

Real-time Decentralized and Robust Voltage Control in Distribution Networks

Guannan Qu, Na Li, and Munther Dahleh

Abstract—Voltage control plays an important role in the operation of electricity distribution networks, especially when there is a large penetration of renewable energy resources. In this paper, we focus on voltage control through reactive power compensation and study how different information structures affect the control performance. In particular, we first show that using only voltage measurements to determine reactive power compensation is insufficient to maintain voltage in the acceptable range. Then we propose two fully decentralized and robust algorithms by adding additional information, which can stabilize the voltage in the acceptable range. The one with higher complexity can further minimize a cost of reactive power compensation in a particular form. Both of the two algorithms use only local measurements and local variables and require no communication. In addition, the two algorithms are robust against heterogeneous update rates and delays.

I. INTRODUCTION

Voltages in a distribution feeder fluctuate according to the feeder loading condition. The primary purpose of voltage control is to maintain acceptable voltages (plus or minus 5% around nominal values) at all buses along the distribution feeder under all possible operating conditions. Traditionally the voltage control is achieved by re-configuring transformer taps and capacitors banks (Volt/Var control) [1], [2] based on local measurements (usually voltages) at a slow time scale. This control setting works under normal circumstances, because the change of the loading condition is relatively mild and predictable.

Due to the increasing penetration of distributed energy resources (DER) such as photovoltaic and wind generation in the distribution networks, the operating conditions (supply, demand, voltages, etc) of the distribution feeder fluctuate fast and by a large amount. The conventional voltage control lacks flexibility to respond to those conditions and they may not produce the desired results. This raises important issues on the network security and reliability. To overcome the challenges, new technologies are being proposed and developed, e.g., the inverter design for voltage control. Inverters connect DERs to the grid and adjust the reactive power outputs to stabilize the voltages at a fast time scale [3], [4]. The new technologies

Some preliminary results have been published at the 2014 Allerton Conference on Communication, Control and Computing. This journal version contains several new results that are not contained in the conference paper, especially the robustness of the control algorithms and the related proofs.

G. Qu and N. Li are with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA (Emails: gqu@g.harvard.edu, nali@seas.harvard.edu). M. Dahleh are with the Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA (Email: dahleh@mit.edu.) *Corresponding Author: N. Li*

will enable realtime distributed voltage control that is needed for the future power grid.

One key element to implement those new technologies is the voltage control algorithms which satisfy certain information constraints yet guarantee the overall system performance. In general, in the low/medium voltage distribution networks, only a small portion of buses are monitored, individuals are unlikely to announce their generation or load profile, the availability of DERs are fluctuating and uncertain, and even the grid parameters and the topology are only partially known. All of these facts make decentralized algorithms necessary for the voltage control. Each control component adjusts its reactive power input based on the local signals that are easy to measure, calculate, and/or communicate. Such local information dependence facilitates the realtime implementation of the algorithms. In fact, there exist classes of inverter-based local voltage control schemes that only use local voltage measurements [5], [6]. For example, voltage droop control in microgrids adjust reactive power injection based on the local voltage deviation from its nominal value. However, it remains as a daunting challenge to guarantee the performance of the control rules, i.e., to stabilize the voltages within the acceptable range under all possible operating conditions.

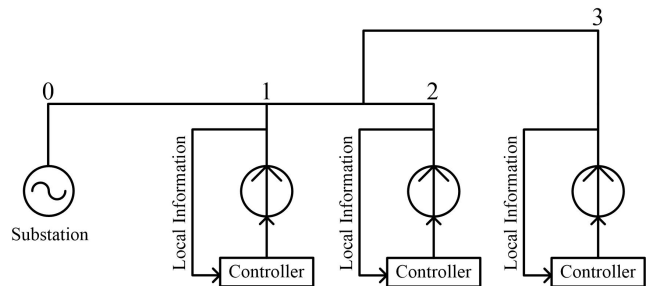


Figure 1: An example of a distribution network with 1 substation and 3 buses. Each bus is equipped with a controller that only takes local information as its input.

In this paper, we focus on voltage control through reactive power compensation.¹ To facilitate the design and analysis of voltage control, we use a linear branch flow model similar to the Simplified DistFlow equations introduced in [7]. The linear branch flow model and the local Volt/Var control form a closed loop dynamical system (Equation (4)). Then we study three types of decentralized voltage control using different local information. The scheme of the decentralized voltage control is shown in Figure 1. Note that there is no communication in

¹Thus, the two terms, voltage control and Volt/Var control, are interchangeable in this paper.

the scheme. In particular, we first show that using only voltage measurements to determine reactive power compensation (no matter whether in a centralized or decentralized way) is insufficient, because there always exist operating conditions under which such voltage control fails to maintain acceptable voltages. This implies that most of existing voltage control algorithms such as [5], [8] that only uses voltage measurements are not able to maintain voltages within the acceptable range under certain circumstances. Then we propose two decentralized algorithms by adding additional information into the control design.² The additional information can be either measured locally or computed locally, meaning that the two algorithms are fully decentralized, requiring no communication. With the aid of the additional information, both of the two algorithms can stabilize voltages in the acceptable range; and the one with higher complexity can further minimize a cost of the reactive power compensation in a particular form. Moreover, the two algorithms are robust against heterogeneous update rates and delays. The heterogeneous update rates means that each control device can have its own time clock to choose when to update their reactive power injection and the clock rates can be different. The delays include both the measurement and the actuation delays that widely exist in real systems. The robustness makes the two algorithms suitable for practical applications. Overall, our results imply that with the aid of right local information, local voltages actually carry out the whole network information for Volt/Var control.

A. Related Work

Research has been proposed and conducted to improve the existing voltage control to mitigate the voltage fluctuation impact. To name a few, [9] studies the active power curtailment to mitigate the voltage rise impact caused by DER; [10] studies the distributed VAR control to minimize power losses and stabilize voltages; [11] reverse-engineers the IEEE 1547.8 standard and study the equilibrium and dynamics of voltage control; [12] proposes two stage voltage control. Compared with the work in this paper, they usually require certain amount of communication or lack theoretical guarantee of performance regarding stabilizing the voltage within the acceptable range. A recent paper [13] proposes a similar algorithm as our algorithm 1. Besides the different approximation model used in [13], the control objective is also different.³ Another similar paper [14] also proposes a linear decentralized control algorithm for voltage control using a similar power flow approximation model to ours. However, their control algorithm only uses voltage measurement, making the algorithm fall into the category of

²For the sake of clarity, we will refer the three algorithms as algorithm 0, 1, 2 correspondingly.

³The objective of the voltage control in [13] is to drive the voltage profile to a predefined point, whereas the objective in this paper is to stabilize the voltage within a range. Also, because of the different approximation models, the theoretical guarantees of [13] and our paper are slightly different. For instance, they show that the convergence depends on the system operating condition, whereas we show that the convergence is regardless of the operating condition. Though we apply the full nonlinear AC model to simulate our algorithms and the simulation results are consistent with the theoretical analysis, it remains as an interesting question about how to choose the power flow approximation model and how good the approximation is.

our algorithm 0, which is insufficient if the objective is to stabilize the voltage within the acceptable range.

In addition to the work mentioned above, in power engineering community, much work has been done in voltage control for microgrids [5], [15], [15]–[19]. See [6] for a comprehensive review. The existing methods generally fall under two layers, primary control (droop control) and secondary control.⁴ The droop control can be viewed as a linear controller in which the reactive power injection is a linear function of the measured voltage. In section III, we will show that this kind of methods do not work under some operation conditions. The secondary control is essentially an integral controller that eliminates the deviation between the voltage and a reference point. This method can be viewed as a special case of algorithm 1 in this paper by shrinking the acceptable range to the reference point.

The remaining of this paper is organized as follows: Section II presents an AC power flow model, its linear approximation, and the formulation of the Volt/Var control; Section II-C illustrates the impossibility result of merely using voltage information in the control; Section IV presents one decentralized robust algorithm to maintain acceptable voltages; Section V presents one decentralized robust algorithm to maintain acceptable voltages and also reach certain optimality as to the reactive power support; Section VI simulates the algorithms to complement our analysis; Section VII concludes the paper.

II. PRELIMINARIES: POWER FLOW MODEL AND PROBLEM FORMULATION

Due to space limit, we introduce here an abridged version of the branch flow model; see, e.g., [20], [21] for more details.

A. Branch flow model for radial networks

Consider a radial distribution circuit that consists of a set N of buses and a set E of distribution lines connecting these buses. We index the buses in N by $i = 0, 1, \dots, n$, and denote a line in E by the pair (i, j) of buses it connects. Bus 0 represents the substation and other buses in N represent branch buses. For each line $(i, j) \in E$, let $I_{i,j}$ be the complex current flowing from buses i to j , $z_{ij} = r_{ij} + \mathbf{i}x_{i,j}$ be the impedance on line (i, j) , and $S_{ij} = P_{ij} + \mathbf{i}Q_{i,j}$ be the complex power flowing from buses i to bus j . On each bus $i \in N$, let V_i be the complex voltage and $s_i = p_i + \mathbf{i}q_i$ be the complex power injection, i.e., the generation minus consumption. As customary, we assume that the complex voltage V_0 on the substation bus is given and fixed at the nominal value.

