Real-time Robust Decentralized 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 electric vehicles 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 will enable realtime distributed voltage control that is needed for the future power grid.

Corresponding Author: N. Li

G. Qu and N. Li are with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA (Emails: nali@seas.harvard.edu, gqu@g.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.)

One key element to implement those new technologies is the voltage control rules 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, to calculate, or to communicate. The local information dependence facilitates the realtime implementation of those algorithms. In fact, there exist classes of inverter-based local voltage control schemes that only use local voltage measurements. For example, the IEEE 1547.8 standard [5] proposes decentralized voltage control rules based on the local voltage deviations from its nominal values. 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.

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 [6]. 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 voltage control with different information structure. In particular, we first show that using only voltage measurements to determine reactive power compensation is insufficient, because there always exist operating conditions under which such voltage control fails to maintain acceptable voltages. Then we proposes two decentralized algorithms by adding additional information into the control design. As shown in 1, 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 voltage in the acceptable range; and the one with higher complexity can further minimize the cost of the reactive power compensation in a particular form. Our results implies that with the aid of right local information, local voltages carry out the whole network information for Volt/Var control. In addition, the two algorithms are robust against

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

heterogeneous update rates and delays. This means that on one hand, the algorithms does not require the buses update at each time step. They might update once in a few time steps. And the control devices underlying the updates can have different clock rates. On the other hand, the algorithms consider delays resulting from measurement devices (e.g. voltage sensors) and implementation devices (e.g. inverters). Our results show that despite the heterogeneous update rates and the delays, our algorithms will still converge. We will also discuss how the heterogeneous update rates and the delays will affect the selection of stepsize that guarantees convergence.

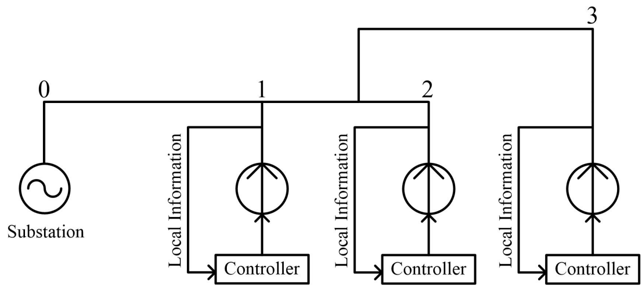

Figure 1: An example of a distribution network with 3 buses and 1 substation. Each bus is equipped with a controller that only takes local information as its input.

Research has been proposed and conducted to improve the existing voltage control to mitigate the voltage fluctuation impact. To name a few, [7] studies the active power curtailment to mitigate the voltage rise impact caused by DER; [8] studies the distributed VAR control to minimize power losses and stablizie voltages; [9] reverse-engineers the IEEE 1547.8 standard and study the equilibrium and dynamics of voltage control; [10] 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. A recent paper [11] proposes a similar algorithm as the first algorithm in this paper. The difference lies in the following: [11] uses a nonlinear power flow model while this paper uses a linear power flow model; [11] drives the voltage profile to a reference point while this paper stabilizes the voltage profile into the acceptable range; [11] characterizes convergence guarantee under each specific active power profile P, while we characterize convergence guarantee regardless of P; [11] studies a specific case of the impact of network topology on convergence while this paper does not. Another paper [12] is also similar to this paper. The difference is two-fold. 1) [12] uses a linear control algorithm, i.e. the reactive power is a linear function of the measured voltage. This algorithm falls under the class of algorithms described in Section III and we will further discuss their link in Remark 1. 2) [12] uses a stronger assumption in the case that the buses have heterogeneous update rates (referred to as the asynchronous algorithm in [12]). The stronger assumption leads to an interesting result that the asynchronous algorithm enjoys higher degree of robustness. We will link this result to our algorithms in Remark 3 and Remark 4.

