DYNAMIC DECENTRALIZED VOLTAGE CONTROL FOR POWER DISTRIBUTION NETWORKS

Hao Jan Liu, Wei Shi, and Hao Zhu

Department of ECE and CSL, University of Illinois at Urbana-Champaign
(Invited Paper)

ABSTRACT

Voltage regulation in power distribution networks has been increasingly challenged by the integration of volatile and intermittent distributed energy resources (DERs). These resources can also provide limited reactive power support that can be used to optimize the network-wide voltage. A decentralized voltage control scheme based on the gradient-projection (GP) method is adopted to minimize a voltage mismatch error objective under limited reactive power. Coupled with the power network flow, the local voltage directly provides the instantaneous gradient information. This paper aims to quantify the performance of this decentralized GP-based voltage control under dynamic system operating conditions modeled by an autoregressive process. Our analysis offers the tracking error bound on the instantaneous solution to the transient optimizer. Under stochastic processes that have bounded iterative changes, the results can be extended to general *constrained dynamic optimization* problems with smooth strongly convex objective functions. Numerical tests using a 21-bus network have been performed to validate our analytical results.

1. INTRODUCTION

Recent deployment of smart grid technologies has witnessed increasing penetration of distributed energy resources (DERs) such as the rooftop photovoltaic panels and batteries of electric vehicles in the power distribution networks. As these resources are generally volatile and intermittent, they greatly challenge the operational goal of maintaining a satisfactory voltage level per power system reliability standards. Through advancements in power electronics, DERs is also an excellent resource of *reactive power*, a quantity that is known to have a significant impact on the network voltage level. Hence, a plausible solution is to design effective voltage control strategies requiring minimal hardware implementations to utilize the reactive power from DERs.

The communication infrastructure deployed in power distribution networks is, and will continue to be lacking in the foreseen future. In this sense, voltage control designs by minimizing the global voltage mismatch error is often not feasible as they strongly depend on the quality of communication links either between a control center and remote devices [1], or among neighboring devices [2–4]. Therefore, decentralized voltage control frameworks are more favored for distribution network operations to tackle this challenge; see e.g., [5–8]. Our earlier work [8] has proposed several decentralized voltage control designs along with the convergence analysis for static system scenarios. Thanks to the physical power network coupling, the most up-to-date gradient direction for a centralized weighted voltage error objective can be obtained through the local voltage measurement. Interestingly, the decentralized voltage control approach using this measurement is equivalent to the classical gradient-projection (GP) method which accounts for the reactive power resource limits.

This paper offers the analytical performance of the decentralized GP-based voltage control design under dynamic power network operating conditions. This problem boils down to a *stochastic optimization* one since the objective function is time-varying; see e.g., [9]. Under stochastic approximation frameworks, recent work in [10, 11] has developed the stochastic (sub-)gradient decent method, which has been adopted by [12] for solving this voltage control problem. However, the performance analysis has been focused on the convergence to the optimal solution, which minimizes the expected objective function [9]. Aiming at the error bound in tracking the instantaneous optimal solution, our analysis is more closely relevant to the body of work on *dynamic convex optimization*; see e.g., [13, 14]. Although some of these dynamic optimization methods involve gradient descent updates [15], the formulation of constrained optimization has not been considered. However, our voltage control problem has to account for reactive power resource limits in order to provide feasible control inputs. The contribution of this work lies in the development of *constrained* dynamic optimization framework for which the projection operator in the GP updates has to be explicitly accounted for. Motivated by the specific dynamic voltage control problem, we derive the tracking error bounds for a quadratic objective function under an autoregressive dynamic process. The analytical results can be extended to generic constrained dynamic op-

This material is based upon work supported by the Department of Energy under Award Number DE-OE000078. Contact authors. Emails: {haoliu6,wilburs,haozhu} @illinois.edu.