The branch flow model was first proposed in [1], [2] to model power flows in a steady state in a radial distribution circuit:

$$-p_j = P_{ij} - r_{ij}\ell_{ij} - \sum_{k:(j,k) \in E} P_{jk}, \quad j = 1, \dots, n \quad (1a)$$

⁴There is usually a third layer of microgrid control that focuses on the economical operation of microgrid. This layer is beyond the interest of the discussion here.

$$-q_j = Q_{ij} - x_{ij}\ell_{ij} - \sum_{k:(j,k) \in E} Q_{jk}, \quad j = 1, \dots, n \quad (1b)$$

$$v_j = v_i - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) + (r_{ij}^2 + x_{ij}^2)\ell_{ij}, \quad (i, j) \in E \quad (1c)$$

$$\ell_{ij} = \frac{P_{ij}^2 + Q_{ij}^2}{v_i}, \quad (i, j) \in E, \quad (1d)$$

where $\ell := |I_{ij}|^2$, $v_i := |V_i|^2$. Equations (1) define a system of equations in the variables $(P, Q, \ell, v) := (P_{ij}, Q_{ij}, \ell_{ij}, (i, j) \in E, i = 1, \dots, n)$, which do not include phase angles of voltages and currents. Given an (P, Q, ℓ, v) these phase angles can be uniquely determined for radial networks. This is not the case for mesh networks; see [20] for exact conditions under which phase angles can be recovered for mesh networks.

B. Linear approximation of the branch flow model

Real distribution circuits usually have very small r, x , i.e. $r, x \ll 1$, while $v \sim 1$. Thus real and reactive power losses are typically much smaller than power flows P_{ij}, Q_{ij} . Following [7], we neglect the higher order real and reactive power loss terms in (1) by setting $\ell_{ij} = 0$ and approximate P, Q, v using the following linear approximation, known as Simplified Distflow introduced in [7].

$$-p_j = P_{ij} - \sum_{k:(j,k) \in E} P_{jk}, \quad j = 1, \dots, n \quad (2a)$$

$$-q_j = Q_{ij} - \sum_{k:(j,k) \in E} Q_{jk}, \quad j = 1, \dots, n \quad (2b)$$

$$v_j = v_i - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}), \quad (i, j) \in E \quad (2c)$$

From (2), we can derive that the voltage $v = (v_1, \dots, v_n)^T$ and power injection $p = (p_1, \dots, p_n)$, $q = (q_1, \dots, q_n)$ satisfy the following equation:

$$v = Rp + Xq + v_0 \quad (3)$$

where $R = [R_{ij}]_{n \times n}$, $X = [X_{ij}]_{n \times n}$ are given as follows:

$$R_{ij} := 2 \sum_{(h,k) \in \mathcal{P}_i \cap \mathcal{P}_j} r_{hk},$$

$$X_{ij} := 2 \sum_{(h,k) \in \mathcal{P}_i \cap \mathcal{P}_j} x_{hk}.$$

Here $\mathcal{P}_i \subset E$ is the set of lines on the unique path from bus 0 to bus i . The detailed derivation is given in [22]. Since $r_{ij} > 0$, $x_{ij} > 0$ for all i, j , R, X have the following properties.

Lemma 1. R, X are positive definite and positive matrices.⁵

Proof. We refer readers to [22] for the detailed proof. \square

C. Problem formulation

Before rigorously formulating the Volt/Var control problem, we separate q into two parts, $q^c = (q_1^c, \dots, q_n^c)$ and $q^e = (q_1^e, \dots, q_n^e)$, where q^c denotes the reactive power injection governed by the Volt/Var control components and q^e denotes

any other reactive power injection.⁶ Let $v^{par} \triangleq Rp + Xq^e + v_0$, then,

$$v = Xq^c + v^{par}.$$

The goal of Volt/Var control on a distribution network is to provision reactive power injections q^c to maintain the bus voltages v within a tight range $[v, \bar{v}]$ under any operating condition given by v^{par} . Without causing any confusion, in the rest of the paper, we will simply use q_i instead of q_i^c to denote the reactive power pulled by the Volt/Var control devices. The Volt/Var control can be modeled as a control problem on a quasi-dynamical system with state v and controller q ; that is, given the current state $v(t)$ and other available information, the controller determines a new reactive power injections $q(t)$ and the new $q(t)$ results in a new voltage profile $v(t+1)$ according to (3). Mathematically, the Volt/Var control problem is formulated as the following closed loop dynamical system,

$$v(t+1) = Xq(t) + v^{par}; \quad (4a)$$

$$q(t) = u(\text{information at time } t). \quad (4b)$$

where $u = (u_1, \dots, u_n)$ is the Volt/Var controller. The objective of Volt/Var control is to design u to drive the system voltage $v(t)$ to the acceptable range $[v, \bar{v}]$ under any system operating condition which is given by v^{par} . Rigorously speaking, it requires that

$$\lim_{t \rightarrow \infty} \text{dist}(v(t), [v, \bar{v}]) = 0.$$

Here $\text{dist}(y, Z) := \min_{z \in Z} \|y - z\|$ where y is a point and Z is a set. Note that (4a) is governed by the system intrinsic dynamics (Kirchoff's Law) and not able to controlled or tuned, which makes the Volt/Var control challenging.

The problem we address in this paper is to design voltage control algorithm u where each u_i only uses local information, as shown in Figure 1. In general, in electricity distribution networks, only a small portion of buses are monitored, individuals are unlikely to announce their generation or load profile, the availability of DERs are fluctuating and uncertain, and even the grid parameters and the topology are only partially known. All of these facts make decentralized algorithms necessary for the voltage control, i.e., each control component adjusts its reactive power input based on the local signals that are easy to measure or to communicate.

III. VOLTAGE CONTROL USING ONLY VOLTAGE MEASUREMENTS: IMPOSSIBILITY RESULT

We first study such Volt/Var control rules that use merely voltage measurements as the control information. This type of control has been proposed and discussed in many existing literature and applications [6]. For example, the IEEE 1547.8 standard [8] proposes decentralized voltage control for inverters using the deviation of the local voltages from the nominal value: an inverter monitors its terminal voltage and sets its reactive power generation based on a static and predefined Volt/Var curve. If the predefined Volt/Var curve is linear, the

⁶For easy exposition, we assume that there is a q_i^e at each bus i . But the algorithm extends to the scenario that only a subset of buses have Volt/Var control component.

⁵A matrix is a positive matrix iff each item is positive.

controller is called a droop controller. Regarding a detailed theoretic analysis of droop controllers, please see [14]. The scheme of such controllers is shown in Figure 2. Besides the IEEE 1547.8 standard, there are other researches promoting adapting reactive power injection according to the voltage [5], [9], [14]. However, in the following, we show that this type of controller is insufficient to drive the voltage to the acceptable range even if controller stabilizes the system.

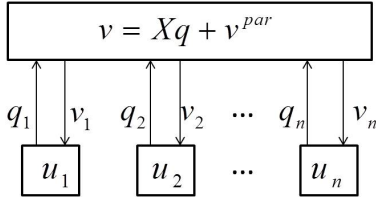


Figure 2: IEEE 1547.8 standard: Decentralized Volt/Var control using local voltage measurements.

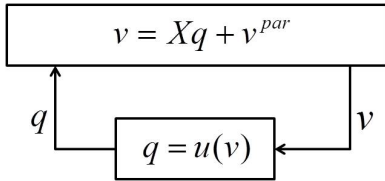


Figure 3: Volt/Var control using merely voltage information: This type of controller is insufficient for Volt/Var control.

In fact, we will demonstrate that as long as u is in the following form,

$$q(t) = u(v(t)), \quad (5)$$

and u maps the bounded set $[\underline{v}, \bar{v}]$ to a bounded set, then it is impossible for such u to maintain acceptable voltages under all the operating condition, no matter whether u is in a centralized or decentralized form. This is formally stated in the following proposition.

Proposition 2. *For any u in the form of (5) that maps $[\underline{v}, \bar{v}]$ to a bounded set, there exist v^{par} such that this controller is not able to stabilize the voltage v in the acceptable range $[\underline{v}, \bar{v}]$.⁷*

Proof. The proof is straightforward. Substituting (5) into (4a), we have:

$$v(t+1) - Xu(v(t)) = v^{par}.$$

Given a v^{par} , if u stabilizes the voltage in the acceptable range $[\underline{v}, \bar{v}]$, i.e., $\lim_{t \rightarrow \infty} \text{dist}(v(t), [\underline{v}, \bar{v}]) = 0$, then v^{par} should be at least in the set of $M := \{v - Xu(\bar{v}) : v, \bar{v} \in [\underline{v}, \bar{v}]\}$ which is bounded because u maps $[\underline{v}, \bar{v}]$ to a bounded set. Thus we know there exist v^{par} that the controller is not able to stabilize the voltage in the acceptable range. \square

This proposition tells us that almost any Volt/Var control that depends merely on voltage information is not suitable to maintain acceptable voltages, no matter whether it is decentralized or centralized. Thus we should consider adding (or using) other information to design controller u . In the rest of

⁷Note that any continuous function maps $[\underline{v}, \bar{v}]$ into a bounded set.

the paper, we will show that if we use both the information of the current q and v , then a fully decentralized algorithm in the form of $u_i(v_i(t), q_i(t-1))$ is able to maintain acceptable voltages; further, if we introduce some auxiliary variables $\lambda_i, \forall i$, then a fully decentralized algorithm in the form of $u_i(v_i(t), \lambda_i(t-1))$ is able to both maintain acceptable voltages and minimize a cost of reactive power compensation in a particular form.

IV. A DECENTRALIZED ALGORITHM TO REACH THE ACCEPTABLE VOLTAGE RANGE

In this section, we will show that by using both the local voltage and reactive power information, a fully decentralized algorithm in the form of $q_i(t) = u(v_i(t), q_i(t-1))$ is able to stabilize the voltage within the acceptable range. The scheme of the algorithm is shown in Figure 4. Moreover, this algorithm is actually robust to heterogeneous update rates and delays. Firstly, each control device is free to choose when to update the reactive power injection; secondly, it is allowed to have measurement and actuation delay. This robustness makes the algorithm suitable for practical applications because in practice different buses may have different control devices which operate at different time rates and delay widely exists in real systems.