In addition to the work mentioned above, in power engineering community, much work has been done in voltage control for microgrid (which can be viewed as a special distribution system). See [13] for a comprehensive review. The existing methods generally fall under two layers, primary control (droop control) and secondary control. The droop control can be viewed as a linear controller in which the reactive power output is a linear function of the measured voltage. In section III, we will show that this kind of methods do not work under certain circumstances. The secondary control is essentially an integral controller that eliminates the deviation of the measured voltage from a reference point. These methods can be viewed as a special case of the first algorithm in this paper by shrinking the acceptable range to the reference point. But no theoretic analysis that incorporate network constraints has been performed to guarantee convergence of such integral controllers.

The remaining of this paper is organized as follows: Section II present an AC power flow model, its linear approximation, and the formulation of the Volt/Var control; Section II-C illustrate 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 simulate the algorithms to complement our analysis; Sectoin 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., [14], [15] 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:

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

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

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

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

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 [14] 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. $r, x << 1$, while $v \sim 1$. Thus real and reactive power losses are typically much smaller than power flows P_{ij}, Q_{ij}. Following [6], 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 [6].

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

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

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

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 [16]. 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.*[2]

Proof. We refer readers to [16] for the detailed proof. ☐

[2]A matrix is a positive matrix iff each item is positive.

C. Problem formulation

Before rigorously formulating the Volt/Var control problem, we separate q into two part, $q^c = (q_1^c,\dots,q_n^c)$ 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.[3] 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 $[\underline{v}, \bar{v}]$ 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. 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 lead the system voltage $v(t)$ to reach the acceptable range $[\underline{v}, \bar{v}]$ under any system operating condition which is given by v^{par}. Mathematically, it requires that

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

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 will address in this paper is the information requirement of the controller u in order to stabilize the voltage in the acceptable range. 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 demand decentralized algorithms 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

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

of control has been proposed and discussed in many existing literature and applications. For example, the IEEE 1547.8 standard [5] 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. The scheme is shown in Figure 2. Besides the IEEE 1547.8 standard, there are other researches promoting adapting reactive power injection according to the voltage [7]. However, in the following, we show that this type of controller is insufficient for Volt/Var control.

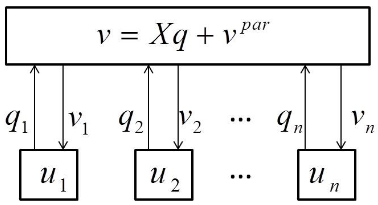

Figure 2: IEEE 1547.8 standard: Decentralized Volt/Var control using local voltage measurements.

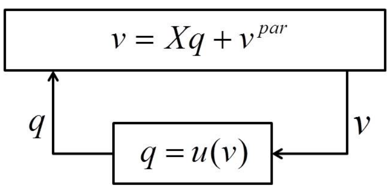

Figure 3: Volt/Var control using merely voltage information: This type of controller is insufficient for Volt/Var control.

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

$$q(t) = u(v(t)), \tag{5}$$

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 $[\underline{v}, \bar{v}]$.*[4]

Proof. The proof is straghtforward. Subsistuting (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., $\lim_{t\to\infty} \text{dist}(v(t), [\underline{v}, \bar{v}]) = 0$, then v^{par} should be at least in the set of $M := \{v - Xu(\tilde{v}) : v, \tilde{v} \in [\underline{v}, \bar{v}]\}$ 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. □

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

This proposition tells us that any Volt/Var control depending 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 the paper, we will show that if we use both the information of the current q and v, then a fully decentralized in the form of $u_i(p_i(t), v_i(t))$ is able to maintain acceptable voltages; further, if we introduce some auxiliary variables, then a fully decentralized algorithm in the form of $u_i(\text{virtual variables at } i, v_i)$ is able to both maintain acceptable voltages and minimize a cost of reactive power compensation in a particular form.

