady-state tracking error off the convergence speed. Under this trade-off, the optimal selection of ϵ would also depend on the dynamics of the AR(1) process. If the nominal voltage evolves very fast, it is preferred to have a large ϵ for a better tracking performance. Otherwise, if the dynamics of the nominal voltage series has a large time constant, we can afford to have a small ϵ in order to achieve a better tracking error performance. By analyzing the bounds in Theorem 2, it is possible to provide a general
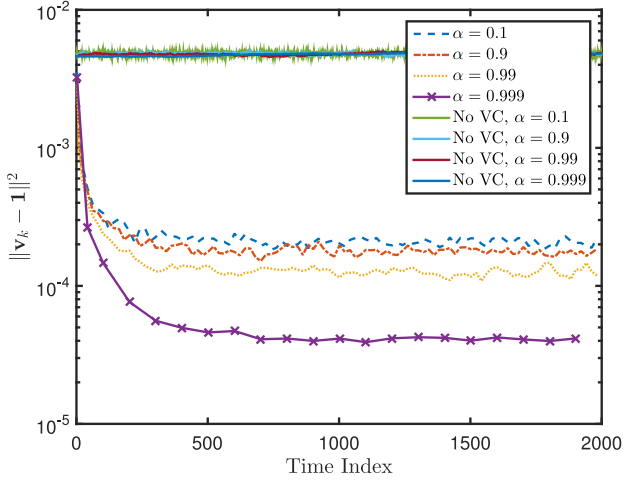
Fig. 7. Iterative voltage mismatch error performance averaged over 30 random realizations for the voltage control scheme with various values of forgetting factor α.
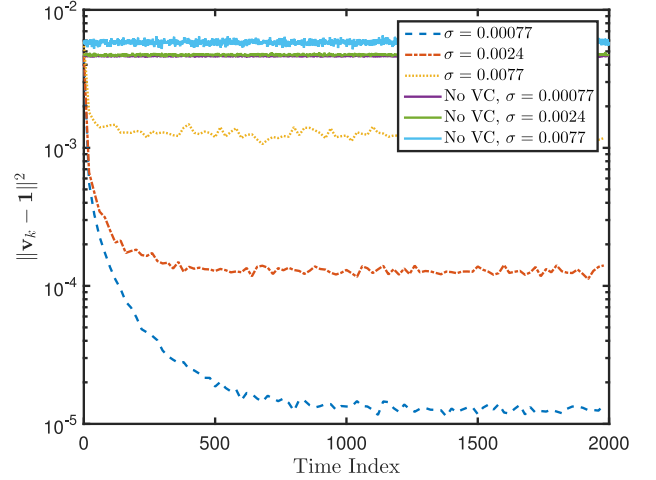

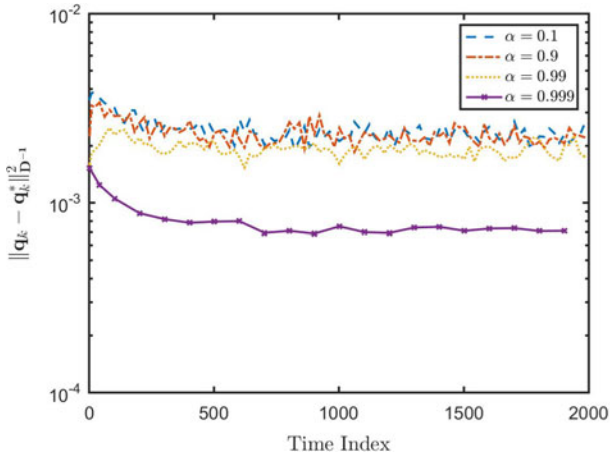Fig. 9. Iterative voltage mismatch error performance averaged over 30 random realizations under various σ values.

Fig. 8. Iterative tracking error averaged over 30 random realizations for the voltage control scheme with various values of forgetting factor α.
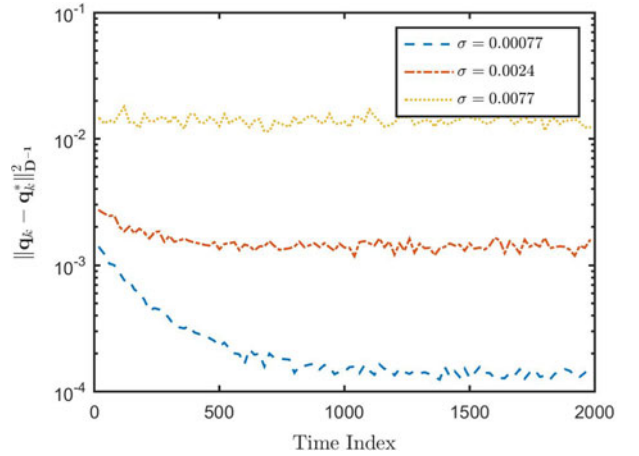

Fig. 10. Iterative tracking error averaged over 30 random realizations under various σ values.

guideline on selecting a proper ϵ value based on the dynamics of $\overline{\mathbf{v}}_k$.