We now present the first algorithm which we name as algorithm 1 and then prove its convergence.

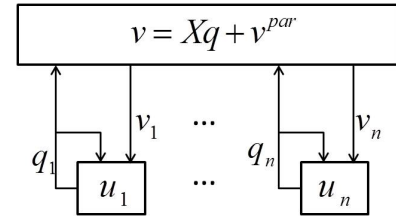


Figure 4: Volt/Var control using local information of local voltage and reactive power injection: This type of controller is able to maintain acceptable voltages.

A. Algorithm 1

For each bus i , we introduce a infinite time step set $T_i \subseteq \{1, 2, \dots\}$ to represent the time steps when bus i updates the reactive power rejection.

For $t \notin T_i$, no new decision will be made, which means the reactive injection stays the same as the previous time step, i.e.

$$q_i(t) = q_i(t-1) \quad (6)$$

For $t \in T_i$, bus i updates according to the following update rule,

$$q_i(t) = q_i(t-1) - \epsilon d_i(v_i(t - \tau_i(t))) \quad (7)$$

where given a voltage v_i ,

$$d_i(v_i) := [v_i - \bar{v}_i]^+ - [\underline{v}_i - v_i]^+ \quad (8)$$

and ϵ is a positive constant stepsize. Here $[\cdot]^+$ is defined as $[a]^+ := \max(a, 0)$. $\tau_i(t) \geq 0$ is the measurement delay of bus

i at t . $v_i(t - \tau_i(t))$ represents the voltage measurement bus i received from a voltage sensor at time t , which is the actual voltage $\tau_i(t)$ steps before t .⁸

The algorithm in (7) says that if the voltage at bus i exceeds its upper limit, bus i decreases the reactive power injection; in contrast, if the local voltage falls below its lower limit, bus i increases its reactive power injection. This algorithm is very simple and intuitive, yet we will prove that regardless of v^{par} and despite the heterogeneous update rates and the delays, the voltage will asymptotically converge to the acceptable range $[\underline{v}, \bar{v}]$ under the controller specified in (6) and (7). Before that, we discuss the properties of this algorithm which make it attractive to real-time and scalable implementation.

- 1) We note that in the algorithm, each bus i uses only the local voltage measurement $v_i(t - \tau_i(t))$ and its previous reactive power injection $q_i(t-1)$. This algorithm is fully decentralized, requiring no communication.
- 2) The controller (7) is similar to an integral controller. The implementation is simple.
- 3) The algorithm does not require any information about the system operating condition, represented by v^{par} . This makes the algorithm practical because due to the large volatility and uncertainty of renewable energy, time-varying nature of demand, and the privacy concern of consumers, v^{par} is not available and fluctuates fast and by a large amount.
- 4) Though the convergence of the algorithm depends on the value of the step size ϵ , as shown in the next section, the criterion for choosing ϵ to guarantee the convergence is independent of v^{par} . As a result, once we have incorporated the algorithm into the hardware design of Volt/Var control, the Volt/Var control will work under any system operating operation.

Before we proceed to the convergence result, we make the following assumption on the update time steps T_i and the measurement delays $\tau_i(t)$.

Assumption 1. For any i , the difference between consecutive update time steps in T_i is upper bounded by T_a . For any i and t , the time delay $\tau_i(t)$ is upper bounded by T_d .

The first part of the assumption means that each bus should update at least once within T_a steps; the second part means that the delays are upper bounded by T_d . This assumption is mild and reasonable, and can be easily guaranteed in practice. Under this assumption, we have the following convergence result.

Theorem 3. For Algorithm 1 in (6) and (7), if Assumption 1 holds and $\epsilon < \frac{2}{\sigma_{max}(X) + 2\|X\|_F T_d}$, then $d(v(t))$ will converge to 0, and $v(t)$ will converge to a point $v^* \in [\underline{v}, \bar{v}]$.

From Theorem 3 we can see that the maximum delay time T_d contributes a linear term in the denominator of the maximum stepsize that guarantees the convergence. The slope of the linear term is the Frobenius norm of X , $\|X\|_F$. The theorem shows that the larger $\|X\|_F$ is (which is usually a

⁸When bus i is not updating at t , we would abuse the notation $\tau_i(t)$ to be an arbitrary nonnegative integer despite it will have no effect on the algorithm.

result of a larger network) or the larger T_d is, the smaller the stepsize should be.

B. Proof of the convergence

In this section, we prove that the voltage asymptotically converges to the acceptable range $[\underline{v}, \bar{v}]$ under the controller specified in (6) and (7). We first introduce a Lyapunov function ψ , which will decrease along the trajectory of the control algorithm. Then, using a similar technique as in [23], we will show that $\|q(t) - q(t-1)\| \rightarrow 0$. Combining this with the fact that each bus has to update at least once within T_a steps, we will reach the conclusion that $d(v(t))$ asymptotically converges to 0.

We first introduce some notations. Let $\alpha(t) = q(t) - q(t-1)$. Define a n -vector $\hat{d}(v) = v - \frac{\bar{v} + \underline{v}}{2}$, i.e. its i th component is $\hat{d}_i(v_i) = v_i - \frac{\bar{v}_i + \underline{v}_i}{2}$. Define a Lyapunov function $\psi(q) = \frac{1}{2} \hat{d}(v(q))^T X^{-1} \hat{d}(v(q))$ where $v(q) = Xq + v^{par}$. Clearly, $\nabla \psi(q) = \hat{d}(v(q))$ and $\nabla^2 \psi(q) = X$. Define $\gamma(t)$ to be a n -vector with its i th component as $\gamma_i(t) = \hat{d}_i(v_i(t - \tau_i(t)))$.

Lemma 4. $\|\gamma(t) - \nabla \psi(q(t-1))\| \leq \|X\|_F \sum_{t'=t-T_d}^{t-1} \|\alpha(t')\|$

Proof. For each i ,

$$\begin{aligned} |\gamma_i(t) - \nabla_i \psi(q(t-1))| &= |\hat{d}_i(v_i(t - \tau_i(t))) - \hat{d}_i(v_i(t))| \\ &= |v_i(t - \tau_i(t)) - v_i(t)| \\ &= |X_i^T(q(t-1) - q(t-1 - \tau_i(t)))| \\ &\leq \|X_i\| \|q(t-1) - q(t-1 - \tau_i(t))\| \\ &= \|X_i\| \left\| \sum_{t'=t-\tau_i(t)}^{t-1} \alpha(t') \right\| \\ &\leq \|X_i\| \sum_{t'=t-\tau_i(t)}^{t-1} \|\alpha(t')\| \\ &\leq \|X_i\| \sum_{t'=t-T_d}^{t-1} \|\alpha(t')\| \end{aligned}$$

where X_i is the i th row of X . Sum over i and the lemma follows. \square

Lemma 5.

$$\gamma(t)^T \alpha(t) = -\frac{1}{\epsilon} \|\alpha(t)\|^2 - \sum_{i=1}^n M_i |\alpha_i(t)|$$

where $M_i := \frac{\bar{v}_i - \underline{v}_i}{2}$.

Proof. For each i , $\gamma_i(t) = \hat{d}_i(v_i(t - \tau_i(t)))$. For notational simplicity, we denote $t - \tau_i(t)$ by t' . If bus i updates at t , then $\alpha_i(t) = -\epsilon d_i(v_i(t'))$. If $d_i(v_i(t')) > 0$, then $\hat{d}_i(v_i(t')) = d_i(v_i(t')) + \frac{\bar{v}_i - \underline{v}_i}{2}$ and $\hat{d}_i(v_i(t')) d_i(v_i(t')) = d_i(v_i(t'))^2 + \frac{\bar{v}_i - \underline{v}_i}{2} d_i(v_i(t'))$, i.e. $\gamma_i(t) \alpha_i(t) = (-\frac{1}{\epsilon}) \alpha_i(t)^2 - \frac{\bar{v}_i - \underline{v}_i}{2} |\alpha_i(t)|$. If $d_i(v_i(t')) < 0$, then $\hat{d}_i(v_i(t')) = d_i(v_i(t')) - \frac{\bar{v}_i - \underline{v}_i}{2}$ and hence $\hat{d}_i(v_i(t')) d_i(v_i(t')) = d_i(v_i(t'))^2 - \frac{\bar{v}_i - \underline{v}_i}{2} d_i(v_i(t'))$, i.e. $\gamma_i(t) \alpha_i(t) = (-\frac{1}{\epsilon}) \alpha_i(t)^2 - \frac{\bar{v}_i - \underline{v}_i}{2} |\alpha_i(t)|$. If $d_i(v_i(t')) = 0$ or bus i does not update, we also have, $\gamma_i(t) \alpha_i(t) =$

$(-\frac{1}{\epsilon})\alpha_i(t)^2 - \frac{\bar{v}_i - \underline{v}_i}{2}|\alpha_i(t)|$ because both sides of the equality are zero. Sum the equality over i ,

$$\begin{aligned}\gamma(t)^T \alpha(t) &= \sum_{i=1}^n \left[-\frac{1}{\epsilon} \alpha_i(t)^2 - \frac{\bar{v}_i - \underline{v}_i}{2} |\alpha_i(t)| \right] \\ &= -\frac{1}{\epsilon} \|\alpha(t)\|^2 - \sum_{i=1}^n M_i |\alpha_i(t)|\end{aligned}$$

Now we are ready to prove Theorem 3.