Remark 1. *The algorithm in [12], in which the reactive injection is a linear function of the measured voltage, is a special case of (5). By Proposition 2, there are cases in which the algorithm in [12] will not stabilize the voltages in the acceptable range. This is consistent with the results in [12]: the equilibrium point of the algorithm in [12] is dependent on the initial voltage profile, and can fall outside of the acceptable range if the initial voltage profile is too far away from the reference voltage.*

IV. A DECENTRALIZED ALGORITHM TO REACH THE ACCEPTABLE VOLAGE 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 in the acceptable range. The scheme of the algorithm is shown in Figure 4.

The algorithms introduced in this section as well as in section V will incorporate robustness against heterogeneous update rates and delays. First, the algorithm considers that the buses do not necessarily update their reactive power injections at each time step. Instead, they might update only once in a few time steps, and their pace of updating can be different. Second, the algorithm also considers measurement and implementation delays. In details, we assume there might be a time delay between the measuring of voltage and the receiving of the measurement by the device, as well as a time delay between the calculation of a target reactive injection and the achievement of the target injection. Equiping our algorithms with the robustness will make our algorithms more realistic for two reasons: firstly, different buses usually have heterogeneous control devices that operate on different clocks; secondly, delays widely exist in real systems. We now formulate the first algorithm which we will name as algorithm I and then prove its convergence.

A. Algorithm

For each bus i, we introduce a infinite time step set $T_i \subseteq \{1, 2, \ldots\}$ to represent the time steps bus i updates. For $t \notin T_i$, no update will be made, which means the reactive injection stays the same as the previous time step, i.e.

$$q_i(t+1) = q_i(t) \tag{6}$$

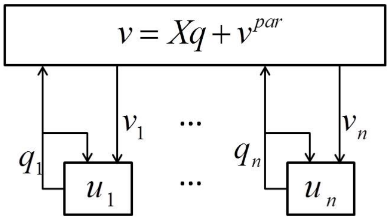

Figure 4: Volt/Var control using local information of local voltage and reactive power injection: This type of controller is able to maintain acceptable voltages.

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

$$q_i(t+1) = q_i(t) - \epsilon d_i(v_i(t+1-\tau_i(t))) \tag{7}$$

where for a voltage value v,

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

and ϵ is a positive constant stepsize. Here $[]^+$ is defined as $[a]^+ := \max(a, 0)$.

The $\tau_i(t)$ in (7) is the measurement delay of bus i at t. $v_i(t+1-\tau_i(t))$ 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. When 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.

The algorithm in (7) says that if the locally measured voltage at bus i is above its upper limit, bus i decreases the reactive power injection; in contrast, if the local voltage is 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 that 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 the algorithm (6) and (7) which make it attractive to real-time and scalable implementation.

i) We note that in the algorithm, each bus i uses only the local voltage measurement $v_i(t+1-\tau_i(t))$ and its previous reactive power injection $q_i(t)$, and the control is similar to an integral controller. Thus, the algorithm does not require any communication and the implementation is simple.
ii) The algorithm does not require any system operating information about v^{par}. This makes the algorithm practical because due to the large volatility and uncertainty of renewable energy, time-varying nature of uncertainty, and the privacy concern of consumers, v^{par} is not available and has huge uncertainty.
iii) Though the convergence of the algorithm depends on the step size value ϵ as shown in the next section, how to choose 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. *$\forall i$, the difference between consecutive update time steps in T_i is upper bounded by T_a. $\forall i$ and $\forall t$, $\tau_i(t)$ are 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 uppder bounded by T_d. This assumption is mild and reasonable, and can be easily guaranteed in real operation. With the assumption, we have the following convergence result.