2) The Forgetting Factor α: We have also varied the parameter α for the AR(1) process used to generate the nominal voltage series, with values ranging from 0.1 to 0.999. The step-size ϵ is fixed at 0.0031. For comparison purposes, the variance of the nominal voltage at every bus is aligned to be the same for different α values by setting it to be $\sigma^2/(1 - \alpha^2) = 10^{-5}$. Hence, when $\alpha = 0.999$ very closely approaches 1, the $\{\overline{\mathbf{v}}_k\}$ series would almost stays flat with minimal temporal variations. Accordingly, the consecutive voltage mismatch error bound B_1 in Prop. 2 decreases as α approaches its upper bound 1. Fig. 7 plots the iterative voltage mismatch error for various α values, while Fig. 8 again plots the corresponding weighted tracking error performance. These curves are also averaged over 30 random realizations. The performance under either mismatch error metrics improves with a larger α value since the constant B_1 and thereby B_2 would decrease. Accordingly, this leads to a smaller

Θ value and reduces the steady-state tracking error. This numerical result points out that B_2, which bounds the optimizer drift, could be related to the consecutive voltage difference B_1.

3) The Noise Variance σ^2: Similar test has been conducted with varying parameter σ for the AR(1) process ranging from 7.7×10^{-4} to 7.7×10^{-3}. With the same step-size $\epsilon = 0.0031$ and parameter $\alpha = 0.1$, Fig. 9 and Fig. 10 plot the average voltage mismatch and weighted tracking error over 30 realizations. Similar to the observations under various α values, using a smaller σ^2 would decrease the steady-state error bounds since the constant Θ tends to be positively related to the parameter σ^2. As the parameter α is fixed, the variance of the nominal voltage at every bus decreases as the noise variance diminishes. Accordingly, the performance of no voltage control slightly improves with smaller noise variance.

All numerical results in this test case have verified our analytical bounds on the tracking error performance. To sum up, the convergence speed depends on the choice of step-size ϵ. Depending on the time constant of nominal voltage dynam-
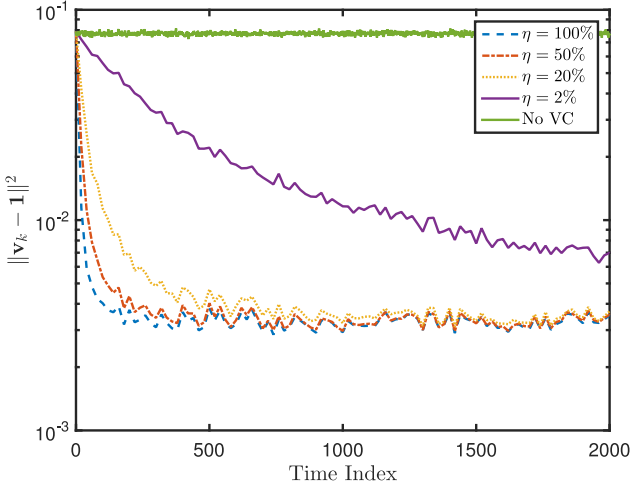
Fig. 11. Iterative voltage mismatch error performance on the IEEE 123-bus system under both the asynchronous updates and the dynamic network operating conditions with different η values.
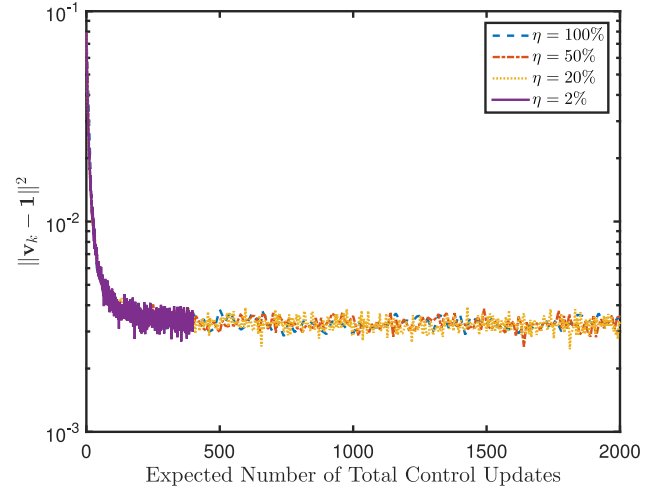