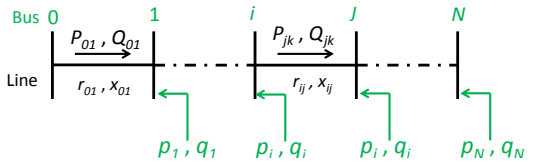

Fig. 1. A radial distribution network with bus and line associated variables.

timization problems with smooth strongly convex objective functions under stochastic processes that have bounded iterative changes.

2. SYSTEM MODELING

A power distribution network is modeled using a tree-topology graph $(\mathcal{N}, \mathcal{E})$ with the set of buses (nodes) $\mathcal{N} := \{0, ..., N\}$ 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 active (reactive) power injection, respectively. All system quantities are in per unit (p.u.). 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} represent 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., $v_j \cong 1$, $\forall j$. For each (i, j), the linearized power flow [16] asserts the bus power balance and line voltage drop under these assumptions, as given by

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

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

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

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

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., [17, pg. 6]). We set $M^o_{i\ell} = 1$ and $M^o_{j\ell} = -1$ if $j \in \mathcal{N}_i^+$. Let $\mathbf{m}_0^T$ represent the first row of $\mathbf{M}^o$ corresponding to bus 0 while the rest of the rows are in the $N \times N$ submatrix $\mathbf{M}$. Since the network is connected under the tree-topology assumption, $\mathbf{M}$ is full-rank and invertible [17]. Upon concatenating all scalar variables into vector form, one can represent (1) as

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

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

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

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. One can solve for $\mathbf{P}$ and $\mathbf{Q}$ in (2) by viewing the uncontrollable $\mathbf{p}$ as a constant 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 $\mathbf{X} := \mathbf{M}^{-T}\mathbf{D}_x\mathbf{M}^{-1}$ represents the network reactance matrix which is positive definite. The linear model (3) constitutes the basis for developing decentralized voltage control designs.

3. 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 $\bar{\mathbf{v}}_k$ 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

$$\mathbf{q}_k^* = \arg\min_{\mathbf{q}\in\mathcal{Q}} f_k(\mathbf{q}) := \frac{1}{2}\|\mathbf{Xq} + \bar{\mathbf{v}}_k - \boldsymbol{\mu}\|_{\mathbf{B}}^2 \tag{4}$$

where the weighted squared norm $\|\mathbf{x}\|_{\mathbf{B}}^2 := \mathbf{x}^T\mathbf{Bx}$, $\forall \mathbf{x}$. The constraint set $\mathcal{Q} := \{\mathbf{q}|\mathbf{q} \in [\underline{\mathbf{q}}, \bar{\mathbf{q}}]\}$ accounts for the limits of local reactive power resources at every bus [6]. Although the quadratic convex problem (4) minimizes a weighted 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.

Thanks to the separable structure of the box constraint $\mathcal{Q}$, the gradient-projection (GP) method [18, Sec. 2.3] can be invoked to solve (4). Upon forming its instantaneous gradient $\nabla f_k(\mathbf{q}_k) := \mathbf{Xq}_k + \bar{\mathbf{v}}_k - \boldsymbol{\mu}$, the GP iteration for a given positive stepsize $\epsilon > 0$ is

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

where the projection operator $\mathbb{P}$ bounds any input to be within $\mathcal{Q}$, and $\mathbf{D} := [\mathrm{diag}(\mathbf{X})]^{-1}$ approximates the Newton gradient [18, Sec. 3 .3]. 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 $\mathbf{q}_k \in \mathcal{Q}$ to be the control input at any time k, the instantaneous voltage becomes $\mathbf{v}_k = \mathbf{Xq}_k + \bar{\mathbf{v}}_k$ based on (3). Thanks to the physical power network coupling, $\mathbf{v}_k$ always provides the most 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

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

which can be completely decoupled into decentralized updates at each bus because $\mathbb{P}$ is separable. With a proper stepsize ϵ choice, the decentralized update (6) converges to the optimizer $\mathbf{q}_k^*$ to the problem (4) for a constant $\bar{\mathbf{v}}_k = \bar{\mathbf{v}}$ [8].