Theorem 3. *For algorithm I (6) and (7), if Assumption 1 holds and $\epsilon < \frac{2}{\sigma_{max}(X)+2\|X\|_F T_d}$, $d(v(t))$ will converge to 0.*

Remark 2. *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 convergence. The slope of the linear term, $\|X\|_F$, reflects the intuition that the larger the network is (which results in a larger $\|X\|_F$), the bigger effect the delays will have on the convergence.*

Remark 3. *The asynchronous setting of [12] uses a slightly stronger setting than Assumption 1: [12] assumes at each time step, at most M buses update. With this assumption, [12] shows that the stability region (in terms of the maximum stepsize that guarantees convergence) of the asynchronous algorithm is greater than the synchronous algorithm (synchronous here means all buses update at all time steps with undelayed information). If we remove the stronger assumption, the result in [12] will become that the asynchronous algorithm has the same stability region as the synchronous algorithm. This is consistent with Theorem 3. This can be seen, by noticing that the maixmum stepsize given in Theorem 3 is independent of T_a, and if we set $T_a = T_d = 0$, algorithm I becomes a synchronous algorithm.*

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 [17], we will show that $\|q(t+1) - q(t)\| \to 0$. Combine this with the fact that the buses have to update at least once within T_a steps, we will reach the conclusion that $d(v(t))$ will asymptotically converge to 0.

We first define some notations. Let $\alpha(t) = q(t+1) - q(t)$ and $d'(v) = v - \frac{\bar{v}+\underline{v}}{2}$, i.e. $d'_i(v_i) = v_i - \frac{\bar{v}_i+\underline{v}_i}{2}$. Define Lyapunov function $\psi(q) = \frac{1}{2} d'(v(q))^T X^{-1} d'(v(q))$ where $v(q) = Xq + v^{par}$. Clearly, $\nabla\psi(q) = d'(v(q))$ and $\nabla^2\psi(q) = X$. Define $\gamma(t)$ as $\gamma_i(t) = d'_i(v_i(t+1-\tau_i(t)))$.

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

Proof. For each i,

$$\begin{aligned}|\gamma_i(t) - \nabla_i\psi(q(t))| &= |d_i'(v_i(t+1-\tau_i(t))) - d_i'(v_i(t+1))| \\ &= |v_i(t+1-\tau_i(t)) - v_i(t+1)| \\ &= |X_i^T(q(t) - q(t-\tau_i(t)))| \\ &\le \|X_i\|\|q(t) - q(t-\tau_i(t))\| \\ &= \|X_i\|\|\sum_{t'=t-\tau_i(t)}^{t-1}\alpha(t')\| \\ &\le \|X_i\|\sum_{t'=t-\tau_i(t)}^{t-1}\|\alpha(t')\| \\ &\le \|X_i\|\sum_{t'=t-T_d}^{t-1}\|\alpha(t')\|\end{aligned}$$

Where X_i is the i the row of X. Sum over i and the lemma follows. □

Lemma 5. $\gamma(t)^T\alpha(t) \le (-\frac{1}{\epsilon})\|\alpha(t)\|^2$

Proof. For each i, $\gamma_i(t) = d_i'(v_i(t+1-\tau_i(t)))$. For notational simplicity, we denote $t+1-\tau_i(t)$ by t'. If bus i updates, then $\alpha_i(t) = -\epsilon d_i(v_i(t'))$. If $d_i(v_i(t')) > 0$, then $d_i'(v_i(t')) = d_i(v_i(t')) + \frac{\bar{v}_i - \underline{v}_i}{2}$ and $d_i'(v_i(t'))d_i(v_i(t')) \ge d_i(v_i(t'))^2$, i.e. $\gamma_i(t)\alpha_i(t) \le (-\frac{1}{\epsilon})\alpha_i(t)^2$. If $d_i(v_i(t')) < 0$, then $d_i'(v_i(t')) = d_i(v_i(t')) - \frac{\bar{v}_i - \underline{v}_i}{2}$ and hence $d_i'(v_i(t'))d_i(v_i(t')) \ge d(v_i(t'))^2$, i.e. $\gamma_i(t)\alpha_i(t) \le (-\frac{1}{\epsilon})\alpha_i(t)^2$. If $d_i(v_i(t')) = 0$ or bus i does not update, we also have, $\gamma_i(t)\alpha_i(t) \le (-\frac{1}{\epsilon})\alpha_i(t)^2$ because both sides of the inequality are zero. Sum the inequality over i and we're done. □

Now we are ready to prove Theorem 3.