Proof. With Lemma 4 and Lemma 5 we have

$$\begin{aligned}& \psi(q(t)) - \psi(q(t-1)) \\ & \leq \nabla \psi(q(t-1))^T \alpha(t) + \frac{1}{2} \alpha(t)^T \nabla^2 \psi(q(t-1)) \alpha(t) \\ & \leq \|\nabla \psi(q(t-1)) - \gamma(t)\| \|\alpha(t)\| + \gamma(t)^T \alpha(t) + \frac{1}{2} \alpha(t)^T X \alpha(t) \\ & \leq \|X\|_F \sum_{t'=t-T_d}^{t-1} \|\alpha(t')\| \|\alpha(t)\| \\ & \quad - \sum_{i=1}^n M_i |\alpha_i(t)| + \left(-\frac{1}{\epsilon} + \frac{1}{2} \sigma_{\max}(X)\right) \|\alpha(t)\|^2 \\ & \leq \frac{1}{2} \|X\|_F \sum_{t'=t-T_d}^{t-1} (\|\alpha(t')\|^2 + \|\alpha(t)\|^2) \\ & \quad - \sum_{i=1}^n M_i |\alpha_i(t)| + \left(-\frac{1}{\epsilon} + \frac{1}{2} \sigma_{\max}(X)\right) \|\alpha(t)\|^2\end{aligned}$$

Sum from $t = 1$ to $t = T$, we have

$$\begin{aligned}& \psi(q(T)) - \psi(q(0)) \\ & \leq \frac{1}{2} \|X\|_F \sum_{t=1}^T \sum_{t'=t-T_d}^{t-1} (\|\alpha(t')\|^2 + \|\alpha(t)\|^2) \\ & \quad + \left(-\frac{1}{\epsilon} + \frac{1}{2} \sigma_{\max}(X)\right) \sum_{t=1}^T \|\alpha(t)\|^2 - \sum_{t=1}^T \sum_{i=1}^n M_i |\alpha_i(t)| \\ & = \frac{1}{2} \|X\|_F \sum_{t=1}^T \sum_{t'=t-T_d}^{t-1} \|\alpha(t')\|^2 - \sum_{t=1}^T \sum_{i=1}^n M_i |\alpha_i(t)| \\ & \quad + \left(-\frac{1}{\epsilon} + \frac{1}{2} \sigma_{\max}(X) + \frac{T_d}{2} \|X\|_F\right) \sum_{t=1}^T \|\alpha(t)\|^2 \\ & \leq \left(-\frac{1}{\epsilon} + \frac{1}{2} \sigma_{\max}(X) + T_d \|X\|_F\right) \sum_{t=1}^T \|\alpha(t)\|^2 \\ & \quad - \sum_{t=1}^T \sum_{i=1}^n M_i |\alpha_i(t)|\end{aligned}\tag{9}$$

Since $\psi(q(T))$ is lower bounded and $M_i \geq 0$, if $\epsilon < \frac{2}{\sigma_{\max}(X) + 2\|X\|_F T_d}$, $\sum_{t=1}^T \|\alpha(t)\|^2$ is upper bounded, hence $\alpha(t) \rightarrow 0$. For each i , denote the elements in T_i by an increasing sequence $\{t_i(n)\}_{n=1}^\infty$. Denote $t'_i(n) = t_i(n) - \tau_i(t_i(n))$, then by (7), $d_i(v_i(t'_i(n))) \rightarrow 0$ as $n \rightarrow \infty$. For each t , by assumption 1, $\exists n_t \in \mathbb{N}$ s.t. $t - T_a - T_d \leq t'_i(n_t) \leq t$. So

$$\begin{aligned}|d_i(v_i(t)) - d_i(v_i(t'_i(n_t)))| &\leq |v_i(t) - v_i(t'_i(n_t))| \\ &\leq \|X_i\| \|q(t-1) - q(t'_i(n_t) - 1)\|\end{aligned}$$

$$\leq \|X_i\| \sum_{t'=t-T_a-T_d}^{t-1} \|\alpha(t')\| \tag{10}$$

Where the first inequality follows from the fact that $d_i(v)$ is a Lipschitz function with Lipschitz coefficient 1. Let $t \rightarrow \infty$, then $n_t \rightarrow \infty$, so $d_i(v_i(t'_i(n_t))) \rightarrow 0$. Also notice the right hand side of (10) converges to 0 (because it's a finite sum of sequences each of which converges to 0), we have $d_i(v_i(t)) \rightarrow 0$. This holds for all i , so $\|d(v(t))\| \rightarrow 0$. \square

For each i , if $M_i > 0$ (i.e. $\bar{v}_i > \underline{v}_i$), by (9) $\sum_{t=1}^T |\alpha_i(t)|$ is upper bounded, hence this series converges. This says $|q_i(T+k) - q_i(T)| \leq \sum_{t=T+1}^{T+k} |\alpha_i(t)| \rightarrow 0$ as $T \rightarrow \infty$. So $q_i(t)$ is a Cauchy sequence and it converges to a point q_i^* [24]. For i that $M_i = 0$ (i.e. $\bar{v}_i = \underline{v}_i = v_i^*$), we have $\hat{d}_i(v_i) = d_i(v_i)$, so $\hat{d}_i(t)$ converges to 0, i.e. $v_i(t)$ converges to v_i^* . Now we have for each i either $q_i(t)$ converges or $v_i(t)$ converges. Without loss of generality, we assume for $i = 1, \dots, m$, $q_i(t)$ converges and for $i = m+1, \dots, n$, $v_i(t)$ converges. Define a vector $z(t) := (v_1(t), \dots, v_m(t), q_{m+1}(t-1), \dots, q_n(t-1))^T - (v_1^{par}, \dots, v_m^{par}, 0, \dots, 0)^T$. Due to the linear relationship between $v(t)$ and $q(t-1)$, we have:

$$z(t) = \begin{pmatrix} I_m & 0 \\ X_m^1 & X_m^2 \end{pmatrix} q(t-1) := \Lambda q(t-1)$$

where I_m is the m -dimensional identity matrix, X_m^1 and X_m^2 together make the $m+1$'s through the n 's row of X . Since X is positive definite, we have X_m^2 is invertible, which leads to Λ is invertible. Since $z(t)$ converges, we have $q(t)$ converges, and further $v(t)$ converges to a point v^* . Since $\|d(v(t))\| \rightarrow 0$, v^* lies in the acceptable range. \square

V. A DECENTRALIZED ALGORITHM TO REACH AN OPTIMAL FEASIBLE POINT

In this section, we will introduce local auxiliary variables and design a new control algorithm, which we name as algorithm 2. Algorithm 2 will not only drive the voltage into an equilibrium point in $[\underline{v}, \bar{v}]$, but also guarantee that the equilibrium point will minimize a cost of reactive power provision in a certain form. We first provide the algorithm and then discuss its convergence and the optimality of the equilibrium point.

As in algorithm 1, we use the set T_i to represent the time steps that bus i updates its reactive power injection. For each bus i , we introduce two local auxiliary variables, $\bar{\lambda}_i$ and $\underline{\lambda}_i$.

At each time t , if $t \notin T_i$, all the variables stay the same as the previous time step:

$$\bar{\lambda}_i(t) = \bar{\lambda}_i(t-1) \tag{11a}$$

$$\underline{\lambda}_i(t) = \underline{\lambda}_i(t-1) \tag{11b}$$

$$q_i(t) = q_i(t-1) \tag{11c}$$

If $t \in T_i$,

$$\bar{\lambda}_i(t) = [\bar{\lambda}_i(t-1) + \epsilon(v_i(t - \tau_i(t)) - \bar{v}_i)]^+ \tag{12a}$$

$$\underline{\lambda}_i(t) = [\underline{\lambda}_i(t-1) + \epsilon(\underline{v}_i - v_i(t - \tau_i(t)))]^+ \tag{12b}$$

$$q_i(t) = \underline{\lambda}_i(t - \tau'_i(t)) - \bar{\lambda}_i(t - \tau'_i(t)) \tag{12c}$$

where $\tau_i(t)$ and $\tau'_i(t)$ capture the measurement, computation, and actuation delay.⁹ The diagram of the algorithm is shown in Figure 5.

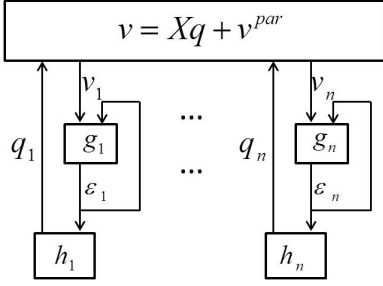


Figure 5: Volt/Var control using local information of voltage, reactive power injection, and additional auxiliary variables: This type of controller is able to guarantee both of feasibility and optimality.

In this algorithm, the control of the local reactive power depends only on the two local auxiliary variables and the updating rule of the two variables depends on their own values and the current local voltage measurement, requiring no communication and no information about v^{par} . Moreover, the updating rule is also similar to an integral controller with saturation. As a result, this algorithm shares the same properties with algorithm 1, making it also attractive to practical applications.

Before we proceed to the convergence analysis of algorithm 2, we point out its relation to the dual gradient algorithm of an optimization problem. To simplify illustration, we consider the following synchronous and delay-free version of algorithm 2:

$$\bar{\lambda}_i(t) = [\bar{\lambda}_i(t-1) + \epsilon(v_i(t) - \bar{v}_i)]^+ \quad (13a)$$

$$\underline{\lambda}_i(t) = [\underline{\lambda}_i(t-1) + \epsilon(\underline{v}_i - v_i(t))]^+ \quad (13b)$$

$$q_i(t) = \underline{\lambda}_i(t) - \bar{\lambda}_i(t) \quad (13c)$$

We have the following theorem regarding the convergence of algorithm (13):

Theorem 6. *If $\epsilon < \frac{1}{\sigma_{\max}(X)}$, $q(t)$ in algorithm (13) converges to the optimal point q^* of the following optimization problem:*

$$\min_q \quad \frac{1}{2} q^T X q \quad (14a)$$

$$s.t. \quad Xq + v^{par} \leq \bar{v}, \quad (14b)$$

$$Xq + v^{par} \geq \underline{v}, \quad (14c)$$

and $v(t)$ converges to the corresponding voltage $v^* = Xq^* + v^{par}$, which lies in the acceptable range $[\underline{v}, \bar{v}]$.