Fig. 12. Voltage mismatch error versus the total number of updates across the IEEE 123-bus system for the asynchronous voltage control scheme under dynamic operating conditions with various choices of duty cycle η.

ics, the step-size needs to be properly chosen trading off the convergence speed and the steady-state error performance. The dynamics of nominal voltage series based on the AR(1) process parameters does not affect the convergence speed per se, yet more significantly related to the steady-state tracking error performance. Note that the analytical bounds of Theorem 2 are not tight, because of the scalar β' used to eliminate the cross-product terms from the squared sum norm. However, they are very effective to characterize the error performance of dynamic decentralized voltage control scheme while facilitating the selection of step-size.

Test Case 3: We have also tested the decentralized voltage control scheme using a realistic multi-phase distribution network, namely the IEEE 123-bus system model [41]. This test case incorporates both the asynchronous updates among multiple buses and dynamical nominal voltage, with similar settings as in the earlier two test cases. Moreover, the line segments of the 123-bus system are all lossy and involves inter-phase mutual couplings. Hence, this test provides an accurate representation of how the decentralized voltage control would perform in practice with uncertainties in the DER hardware and network operating conditions.

Fig. 11 plots the iterative network voltage mismatch error under various choices of duty cycle parameters η. The step-size $\epsilon = 0.01$ has been chosen for every scenario to ensure stability. Different from the earlier two test cases, the control implementation has incorporated both the asynchronous updates and the dynamic voltage profile. Although we have not be able to derive the tracking error bounds under both sources of uncertainty, Fig. 11 demonstrates that its convergence speed results is similar to the solely asynchronous case as in Fig. 2, while the steady-state tracking error may have similar bounds as in the dynamic control analysis. In addition, Fig. 12 illustrates the voltage mismatch error performance versus the expected number of total control updates. Similar to its single-phase counterpart, the convergence speed in this plot is analogous for a fixed step-size ϵ value, regardless of the asynchronous metric η. Thus, we are

confident that the analysis of this paper can be integrated to a joint framework that characterizes the tracking error performance under both uncertain sources.

VII. CONCLUSION AND FUTURE WORK

This paper develops a decentralized dynamic optimization framework for analyzing the performance of a voltage control scheme based on gradient-projection (GP) methods for online system implementations. By constructing the linearized flow model for power distribution networks, one can design a voltage control scheme by minimizing a surrogate voltage mismatch error using the GP iterations. Thanks to the physical power network coupling, this GP-based scheme boils down to a decentralized voltage control design where every bus can measure its local voltage to obtain the instantaneous gradient direction. Compared to earlier results for a static optimization scenario, we have significantly extended the analysis on convergence conditions and error performance to account for two dynamic scenarios: i) the nodes perform the decentralized update in an asynchronous fashion; and ii) the network operating point is dynamically changing. Assuming the nominal voltage evolves following an AR(1) process, the weighted tracking error can be bounded by an exponentially decreasing term plus a constant term that would depend on the successive difference of the transient optimal solution. Interestingly, the choice of step-size may need to be more conservative depending on the trade-off between the convergence speed and the steady-state tracking error for the dynamic control design. Several numerical tests have been performed to demonstrate and validate our analytical results on the performance of the decentralized voltage control scheme under realistic dynamic scenarios using practical power network models.

Future work includes the development of an integrated framework for performance analysis under both dynamic scenarios simultaneously. We are also interested to investigate further on the impacts of vastly different time-scale among all network

resources, which will help us to characterize the interactions with traditional voltage control devices of slower time-scales.

REFERENCES

[1] P. Carvalho, P. F. Correia, and L. Ferreira, "Distributed reactive power generation control for voltage rise mitigation in distribution networks," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 766–772, May 2008.

[2] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proc. IEEE*, vol. 99, no. 6, pp. 1063–1073, Jun. 2011.