Proof. With Lemma 4 and Lemma 5 we have

$$\begin{aligned}&\psi(q(t+1)) - \psi(q(t)) \\ &\le \nabla\psi(q(t))^T\alpha(t) + \frac{1}{2}\alpha(t)^T\nabla^2\psi(q(t))\alpha(t) \\ &\le \|\nabla\psi(q(t)) - \gamma(t)\|\|\alpha(t)\| + \gamma(t)^T\alpha(t) + \frac{1}{2}\alpha(t)^TX\alpha(t) \\ &\le \|X\|_F\sum_{t'=t-T_d}^{t-1}\|\alpha(t')\|\|\alpha(t)\| + (-\frac{1}{\epsilon} + \frac{1}{2}\sigma_{max}(X))\|\alpha(t)\|^2 \\ &\le \frac{1}{2}\|X\|_F\sum_{t'=t-T_d}^{t-1}(\|\alpha(t')\|^2 + \|\alpha(t)\|^2) \\ &+ (-\frac{1}{\epsilon} + \frac{1}{2}\sigma_{max}(X))\|\alpha(t)\|^2\end{aligned}$$

Sum from $t=0$ to $t=T$, we have

$$\begin{aligned}&\psi(q(T+1)) - \psi(q(0)) \\ &\le \frac{1}{2}\|X\|_F\sum_{t=0}^{T}\sum_{t'=t-T_d}^{t-1}(\|\alpha(t')\|^2 + \|\alpha(t)\|^2) \\ &+ (-\frac{1}{\epsilon} + \frac{1}{2}\sigma_{max}(X))\sum_{t=0}^{T}\|\alpha(t)\|^2 \\ &= \frac{1}{2}\|X\|_F\sum_{t=0}^{T}\sum_{t'=t-T_d}^{t-1}\|\alpha(t')\|^2 \\ &+ (-\frac{1}{\epsilon} + \frac{1}{2}\sigma_{max}(X) + \frac{T_d}{2}\|X\|_F)\sum_{t=0}^{T}\|\alpha(t)\|^2 \\ &\le (-\frac{1}{\epsilon} + \frac{1}{2}\sigma_{max}(X) + T_d\|X\|_F)\sum_{t=0}^{T}\|\alpha(t)\|^2\end{aligned}$$

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

$$\begin{aligned}|d_i(v_i(t)) - d_i(v_i(t_i'(n_t)))| &\le \|v_i(t) - v_i(t_i'(n_t))\| \\ &\le \|X_i\|\|q(t-1) - q(t_i'(n_t) - 1)\| \\ &\le \|X_i\|\sum_{t'=t-T_a-T_d-1}^{t-2}\|\alpha(t')\|\end{aligned} \tag{9}$$

Where the first inequality follows from the fact that $d_i(v)$ is a Lipshitz function with Lipshitz coefficient 1. Let $t \to \infty$, then $n_t \to \infty$, so $d_i(v_i(t_i'(n_t))) \to 0$. Also notice the right hand side of (9) converges to 0, we have $d_i(v_i(t)) \to 0$. In conclusion, $\|d(v(t))\| \to 0$. □

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 will name as Algorithm II. Algorithm II will not only drive the voltage into a point in the feasible set $[\underline{v}, \bar{v}]$, 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 I, we use a time step set T_i to represent the time steps that bus i updates. 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:

$$\bar{\lambda}_i(t+1) = \bar{\lambda}_i(t) \tag{10a}$$
$$\underline{\lambda}_i(t+1) = \underline{\lambda}_i(t) \tag{10b}$$
$$q_i(t+1) = q_i(t) \tag{10c}$$