Proof. Introducing dual variable $\bar{\lambda}$ for (14b) and $\underline{\lambda}$ for (14c), we have the following dual gradient algorithm [25]:

$$\bar{\lambda}(t) = [\bar{\lambda}(t-1) + \epsilon(Xq(t-1) + v^{par} - \bar{v})]^+$$

$$\underline{\lambda}(t) = [\underline{\lambda}(t-1) + \epsilon(\underline{v} - Xq(t-1) - v^{par})]^+$$

$$q(t) = \arg \min_q \left\{ \frac{1}{2} q^T X q + \bar{\lambda}(t)(Xq + v^{par} - \bar{v}) \right.$$

⁹Similar to Algorithm 1, for $t \notin T_i$ we abuse the notation $\tau_i(t)$ and $\tau'_i(t)$ to be any nonnegative integers.

$$+ \underline{\lambda}(t)(\underline{v} - Xq - v^{par}) \}$$

The preceding algorithm is equivalent to the following:

$$\bar{\lambda}(t) = [\bar{\lambda}(t-1) + \epsilon(v(t) - \bar{v})]^+$$

$$\underline{\lambda}(t) = [\underline{\lambda}(t-1) + \epsilon(\underline{v} - v(t))]^+$$

$$q(t) = \underline{\lambda}(t) - \bar{\lambda}(t)$$

which is exactly the algorithm (13).

Through simple derivation, the dual problem of (14) is given by:

$$\max_{\bar{\lambda}, \underline{\lambda} \geq 0} D = -\frac{1}{2}(\underline{\lambda} - \bar{\lambda})^T X(\underline{\lambda} - \bar{\lambda}) + \bar{\lambda}^T(v^{par} - \bar{v}) + \underline{\lambda}^T(\underline{v} - v^{par}) \quad (15)$$

By Lemma 8 in Appendix, we have $\sigma_{\min}(\nabla^2 D) = -2\sigma_{\max}(X)$. Therefore, we know that if $\epsilon < \frac{1}{\sigma_{\max}(X)}$, $\bar{\lambda}(t), \underline{\lambda}(t)$ converge to the dual optimum. The statement of the theorem then follows. \square

The preceding theorem shows that algorithm 2 is the dual gradient algorithm for (14) with heterogeneous update rates and delays. Due to this connection, it is intuitive that despite the heterogeneous updates and the delays, algorithm 2 shall still converge, although the convergence might require tighter bounds on ϵ . Before presenting the convergence of algorithm 2, we require the following mild assumption:

Assumption 2. *$\forall i$, the difference between consecutive update time steps in T_i is upper bounded by T_a . $\forall i$ and $\forall t$, $\tau_i(t)$ and $\tau'_i(t)$ are upper bounded by T_a .*

Theorem 7. *In algorithm 2 (11) and (12), if Assumption 2 holds and $\epsilon < \frac{1}{\sigma_{\max}(X) + 2\|X\|_F(2T_d + T_a)}$, $q(t)$ and $\lambda(t)$ will converge to the optimizer of the primal problem (14) and the dual problem (15) respectively, and $v(t)$ will converge to a point $v^* \in [\underline{v}, \bar{v}]$.*

Different from Theorem 3, in Theorem 7, T_a contributes a linear term in the denominator of the maximum stepsize. This means that the heterogeneous update rates will have a negative effect on choosing the stepsize of algorithm 2.

VI. CASE STUDY

In this section we evaluate algorithm 1 and algorithm 2 on a distribution circuit of South California Edison with a high penetration of photovoltaic (PV) generation [26]. Figure 6 shows a 56-bus distribution circuit. Note that Bus 1 indicates the substation, and there are 1 photovoltaic (PV) generators located at bus 45 and there are shunt capacitors located at bus 19, 21, 30, 53. See [26] for the network data including the line impedance, the peak MVA demand of the loads and the nameplate capacity of the shunt capacitors and the photovoltaic generation.

In the simulation, we assume that there are Volt/Var control components at bus 19, 21, 30, 45, and 53 and those control components can pull in (supply) and out (consume) reactive power. The nominate voltage magnitude is 12kV and the acceptable range is set as [11.4kV, 12.6kV] which is the plus/minus 5% of the nominate value. Though the analysis of this paper is built on the linearized power flow model (2), we

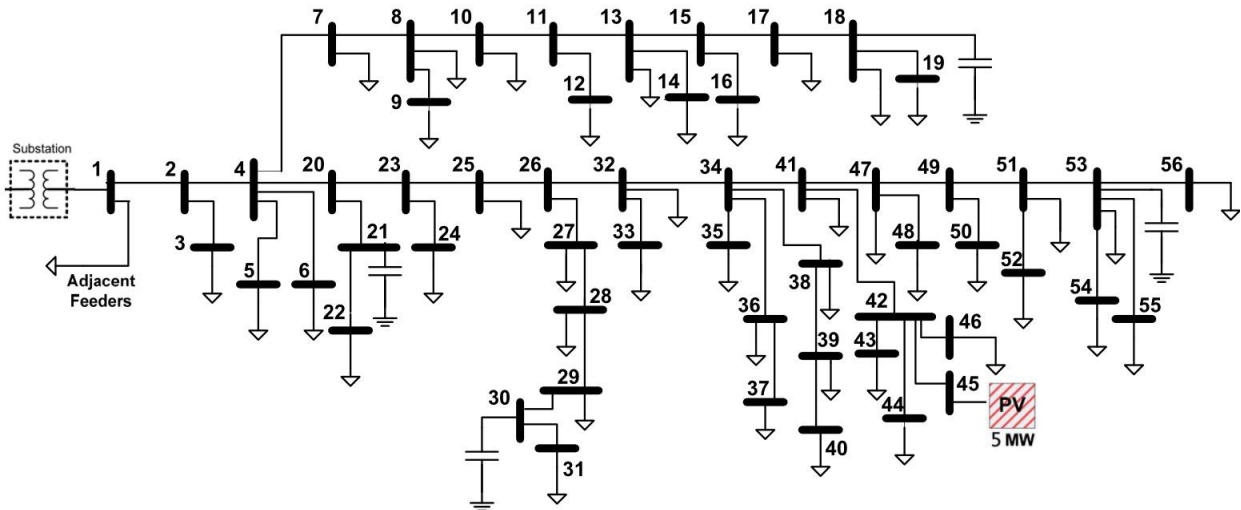


Figure 6: Schematic diagram of two SCE distribution systems.

simulate the voltage control algorithms using the full nonlinear AC power flow model (1) [27].

We simulate three different scenarios: 1) the PV generator is generating a large amount of power but the loads are moderate, resulting in high voltages at some buses (Figure 7 and Figure 8); 2) the PV generator is generating a large amount of power and some buses are having heavy loads, resulting in high voltages at some buses and low voltages at other buses (Figure 9 and Figure 10). In addition to the above two scenarios, we consider a third scenario where the load and the PV generation are constantly fluctuating, drawn from a uniform distribution (Figure 11). The fluctuating speed is considered to be slower than the control algorithm updating speed, which means the controller has a certain amount of time steps to stabilize the voltages after each change in the load and the PV generation. This scenario is introduced to validate the performance of the algorithms under a more realistic setting. The results of case I and case II are summarized in Figure 7 to Figure 10. In the four figures, the buses update at heterogeneous clock rates and with measurement and actuation delay. Due to the delay, buses might make incorrect updates (e.g., in Figure 8, at $t = 20$, all the voltages are already within the acceptable range, but the buses drive the voltages outside of the range). However, despite the heterogeneous update clock rates and the delay, the algorithms still converge asymptotically. Figure 11 verifies the algorithm performance under a fluctuating load and PV generation setting. All the results verify the proposed algorithms' effectiveness.

VII. CONCLUSION

In this paper, we study how different information structures affect the performance of Volt/Var control. In particular, we first show that using only voltage measurements to decide reactive power injection is insufficient to maintain acceptable voltages. Then we propose two robust and fully decentralized algorithms by adding additional information into the control design. Both of them can maintain acceptable voltages; but one is also able to optimize the reactive power injection in terms

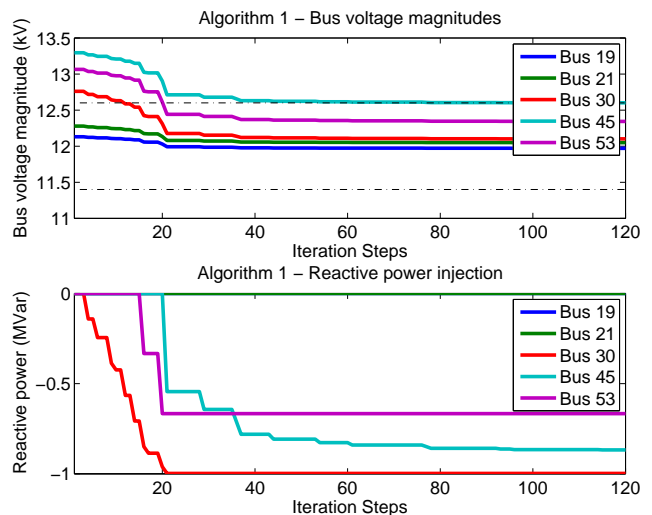


Figure 7: Simulation result of case I, algorithm I. The upper figure shows the voltage profile, and the lower figure shows the reactive power injection. Simulation parameters: $T_a = 25$, $T_d = 15$, $\epsilon = 8$.

of minimizing a cost of the reactive power compensation. Both of the two algorithms use only local measurements and local variables, requiring no communication. They are also robust against heterogeneous update rates and delays. Our results imply that with the aid of right local information, local voltages carry out the whole network information for Volt/Var control.