[3] B. Robbins, C. Hadjicostis, and A. Dominguez-Garcia, "A two-stage distributed architecture for voltage control in power distribution systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1470–1482, May 2013.

[4] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter VAR control in distribution systems with high PV penetration," in *Proc. 2012 IEEE Power Energy Soc. Gen. Meeting*, Jul. 2012, pp. 1–7.

[5] E. Dall'Anese, H. Zhu, and G. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.

[6] P. Sulc, S. Backhaus, and M. Chertkov, "Optimal distributed control of reactive power via the alternating direction method of multipliers," *IEEE Trans. Energy Convers.*, vol. 29, no. 4, pp. 968–977, Dec. 2014.

[7] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *Proc. IEEE 52nd Conf. Decision Control*, Dec. 2013, pp. 4329–4334.

[8] H. Zhu and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794–3803, Sep. 2016.

[9] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.

[10] H. R. Feyzmahdavian and M. Johansson, "On the convergence rates of asynchronous iterations," in *Proc. IEEE 53rd Conf. Decision Control*, 2014, pp. 153–159.

[11] P. Scott and S. Thiébaux, "Dynamic optimal power flow in microgrids using the alternating direction method of multipliers," *ArXiv: 1410.7868*, Oct. 2014. [Online]. Available: https://arxiv.org/abs/1410.7868

[12] S. Gill, I. Kockar, and G. W. Ault, "Dynamic optimal power flow for active distribution networks," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 121–131, Jan. 2014.

[13] K. Slavakis, S.-J. Kim, G. Mateos, and G. Giannakis, "Stochastic approximation vis-a-vis online learning for big data analytics [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 124–129, Nov. 2014.

[14] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. 19th Int. Conf. Comput. Statist.*, Aug. 2010, pp. 177–187. [Online]. Available: http://leon.bottou.org/papers/bottou-2010.

[15] R. Johnson and T. Zhang, "Accelerating stochastic gradient descent using predictive variance reduction," in *Proc. Adv. Neural Inf. Process. Syst. 26*, 2013, pp. 315–323.

[16] V. Kekatos, G. Wang, A. Conejo, and G. Giannakis, "Stochastic reactive power management in microgrids with renewables," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3386–3395, Nov. 2015.

[17] Q. Ling and A. Ribeiro, "Decentralized dynamic optimization through the alternating direction method of multipliers," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1185–1197, Mar. 2014.

[18] Z. J. Towfic and A. H. Sayed, "Adaptive penalty-based distributed stochastic convex optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3924–3938, Aug. 2014.

[19] A. Simonetto, A. Mokhtari, A. Koppel, G. Leus, and A. Ribeiro, "A class of prediction-correction methods for time-varying convex optimization," *arXiv: 1509.05196*, Sep. 2015. [Online]. Available: https://arxiv.org/abs/1509.05196

[20] K. Zhou and S. Roumeliotis, "Multirobot active target tracking with combinations of relative observations," *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 678–695, Aug. 2011.

[21] F. Jakubiec and A. Ribeiro, "D-map: Distributed maximum a posteriori probability estimation of dynamic systems," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 450–466, Jan. 2013.

[22] M. Baran and F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725–734, Jan. 1989.

[23] P. Sulc, S. Backhaus, and M. Chertkov, "Optimal distributed control of reactive power via the alternating direction method of multipliers," *IEEE Trans. Energy Convers.*, vol. 29, no. 4, pp. 968–977, Dec. 2014.

[24] Q. Li, R. Ayyanar, and V. Vittal, "Convex optimization for des planning and operation in radial distribution systems with high penetration of photovoltaic resources," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 985–995, Jul. 2016.

[25] B. A. Robbins and A. D. Dominguez-Garcia, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903–2913, Jul. 2016.

[26] D. B. West, *Introduction to Graph Theory*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[27] B. Robbins, H. Zhu, and A. D. Dominguez-Garcia, "Optimal tap setting of voltage regulation transformers in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 256–267, Jan. 2016.

[28] E. Dall'Anese, H. Zhu, and G. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.

[29] S. H. Low, "Convex relaxation of optimal power flow-part II: Exactness," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 2, pp. 177–189, Jun. 2014.

[30] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Voltage regulation algorithms for multiphase power distribution grids," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3921–3923, Sep. 2016.