4. DYNAMIC PERFORMANCE ANALYSIS

The volatility and intermittence of loads and generations lead to temporal variations in the network operating condition, i.e., a dynamic $\bar{\mathbf{v}}_k$. 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 used to model the short-term voltage dynamics.

(as1) *For a given constant vector $\bar{\mathbf{c}}$, the nominal voltage $\bar{\mathbf{v}}_k$ follows a wide-sense stationary AR(1) process, as given by*

$$\bar{\mathbf{v}}_{k+1} = \mathbf{A}\bar{\mathbf{v}}_k + \boldsymbol{\eta}_{k+1} + \bar{\mathbf{c}} \tag{7}$$

where $\mathbf{A}$ is a constant transition matrix, and $\boldsymbol{\eta}_{k+1}$ represents a zero-mean white noise process with covariance matrix $\boldsymbol{\Sigma}_\eta$. The largest eigenvalue of matrix $\mathbf{A}$ is less than 1 to ensure stability,.

The spatial correlation is often time negligible for real power networks, as demonstrated in [19]. Hence, the AR(1) model (7) can be simplified as follows:

$$\bar{\mathbf{v}}_{k+1} = \alpha\bar{\mathbf{v}}_k + \boldsymbol{\eta}_{k+1} + \bar{\mathbf{c}} \tag{8}$$

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

To quantify the performance of (6) under the dynamic model in (7), we first introduce an equivalent form for the update (6), given by

$$\begin{aligned} & \mathbf{D}^{-\frac{1}{2}}\mathbf{q}_{k+1} \\ = \; & \tilde{\mathbb{P}}[\mathbf{D}^{-\frac{1}{2}}\mathbf{q}_k - \epsilon\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{-\frac{1}{2}}\mathbf{q}_k - \epsilon\mathbf{D}^{-\frac{1}{2}}(\bar{\mathbf{v}}_k - \boldsymbol{\mu})] \end{aligned} \tag{9}$$

where the projection operator $\tilde{\mathbb{P}}[\cdot]$ thresholds any input to the set $\tilde{\mathcal{Q}} := \{\mathbf{q}|\mathbf{q} \in [\mathbf{D}^{-\frac{1}{2}}\underline{\mathbf{q}}, \mathbf{D}^{-\frac{1}{2}}\bar{\mathbf{q}}]\}$.

Lemma 1. *For the AR(1) model under (as1), the expectation of the weighted norm of consecutive difference is bounded; i.e., there exists $B < \infty$ such that*

$$\mathbb{E}\|\bar{\mathbf{v}}_{k+1} - \bar{\mathbf{v}}_k\|^2_{\mathbf{D}} \le B. \tag{10}$$

Proof: For the convenience of exposition, we show it using the simplified model (8). Specifically, its weighted covariance equals to

$$\mathbb{E}\|\bar{\mathbf{v}}_k - \mathbb{E}\bar{\mathbf{v}}_k\|^2_{\mathbf{D}} = \frac{1}{1-\alpha^2}\mathbb{E}\|\boldsymbol{\eta}_k\|^2_{\mathbf{D}} = \frac{\sigma^2\mathrm{Tr}\mathbf{D}}{1-\alpha^2}$$

Accordingly, we can show that

$$\begin{aligned} \mathbb{E}\|\bar{\mathbf{v}}_{k+1} - \bar{\mathbf{v}}_k\|^2_{\mathbf{D}} &= \mathbb{E}\|(\alpha-1)(\bar{\mathbf{v}}_k - \mathbb{E}\bar{\mathbf{v}}_k) + \boldsymbol{\eta}_{k+1}\|^2_{\mathbf{D}} \\ &= \mathbb{E}\left\{(\alpha-1)^2\|\bar{\mathbf{v}}_k - \mathbb{E}\bar{\mathbf{v}}_k\|^2_{\mathbf{D}} + \|\boldsymbol{\eta}_{k+1}\|^2_{\mathbf{D}}\right\} \\ &= \frac{2\sigma^2\mathrm{Tr}\mathbf{D}}{1+\alpha}. \end{aligned}$$

For the simplified model, the upper bound B in (10) is exact. It is possible to follow the same analysis for the general AR(1) model (7) to derive an upper bound B, which would similarly depend on $\boldsymbol{\Sigma}_\eta$ and $\mathbf{A}$.