Regarding future work, there are two additional objectives yet to achieve. First, we have shown that algorithm 2 minimizes a particular form of reactive power cost. However, it would be more desirable if a more general form of cost function can be minimized by the voltage control algorithm. Second, in this paper we do not consider the capacity limit of reactive power injection at each controller, while in practice, we have limit on the reactive power injections, i.e., actuator

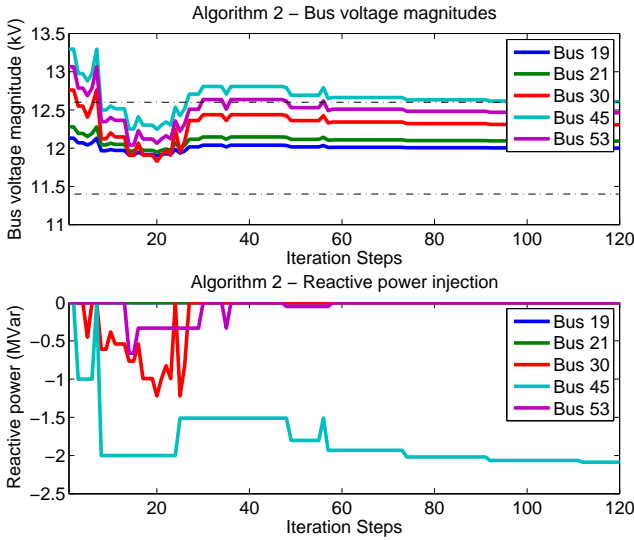


Figure 8: Simulation result of case I, algorithm II. The upper figure shows the voltage profile, and the lower figure shows the reactive power injection. Simulation parameters: $T_a = 25$, $T_d = 15$, $\epsilon = 5$.

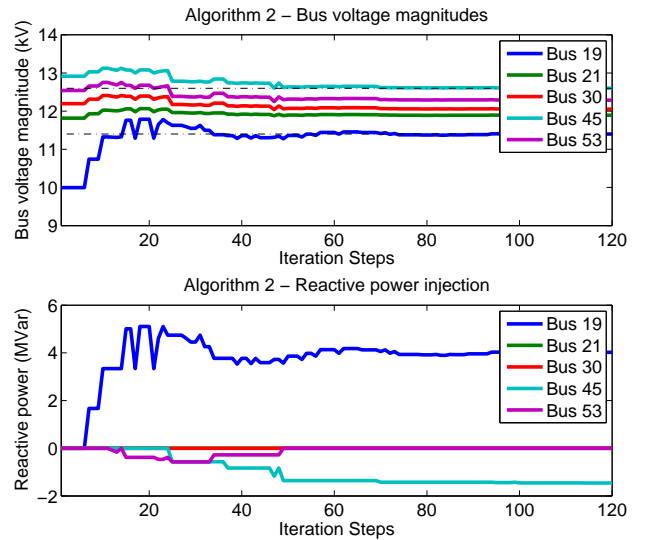


Figure 10: Simulation result of case II, algorithm II. The upper figure shows the voltage profile, and the lower figure shows the reactive power injection. Simulation parameters: $T_a = 25$, $T_d = 15$, $\epsilon = 5$.

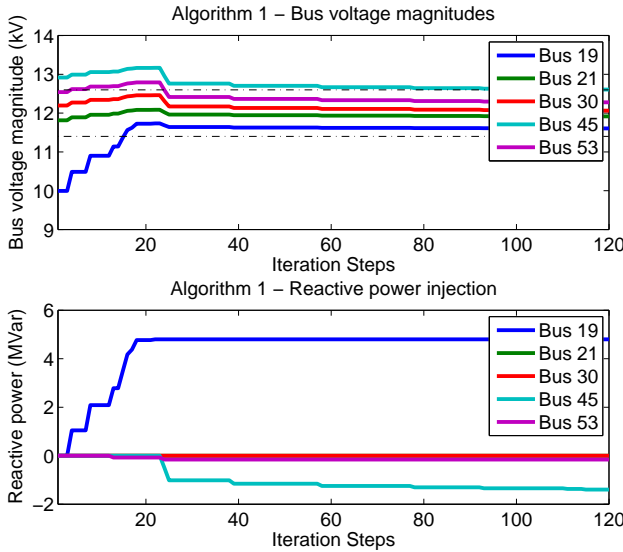


Figure 9: Simulation result of case II, algorithm I. The upper figure shows the voltage profile, and the lower figure shows the reactive power injection. Simulation parameters: $T_a = 25$, $T_d = 15$, $\epsilon = 8$.

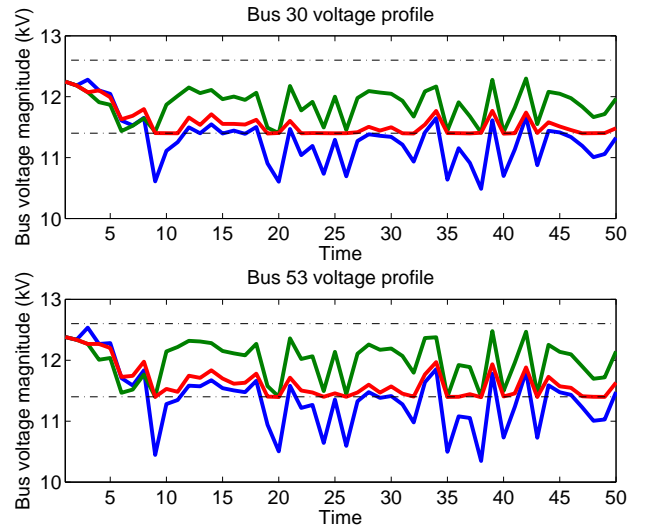


Figure 11: Fluctuating load and PV generation. The voltage profile in this figure is the stabilized voltage after each load and generation change. The upper figure shows the voltage profile of bus 30, and the lower figure shows the voltage profile of bus 53. The blue lines are the voltage profile under no reactive power regulation; the green lines algorithm 1; the blue lines algorithm 2.

saturation. It remains as an open question how to design real-time distributed voltage control with provable performance given the capacity limit of reactive power injection.

REFERENCES

- [1] M. E. Baran and F. F. Wu, "Optimal Capacitor Placement on radial distribution systems," *IEEE Trans. Power Delivery*, vol. 4, no. 1, pp. 725–734, 1989.
- [2] —, "Optimal Sizing of Capacitors Placed on A Radial Distribution System," *IEEE Trans. Power Delivery*, vol. 4, no. 1, pp. 735–743, 1989.

- [3] J. Smith, W. Sunderman, R. Dugan, and B. Seal, "Smart inverter volt/var control functions for high penetration of pv on distribution systems," in *Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES*, 2011, pp. 1–6.
- [4] 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, 2011.
- [5] J. Vasquez, J. Guerrero, M. Savaghebi, J. Eloy-Garcia, and R. Teodorescu, "Modeling, analysis, and design of stationary-reference-frame

droop-controlled parallel three-phase voltage source inverters,” *Industrial Electronics, IEEE Transactions on*, vol. 60, no. 4, pp. 1271–1280, April 2013.

- [6] A. Bidram and A. Davoudi, “Hierarchical structure of microgrids control system,” *Smart Grid, IEEE Transactions on*, vol. 3, no. 4, pp. 1963–1976, Dec 2012.
- [7] M. Baran and F. Wu, “Network reconfiguration in distribution systems for loss reduction and load balancing,” *IEEE Trans. on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr 1989.
- [8] “P1547.8—recommended practice for establishing methods and procedures that provide supplemental support for implementation strategies for expanded use of iecce standard 1547,” *Approved & Proposed IEEE Smart Grid Standards*.
- [9] P. Carvalho, L. Ferreira, and M. Ilic, “Distributed energy resources integration challenges in low-voltage networks: Voltage control limitations and risk of cascading,” *Sustainable Energy, IEEE Transactions on*, vol. 4, no. 1, pp. 82–88, 2013.
- [10] S. Bolognani, G. Cavraro, R. Carli, and S. Zampieri, “A distributed feedback control strategy for optimal reactive power flow with voltage constraints,” *arXiv preprint arXiv:1303.7173*, 2013.
- [11] M. Farivar, L. Chen, and S. Low, “Equilibrium and dynamics of local voltage control in distribution systems,” 2013.
- [12] B. A. Robbins, C. N. Hadjicostis, and A. D. Domínguez-García, “A two-stage distributed architecture for voltage control in power distribution systems,” *Power Systems, IEEE Transactions on*, vol. 28, no. 2, pp. 1470–1482, 2013.
- [13] B. Zhang, A. D. Domínguez-García, and D. Tse, “A local control approach to voltage regulation in distribution networks,” in *North American Power Symposium (NAPS), 2013*. IEEE, 2013, pp. 1–6.
- [14] A. Kam and J. Simonelli, “Stability of distributed, asynchronous var-based closed-loop voltage control systems,” in *PES General Meeting — Conference Exposition, 2014 IEEE*, July 2014, pp. 1–5.
- [15] M. Savaghebi, A. Jalilian, J. Vasquez, and J. Guerrero, “Secondary control scheme for voltage unbalance compensation in an islanded droop-controlled microgrid,” *Smart Grid, IEEE Transactions on*, vol. 3, no. 2, pp. 797–807, June 2012.
- [16] Q. Shafiee, J. Guerrero, and J. Vasquez, “Distributed secondary control for islanded microgrids - a novel approach,” *Power Electronics, IEEE Transactions on*, vol. 29, no. 2, pp. 1018–1031, Feb 2014.
- [17] J. Rocabert, A. Luna, F. Blaabjerg, and P. Rodriguez, “Control of power converters in ac microgrids,” *Power Electronics, IEEE Transactions on*, vol. 27, no. 11, pp. 4734–4749, Nov 2012.
- [18] T. Vandoorn, J. De Koning, B. Meersman, and L. Vandevelde, “Communication-based secondary control in microgrids with voltage-based droop control,” in *Transmission and Distribution Conference and Exposition (T D), 2012 IEEE PES*, May 2012, pp. 1–6.
- [19] J. F. Hu, J. G. Zhu, and G. Platt, “A droop control strategy of parallel-inverter-based microgrid,” in *Applied Superconductivity and Electromagnetic Devices (ASEMD), 2011 International Conference on*, Dec 2011, pp. 188–191.
- [20] M. Farivar and S. Low, “Branch flow model: Relaxations and convexification,” *arXiv:1204.4865v2*, 2012.
- [21] L. Gan, N. Li, U. Topcu, and S. Low, “Branch flow model for radial networks: convex relaxation,” in *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012.
- [22] M. Farivar, L. Chen, and S. Low, “Equilibrium and dynamics of local voltage control in distribution systems,” in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, 2013, pp. 4329–4334.
- [23] S. H. Low and D. E. Lapsley, “Optimization flow control: basic algorithm and convergence,” *IEEE/ACM Transactions on Networking (TON)*, vol. 7, no. 6, pp. 861–874, 1999.
- [24] W. Rudin, *Principles of mathematical analysis*. McGraw-Hill New York, 1964, vol. 3.
- [25] D. P. Bertsekas, “Nonlinear programming,” 1999.
- [26] M. Farivar, R. Neal, C. Clarke, and S. Low, “Optimal inverter var control in distribution systems with high pv penetration,” in *IEEE Power and Energy Society General Meeting*, San Diego, CA, July 2012.
- [27] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, “Matpower: Steady-state operations, planning, and analysis tools for power systems research and education,” *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 12–19, Feb 2011.