[31] H. J. Liu, R. Macwan, N. Alexander, and H. Zhu, "A methodology to analyze conservation voltage reduction performance using field test data," in *Proc. 2014 IEEE Int. Conf. Smart Grid Commun.*, 2014, pp. 529–534.

[32] P. Huber, J. Wiley, and W. InterScience, *Robust Statistics*. New York, NY, USA: Wiley, 1981.

[33] D. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1999.

[34] M. Hassanzadeh, C. Y. Evrenosoğlu, and L. Mili, "A short-term nodal voltage phasor forecasting method using temporal and spatial correlation," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3881–3890, Sep. 2016.

[35] E. Cotilla-Sanchez, P. Hines, C. Barrows, and S. Blumsack, "Comparing the topological and electrical structure of the north american electric power infrastructure," *IEEE Syst. J.*, vol. 6, no. 4, pp. 616–626, Dec. 2012.

[36] H. J. Liu, W. Shi, and H. Zhu, "Dynamic decentralized voltage control for power distribution networks," in *Proc. 2016 IEEE Statist. Signal Process. Workshop*, Jun. 2016, pp. 1–5.

[37] D. Bertsekas, "Incremental gradient, subgradient, and proximal methods for convex optimization: A survey," *Optim. Mach. Learn.*, vol. 2010, pp. 1–38, 2011.

[38] W. Shi, Q. Ling, G. Wu, and W. Yin, "A proximal gradient algorithm for decentralized composite optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6013–6023, Nov. 2015.

[39] D. Davis and W. Yin, "Convergence rate analysis of several splitting schemes," *arXiv:1406.4834*, Jun. 2014. [Online]. Available: https://arxiv.org/abs/1406.4834

[40] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*. New York, NY, USA: Springer, 2013, vol. 87.

[41] "IEEE PES Distribution Test Feeders," Sep. 2010. [Online]. Available: http://www.ewh.ieee.org/soc/pes/dsacom/testfeeders.html
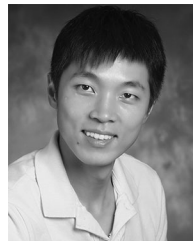
Hao Jan (Max) Liu (S'10) was born in Taipei, Taiwan. He received the B.S. degree from Missouri University of Science and Technology, Rolla, MO, USA, in 2011 and the M.S. degree from the University of Illinois Urbana-Champaign (UIUC), Champaign, IL, USA, in 2013. He is currently working toward the Ph.D. degree in electrical engineering at the UIUC. His current research interests cover the areas of optimization for power distribution networks and smart grid technology, including Volt/VAR and microgrid studies. He received the National Science Foundation East Asia and Pacific Summer Institutes Fellowship in 2015, the 2nd Best Paper Award at the 2016 *North American Power Symposium (NAPS)*, and the Siebel Scholar Award in 2017.
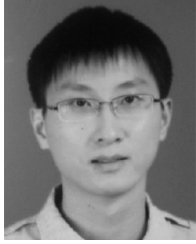
Wei Shi (M'15) received the B.E. degree in automation and the Ph.D. degree in control science and engineering from the University of Science and Technology of China, Hefei, China, in 2010 and 2015, respectively. From 2015 to 2016, he was a Postdoctoral Research Associate at Coordinated Science Laboratory, the University of Illinois at Urbana-Champaign, Urbana, IL, USA. Currently, he is a Postdoctoral Research Associate at Boston University, MA, USA. His research interests include optimization and its applications in signal processing and control.
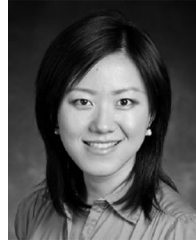
Hao Zhu (M'12) received the B.E. degree from Tsinghua University, Beijing, China, in 2006, and the M.Sc. and Ph.D. degrees from the University of Minnesota, Minneapolis, MN, USA, in 2009 and 2012, respectively, all in electrical engineering. She is an Assistant Professor of electrical and computer engineering at the University of Illinois, Urbana-Champaign (UIUC), Champaign, IL, USA. She was a Postdoc Research Associate on power grid modeling and validation at the UIUC Information Trust Institute before joining the ECE faculty in 2014. Her current research interests include power grid monitoring, power system operations and control, and energy data analytics. She received a Seed Grant from the Siebel Energy Institute and the US AFRL Summer Faculty Fellowship in 2016, and the 2nd Best Paper Award at the 2016 North American Power Symposium. She is currently a member of the steering committee of the IEEE Smart Grid representing the IEEE Signal Processing Society.