Lemma 2. *For any vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ of the same dimension and any positive scalars $a, b > 0$, the following equality holds*

$$\begin{aligned} \|\mathbf{a}+\mathbf{b}+\mathbf{c}\|^2 \; \le & (1+\frac{1}{a})(1+\frac{1}{b})\|\mathbf{a}\|^2 \\ & + (1+\frac{1}{a})(1+b)\|\mathbf{b}\|^2 + (1+a)\|\mathbf{c}\|^2. \end{aligned} \tag{11}$$

Theorem 1. *Under (as1) and for any $0 < \epsilon < 2/M$, where M is the largest eigenvalue of matrix $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$, 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*

$$\mathbb{E}\|\mathbf{q}_k - \mathbf{q}^*_k\|^2_{\mathbf{D}^{-1}} \le \rho^k\mathbb{E}\|\mathbf{q}_0 - \mathbf{q}^*_0\|^2_{\mathbf{D}^{-1}} + \tfrac{1-\rho^k}{1-\rho}\Theta, \; \forall k \tag{12}$$

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

Proof: Let us form the weighted tracking error

$$\begin{aligned} \tilde{\mathbf{q}}_{k+1} &:= \mathbf{D}^{-\frac{1}{2}}\mathbf{q}_{k+1} - \mathbf{D}^{-\frac{1}{2}}\mathbf{q}^*_{k+1} \\ &= \tilde{\mathbb{P}}\left[\mathbf{D}^{-\frac{1}{2}}\mathbf{q}_k - \epsilon\mathbf{D}^{\frac{1}{2}}(\mathbf{X}\mathbf{q}_k + \bar{\mathbf{v}}_k - \boldsymbol{\mu})\right] \\ &\quad - \tilde{\mathbb{P}}\left[\mathbf{D}^{-\frac{1}{2}}\mathbf{q}^*_{k+1} - \epsilon\mathbf{D}^{\frac{1}{2}}(\mathbf{X}\mathbf{q}^*_{k+1} + \bar{\mathbf{v}}_{k+1} - \boldsymbol{\mu})\right]. \end{aligned}$$

Using Lemma 2 and denoting $\hat{b} := (1+\frac{1}{a})(1+\frac{1}{b})$ and $\bar{b} := (1+\frac{1}{a})(1+b)$, the weighted tracking error norm can be further bounded by

$$\begin{aligned} \|\tilde{\mathbf{q}}_{k+1}\|^2 \le & \hat{b}\|\mathbf{I} - \epsilon\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\|^2\|\tilde{\mathbf{q}}_k\|^2 \\ & + \bar{b}\|\mathbf{I} - \epsilon\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\|^2\|\mathbf{q}^*_k - \mathbf{q}^*_{k+1}\|^2_{\mathbf{D}^{-1}} \\ & + (1+a)\epsilon^2\|\bar{\mathbf{v}}_{k+1} - \bar{\mathbf{v}}_k\|^2_{\mathbf{D}}. \end{aligned} \tag{13}$$

The second term in the right-hand side involves the following weighted norm:

$$\begin{aligned} & \|\mathbf{q}^*_k - \mathbf{q}^*_{k+1}\|^2_{\mathbf{D}^{-1}} \\ = \; & \Big\|\tilde{\mathbb{P}}\left[\mathbf{D}^{-\frac{1}{2}}\mathbf{q}^*_k - \epsilon\mathbf{D}^{\frac{1}{2}}(\mathbf{X}\mathbf{q}^*_k + \bar{\mathbf{v}}_k - \boldsymbol{\mu})\right] \\ & - \tilde{\mathbb{P}}\left[\mathbf{D}^{-\frac{1}{2}}\mathbf{q}^*_{k+1} - \epsilon\mathbf{D}^{\frac{1}{2}}(\mathbf{X}\mathbf{q}^*_{k+1} + \bar{\mathbf{v}}_{k+1} - \boldsymbol{\mu})\right]\Big\|^2 \\ \le \; & (1+\tfrac{1}{c})\|\mathbf{I} - \epsilon\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\|^2\|\mathbf{q}^*_k - \mathbf{q}^*_{k+1}\|^2_{\mathbf{D}^{-1}} \\ & + (1+c)\epsilon^2\|\bar{\mathbf{v}}_{k+1} - \bar{\mathbf{v}}_k\|^2_{\mathbf{D}}, \end{aligned}$$

for any scalar $c > 0$. Taking the expectation on both sides leads to the bound

$$\mathbb{E}\|\mathbf{q}^*_k - \mathbf{q}^*_{k+1}\|^2_{\mathbf{D}^{-1}} \le \frac{(1+c)\epsilon^2}{1-(1+c^{-1})\|\mathbf{I} - \epsilon\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\|^2}B \tag{14}$$

where the constant B follows from Lemma 1. Further define $\rho := \hat{b}\|\mathbf{I} - \epsilon \mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\|^2$ and a positive constant

$$\Theta \quad := \left(\frac{\bar{b}(1+c)\epsilon^2}{\|\mathbf{I}-\epsilon \mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\|^{-2}-(1+c^{-1})} + (1+a)\epsilon^2 \right) B.$$

By taking the expectation of (13) and substituting (14), we write the equality as

$$\mathbb{E}\|\tilde{\mathbf{q}}_{k+1}\|^2 \leq \rho \mathbb{E}\|\tilde{\mathbf{q}}_k\|^2 + \Theta. \tag{15}$$

One can eventually show the inequality (12) by applying induction. As long as $0 < \epsilon < 2/M$, we have $\|\mathbf{I} - \epsilon \mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\| < 1$. In addition, for any choice of $a, b > 0$ and accordingly any $\hat{b} = (1+\frac{1}{a})(1+\frac{1}{b}) \to 1$, the inequality (15) holds. Thus, if the step size $0 < \epsilon < 2/M$, we can show that the contraction coefficient $\rho := \hat{b}\|\mathbf{I} - \epsilon \mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\| \in (0,1)$. This completes our proof.

Remark 1. *(Generalizations.)* Although our tracking error bound in Theorem 1 is customized for the quadratic objective, it can be extended for analyzing general GP method for a smooth (gradient of the objective is Lipschitz continuous), strongly convex objective with slight modifications. For this general framework, the stability condition becomes $\|\mathbf{I} - \epsilon \mathbf{D}\mathbf{X}\| < 1$. Moreover, the AR(1) process assumed to model the $\bar{\mathbf{v}}_k$ series can be possibly extended to a general stochastic process that has bounded iterative changes as the constant Θ in (12) is bounded if the condition in (10) holds.

5. NUMERICAL TESTS

We investigate the performance of the decentralized voltage control scheme under time-varying system operating conditions. The desired voltage magnitude μ_j is chosen to be 1 p.u. at every bus j. A single-phase radial power distribution network consisting of 21 buses is used to test the algorithm. The impedance of each line segment is set to be $(0.233 + j0.366)\Omega$. The limits of reactive power resources at every bus are chosen to be $[-100,\ 100]$kVA. For the nominal voltage $\bar{\mathbf{v}}_k$under AR(1) process, we set the mean voltage at bus j to be $\bar{c}_j/(1-\alpha) = 1.025 - \frac{0.05}{19}(j-1)$ and noise variance $\sigma^2 = 6 \times 10^{-6}$.