APPENDIX

A. Lemma for Theorem 6

Lemma 8. *Regarding the Hessian matrix of D we have*
(i) $\nabla D^2 = -A$ where $A = \begin{pmatrix} X & -X \\ -X & X \end{pmatrix}$ and (ii)
(ii) $\sigma_{\max}(A) = 2\sigma_{\max}(X)$.

Proof. It’s easy to verify (i). For (ii), let $u = (a, b)^T$ where $a, b \in \mathbb{R}^n$ and $\|u\|_2 = 1$. Observe that $\|Au\|_2^2 = 2\|Xa - Xb\|_2^2 \leq 4(\|Xa\|_2^2 + \|Xb\|_2^2) \leq 4\sigma_{\max}(X)^2(\|a\|_2^2 + \|b\|_2^2) = 4\sigma_{\max}(X)^2$. Therefore, $\sigma_{\max}(A) \leq 2\sigma_{\max}(X)$. Let p be the vector s.t. $\|p\|_2 = 1$ and $\|Xp\|_2 = \sigma_{\max}(X)$. Let $a = -b = p/\sqrt{2}$. Then $\|u\|_2 = 1$ and $\|Au\|_2 = \|\sqrt{2}(Xp, Xp)^T\|_2 = 2\sigma_{\max}(X)$. Hence $\sigma_{\max}(A) = 2\sigma_{\max}(X)$. \square

B. Proof for Theorem 7

First we introduce a few notations. Define $\lambda(t) = [\bar{\lambda}(t)^T, \underline{\lambda}(t)^T]^T$, $\bar{\pi}(t) = \bar{\lambda}(t) - \bar{\lambda}(t-1)$, $\underline{\pi}(t) = \underline{\lambda}(t) - \underline{\lambda}(t-1)$, and $\pi(t) = [\bar{\pi}(t)^T, \underline{\pi}(t)^T]^T$. Define $\mu(t) = [\bar{\mu}(t)^T, \underline{\mu}(t)^T]^T$ where $[\bar{\mu}(t)]_i = v_i(t - \tau_i(t)) - \bar{v}_i$ and $[\underline{\mu}(t)]_i = \underline{v}_i - v_i(t - \tau_i(t))$ for $i = 1, \dots, n$.

Lemma 9. *For each t , we have*

$$\mu(t)^T \pi(t) \geq \frac{1}{\epsilon} \|\pi(t)\|^2.$$

Proof. Denote each component of $\pi(t)$ (and $\lambda(t)$, $\mu(t)$) by subscript l . For component l , if the update rule is (11), apparently we have $[\mu(t)]_l [\pi(t)]_l \geq \frac{1}{\epsilon} [\pi(t)]_l^2$ because both sides of the inequality is 0. If the update rule is (12), we have $[\lambda(t)]_l = [[\lambda(t-1)]_l + \epsilon[\mu(t)]_l]^+$. Apply Projection Theorem and therefore

$$([\lambda(t-1)]_l + \epsilon[\mu(t)]_l - [\lambda(t)]_l)([\lambda(t-1)]_l - [\lambda(t)]_l) \leq 0$$

which yields $[\mu(t)]_l [\pi(t)]_l \geq \frac{1}{\epsilon} [\pi(t)]_l^2$. Summing this inequality over all components l will lead to the lemma. \square

Lemma 10. *The following inequalities hold:*

(i)

$$\|\nabla D(\lambda(t-1)) - \mu(t)\| \leq 2\|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\|$$

(ii)

$$\|v(t) - v(t-1)\| \leq \sigma_{\max}(X) \sum_{t'=t-T_a-T_d}^t (\|\bar{\pi}(t')\| + \|\underline{\pi}(t')\|)$$

Proof. (i) Through simple calculation,

$$\nabla D(\lambda) = \begin{pmatrix} X(\underline{\lambda} - \bar{\lambda}) + v^{par} - \bar{v} \\ -[X(\underline{\lambda} - \bar{\lambda}) + v^{par}] + \underline{v} \end{pmatrix}$$

Therefore, $\nabla D(\lambda(t-1)) - \mu(t)$ is a $2n$ -dimension vector with $[\nabla D(\lambda(t-1)) - \mu(t)]_i = X_i^T(\underline{\lambda}(t-1) - \bar{\lambda}(t-1)) + v_i^{par} - v_i(t - \tau_i(t))$ and $[\nabla D(\lambda(t-1)) - \mu(t)]_{n+i} = -[\nabla D(\lambda(t-1)) - \mu(t)]_i$ for $i = 1, \dots, n$, where X_i denotes the i th row of X . Then,

$$\begin{aligned} & \|[\nabla D(\lambda(t-1)) - \mu(t)]_i\| \\ &= \|X_i^T(\underline{\lambda}(t-1) - \bar{\lambda}(t-1)) - X_i^T q(t-1 - \tau_i(t))\| \\ &\leq \|X_i\| \|\underline{\lambda}(t-1) - \bar{\lambda}(t-1) - q(t-1 - \tau_i(t))\| \end{aligned}$$

$$\begin{aligned}
&\leq \|X_i\| \sum_{t'=t-2T_d-T_a}^{t-1} (\|\underline{\pi}(t')\| + \|\bar{\pi}(t')\|) \\
&\leq \sqrt{2}\|X_i\| \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\| \tag{16}
\end{aligned}$$

The last inequality follows from the fact that for any non-negative real number a and b we have $a + b \leq \sqrt{2(a^2 + b^2)}$. The second last inequality follows from the fact that for the j th component of $\underline{\lambda}(t-1) - \bar{\lambda}(t-1) - q(t-1 - \tau_i(t))$ we have

$$\begin{aligned}
&\underline{\lambda}_j(t-1) - \bar{\lambda}_j(t-1) - q_j(t-1 - \tau_i(t)) \\
&= (\underline{\lambda}_j(t-1) - \underline{\lambda}_j(t-1 - \tau_i(t) - \zeta_j)) \\
&\quad - (\bar{\lambda}_j(t-1) - \bar{\lambda}_j(t-1 - \tau_i(t) - \zeta_j)) \\
&= \sum_{t'=t-\tau_i(t)-\zeta_j}^{t-1} (\underline{\pi}_j(t') - \bar{\pi}_j(t'))
\end{aligned}$$

where ζ_j is a result of the actuation delay and the non-update time steps caused by (11). Define a n -dimensional vector $\hat{\pi}(t')$ whose j th component is $\underline{\pi}_j(t')$ when $t - \tau_i(t) - \zeta_j \leq t' \leq t-1$ and 0 for other t' . Similarly define $\hat{\pi}(t')$ whose j th component is $\bar{\pi}_j(t')$ when $t - \tau_i(t) - \zeta_j \leq t' \leq t-1$ and 0 for other t' . By the definition we have $\|\hat{\pi}(t')\| \leq \|\underline{\pi}(t')\|$ and $\|\hat{\pi}(t')\| \leq \|\bar{\pi}(t')\|$. Observe that $\tau_i(t) + \zeta_j$ is upper bounded by $2T_d + T_a$ by Assumption 2, we have

$$\begin{aligned}
&\|\underline{\lambda}(t-1) - \bar{\lambda}(t-1) - q(t-1 - \tau_i(t))\| \\
&= \left\| \sum_{t'=t-2T_d-T_a}^{t-1} (\hat{\pi}(t') - \hat{\pi}(t')) \right\| \\
&\leq \sum_{t'=t-2T_d-T_a}^{t-1} (\|\hat{\pi}(t')\| + \|\hat{\pi}(t')\|) \\
&\leq \sum_{t'=t-2T_d-T_a}^{t-1} (\|\underline{\pi}(t')\| + \|\bar{\pi}(t')\|)
\end{aligned}$$

which leads to the second last inequality of (16).