For a fixed $\alpha = 0.1$, Fig. 2 plots the iterative voltage mismatch error performance averaged over 30 random realizations with various choices of ϵ. The case of no voltage control is also plotted with the corresponding error staying constant. The maximum value $\epsilon = 0.0061$ is chosen according to the bound $2/M$ in Theorem 1. As shown clearly in Fig. 2, a larger stepsize ϵ leads to slightly faster convergence of the voltage mismatch. However, the steady-state voltage mismatch error is higher for the largest ϵ. This observation coincides with the analytical results of Theorem 1. 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 off between the steady-state error and the convergence speed.

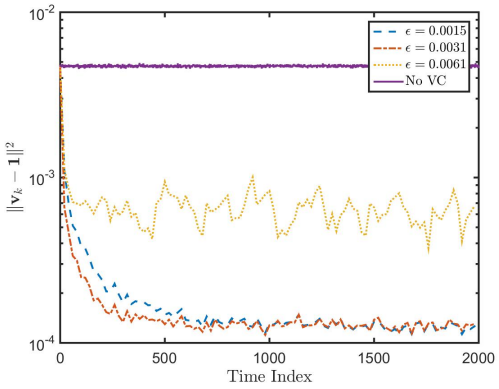

Fig. 2. Iterative voltage mismatch error performance averaged over 30 random realizations for the voltage control scheme with different ϵ values.

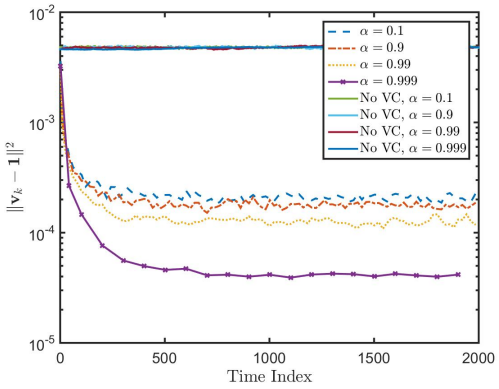

Fig. 3. Iterative voltage mismatch error performance averaged over 30 random realizations for the voltage control scheme with various values of forgetting factor α.

We also vary the forgetting factor α for the AR(1) process with values ranging from 0.1 to 0.999. The stepsize 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 $\sigma^2/(1-\alpha^2) = 10^{-5}$. Note that the expected consecutive voltage difference bound B in Lemma 1 decreases as α approaches its upper bound 1. Fig. 3 plots the corresponding voltage mismatch error performance averaged over 30 random realizations. As expected, the error performance improves with a larger α value since a smaller constant B leads to a decreasing Θ value.

In summary, the choice of stepsize ϵ affects the convergence speed. The stepsize should to be properly chosen trading off between the convergence speed and the steady-state error performance. The dynamics of nominal voltage series based on the AR(1) process parameters does not influence the convergence speed per se, yet more significantly related to the steady-state tracking error performance.

6. REFERENCES

[1] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter var control in distribution systems with high pv penetration," in *2012 IEEE Power and Energy Society General Meeting*, July 2012, pp. 1–7.

[2] 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.

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

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

[5] 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.

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

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

[8] H. Zhu and H. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1–10, 2015.

[9] 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 Processing Magazine*, vol. 31, no. 6, pp. 124–129, 2014.

[10] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of the 19th International Conference on Computational Statistics (COMPSTAT'2010)*, Y. Lechevallier and G. Saporta, Eds. Paris, France: Springer, August 2010, pp. 177–187. [Online]. Available: http://leon.bottou.org/papers/bottou-2010

[11] R. Johnson and T. Zhang, "Accelerating stochastic gradient descent using predictive variance reduction," in *Advances in Neural Information Processing Systems 26*, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 315–323.

[12] 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.

[13] 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, March 2014.

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

[15] A. Y. Popkov, "Gradient methods for nonstationary unconstrained optimization problems," *Automation and Remote Control*, vol. 66, no. 6, pp. 883–891, 2005. [Online]. Available: http://dx.doi.org/10.1007/s10513-005-0132-z

[16] 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.

[17] D. B. West, *Introduction to Graph Theory*, 2nd ed. Upper Saddle River: Prentice hall, 2001.

[18] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

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