Sum (16) over i and $n+i$ where $i = 1, \dots, n$, we have

$$\|\nabla D(\lambda(t-1)) - \mu(t)\| \leq 2\|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\|$$

(ii) By some similar arguments in the derivation of the second last inequality in (16), we have

$$\begin{aligned}
\|v(t) - v(t-1)\| &\leq \sigma_{\max}(X)\|q(t) - q(t-1)\| \\
&\leq \sigma_{\max}(X) \sum_{t'=t-T_a-T_d}^t (\|\bar{\pi}(t')\| + \|\underline{\pi}(t')\|)
\end{aligned}$$

□

Lemma 11. *There exists a positive constant A s.t. when $\epsilon < \frac{1}{A}$, $D(\lambda(t))$ is bounded, and*

$$\|\pi(t)\| \rightarrow 0(t \rightarrow \infty).$$

Proof. Apply Proposition 8, Lemma 9(i) and Lemma 10 to the second order Taylor expansion of $D(\lambda(t))$, we have

$$D(\lambda(t)) = D(\lambda(t-1)) + \nabla D(\lambda(t-1))^T \pi(t)$$

$$\begin{aligned}
&+ \frac{1}{2} \pi^T(t) \nabla^2 D \pi(t) \\
&= D(\lambda(t-1)) + (\nabla D(\lambda(t-1)) - \mu(t))^T \pi(t) \\
&\quad + \mu(t)^T \pi(t) + \frac{1}{2} \pi^T(t) \nabla^2 D \pi(t) \\
&\geq D(\lambda(t-1)) - \|\nabla D(\lambda(t-1)) - \mu(t)\| \|\pi(t)\| \\
&\quad + \left(\frac{1}{\epsilon} - \sigma_{\max}(X)\right) \|\pi(t)\|^2 \\
&\geq D(\lambda(t-1)) - 2\|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\| \|\pi(t)\| \\
&\quad + \left(\frac{1}{\epsilon} - \sigma_{\max}(X)\right) \|\pi(t)\|^2 \\
&\geq D(\lambda(t-1)) + \left(\frac{1}{\epsilon} - \sigma_{\max}(X)\right) \|\pi(t)\|^2 \\
&\quad - \|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} (\|\pi(t)\|^2 + \|\pi(t')\|^2) \tag{17}
\end{aligned}$$

Sum (17) from $t = 1$ to T , we have

$$\begin{aligned}
D(\lambda(T)) &\geq D(\lambda(0)) + \left(\frac{1}{\epsilon} - \sigma_{\max}(X)\right) \sum_{t=1}^T \|\pi(t)\|^2 \\
&\quad - \|X\|_F \sum_{t=1}^T \sum_{t'=t-2T_d-T_a}^{t-1} (\|\pi(t)\|^2 + \|\pi(t')\|^2) \\
&\geq D(\lambda(0)) + \left(\frac{1}{\epsilon} - \sigma_{\max}(X)\right) \sum_{t=1}^T \|\pi(t)\|^2 \\
&\quad - \|X\|_F (2T_d + T_a) \sum_{t=1}^T \|\pi(t)\|^2 \\
&\quad - \|X\|_F \sum_{t=1}^T \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\|^2 \\
&\geq D(\lambda(0)) + \left(\frac{1}{\epsilon} - A\right) \sum_{t=1}^T \|\pi(t)\|^2 \tag{18}
\end{aligned}$$

where $A = \sigma_{\max}(X) + 2\|X\|_F(2T_d + T_a)$. Select ϵ small enough such that

$$\frac{1}{\epsilon} - A > 0$$

Since $D(\lambda(T))$ is upper bounded (assume the primal problem is feasible), by (18), $\sum_{t=0}^T \|\pi(t)\|^2$ is also upper bounded. Because $\sum_{t=0}^T \|\pi(t)\|^2$ is a series consisting of non-negative terms, we must have

$$\|\pi(t)\| \rightarrow 0(t \rightarrow \infty).$$

□

We now prove Theorem 7.

Proof. We first show there must exist one accumulation point of the sequence $\{\lambda(t)\}$. According to (18), $D(\lambda(t))$ is lower bounded by $D(\lambda(0))$, so $\underline{\lambda}(t) - \bar{\lambda}(t)$ is bounded, and $\bar{\lambda}(t)^T (v^{par} - \bar{v}) + \underline{\lambda}(t)^T (\bar{v} - v^{par})$ is lower bounded. Since for each i at least one of $(v_i^{par} - \bar{v}_i)$ and $(\bar{v}_i - v_i^{par})$ is negative, we must have $\underline{\lambda}_i(t)$ and $\bar{\lambda}_i(t)$ are bounded. So the set $\{\lambda \geq 0 | D(\lambda) \geq D(\lambda(0))\}$ is compact, then the existence of accumulation points follows from Weierstrass Theorem [24].

We next show every accumulation point λ^* of $\{\lambda(t)\}$ maximizes the dual problem (15). Let subsequence $\lambda(t_k)$ converges to λ^* with $t_{k+1} - t_k \geq T_a, \forall k$. Define $t_{i,k} = \max\{t \leq t_k - 1 : t + 1 \in T_i\}$. By Assumption 2, $\{t_{i,k}\}_{k=1}^\infty$ satisfies $t_k - T_a \leq t_{i,k} < T_k$ and hence $t_{i,k} \rightarrow \infty$ as $k \rightarrow \infty$. With Lemma 10 (i),

$$\|\nabla D(\lambda(t-1)) - \mu(t)\| \leq 2\|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\| \rightarrow 0$$

Therefore,

$$\lim_k \bar{\mu}_i(t_k + 1) = \lim_k \frac{\partial D(\lambda(t_k))}{\partial \bar{\lambda}_i} = \frac{\partial D(\lambda^*)}{\partial \bar{\lambda}_i}$$

By Lemma 9 (ii), $\|v(t) - v(t-1)\| \rightarrow 0$, hence

$$\|\bar{\mu}_i(t_k + 1) - \bar{\mu}_i(t_{i,k})\| \leq \sum_{t'=t_k-T_a-T_a}^{t_k+1} |v_i(t') - v_i(t'-1)| \rightarrow 0$$

Therefore $\lim_k \bar{\mu}_i(t_{i,k}) = \lim_k \bar{\mu}_i(t_k + 1) = \partial D(\lambda^*)/\partial \bar{\lambda}_i$. Also since

$$|\bar{\lambda}_i(t_k) - \bar{\lambda}_i(t_{i,k})| \leq \sum_{t'=t_k-T_a}^{t_k} |\bar{\pi}_i(t')| \rightarrow 0$$

we have $\lim_k \bar{\lambda}_i(t_k) = \lim_k \bar{\lambda}_i(t_{i,k})$. So,

$$\begin{aligned} [\bar{\lambda}_i^* + \epsilon \frac{\partial \nabla D(\lambda^*)}{\partial \bar{\lambda}_i}]^+ - \bar{\lambda}_i^* &= \lim_k \left[[\bar{\lambda}_i(t_k) + \epsilon \bar{\mu}_i(t_k + 1)]^+ - \bar{\lambda}_i(t_k) \right] \\ &= \lim_k \left[[\bar{\lambda}_i(t_{i,k}) + \epsilon \bar{\mu}_i(t_{i,k})]^+ - \bar{\lambda}_i(t_{i,k}) \right] \\ &= \lim_k \bar{\pi}_i(t_{i,k} + 1) = 0 \end{aligned}$$

Similarly, $[\underline{\lambda}_i^* + \epsilon \frac{\partial \nabla D(\lambda^*)}{\partial \underline{\lambda}_i}]^+ - \underline{\lambda}_i^* = 0$. Apply this equality for all i and we can get $[\lambda^* + \epsilon \nabla D(\lambda^*)]^+ = \lambda^*$. Apply Projection Theorem [25], we have

$$-\epsilon \nabla D(\lambda^*)^T (\lambda - \lambda^*) \geq 0, \quad \forall \lambda \geq 0$$

Which establishes the optimality of λ^* in the dual problem, and thus also the optimality of $\underline{\lambda}^* - \bar{\lambda}^*$ in the primal problem.

Next we will prove the uniqueness of the maximizers of the dual problem (15). Let λ' be a maximizer of $D(\lambda)$, then $\underline{\lambda}' - \bar{\lambda}'$ is the minimizer of the primal problem. Since X is positive definite, the primal problem has a unique minimizer. Hence $\underline{\lambda}' - \bar{\lambda}'$ is unique. By complementary slackness in KKT condition [25], $\forall i, \bar{\lambda}'_i ([X(\underline{\lambda}' - \bar{\lambda}')]_i + v_i^{par} - \bar{v}_i) = 0$ and $\underline{\lambda}'_i (\underline{v}_i - [X(\underline{\lambda}' - \bar{\lambda}')]_i - v_i^{par}) = 0$. Because $([X(\underline{\lambda}' - \bar{\lambda}')]_i + v_i^{par} - \bar{v}_i)$ and $(\underline{v}_i - [X(\underline{\lambda}' - \bar{\lambda}')]_i - v_i^{par})$ cannot be both zero, at least one of $\bar{\lambda}'_i$ and $\underline{\lambda}'_i$ has to be zero. Combining this with the fact that $\underline{\lambda}' - \bar{\lambda}'$ is unique, we have λ' is unique, i.e. the dual problem (15) has a unique maximizer.

At last, since any accumulation point of $\{\lambda(t)\}$ is the maximizer of (15), the accumulation point of $\{\lambda(t)\}$ is unique. According to [24] this means $\lambda(t)$ will converge to the maximizer of the dual problem (15). Also, by (11), (12) and Assumption 2, $q(t)$ will converge to the minimizer of the primal problem (14) and $v(t)$ will converge to a point v^* in $[\underline{v}, \bar{v}]$.

□