If $t \in T_i$,

$$\bar{\lambda}_i(t+1) = \left[\bar{\lambda}_i(t) + \epsilon(v_i(t-\tau_i(t)) - \bar{v}_i)\right]^+ \quad (11a)$$
$$\underline{\lambda}_i(t+1) = \left[\underline{\lambda}_i(t) + \epsilon\left(\underline{v}_i - v_i(t-\tau_i(t))\right)\right]^+ \quad (11b)$$
$$q_i(t+1) = \underline{\lambda}_i(t+1-\tau_i'(t)) - \bar{\lambda}_i(t+1-\tau_i'(t)) \quad (11c)$$

where $\tau_i(t)$ captures the measurement delay, the same as its counterpart in algorithm I. In addition to measurement delay, we introduce $\tau_i'(t)$ to capture the implementation delay, which means that the reactive power injected at time step t is the reactive power calculated $\tau_i'(t)$ steps back from t, i.e. there is a delay in the implementation of the reactive power injection. The same as algorithm I, for $t \notin T_i$ we abuse the notation $\tau_i(t)$ and $\tau_i'(t)$ to be any nonnegative integers. The diagram of the algorithm is shown in Figure 5.

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. Moreover, the updating rule is also similar to an integral controller with saturation. As a result, this algorithm is fully decentralized and it shares the same properties with algorithm I, making it also attractive to practical implementation.

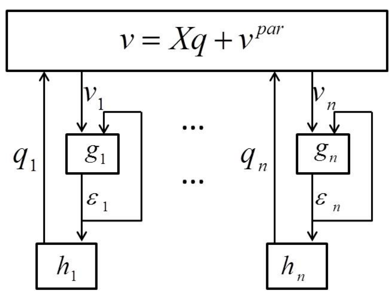

Figure 5: Volt/Var control using local information of voltage, reactive power injection, and additional auxiliary variables: This type of controller is able to guarantee both of feasibility and optimality.

Before we proceed to the convergence analysis of algorithm II, we point out its relation to an optimization problem. To simplify illustration, we consider the following synchronous version (assuming all buses update at all time steps and assuming no delays) of algorithm II:

$$\bar{\lambda}_i(t+1) = \left[\bar{\lambda}_i(t) + \epsilon(v_i(t) - \bar{v}_i)\right]^+ \quad (12a)$$
$$\underline{\lambda}_i(t+1) = \left[\underline{\lambda}_i(t) + \epsilon\left(\underline{v}_i - v_i(t)\right)\right]^+ \quad (12b)$$
$$q_i(t+1) = \underline{\lambda}_i(t+1) - \bar{\lambda}_i(t+1) \quad (12c)$$

Theorem 6. *If $\epsilon < \frac{1}{\sigma_{\max}(X)}$, $q(t)$ in algorithm (12) converges to the optimal point q^* of the following optimization problem:*

$$\min_q \quad \frac{1}{2}q^T X q \quad (13a)$$
$$s.t. \quad Xq + v^{par} \le \bar{v}, \quad (13b)$$
$$Xq + v^{par} \ge \underline{v}, \quad (13c)$$

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

Proof. Introducing dual variable $\bar{\lambda}$ for (13b) and $\underline{\lambda}$ for (13c), we have the following dual gradient method [18]:

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

The preceding algorithm is equivalent to the following:

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

which is exactly the algorithm (12).

Through simple derivation, the dual problem of (13) is given by:

$$\max_{\bar{\lambda},\underline{\lambda}\ge 0} D = -\frac{1}{2}(\underline{\lambda} - \bar{\lambda})X(\underline{\lambda} - \bar{\lambda}) + \bar{\lambda}(v^{par} - \bar{v}) + \underline{\lambda}(\underline{v} - v^{par}) \quad (14)$$

Therefore, we know that if $\epsilon < \frac{2}{|\sigma_{\min}(\nabla D^2)|}$, $\bar{\lambda}(t), \underline{\lambda}(t)$ converge to the dual optimum. Correspondingly, $q(t)$ converges to the optimal point of (13). The conclusion of the theorem follows if $\sigma_{\min}(\nabla D^2) = -2\sigma_{\max}(X)$, which will be proved in the next lemma. □

Lemma 7. *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 $\|Au\|_2^2 = 2\|Xa - Xb\|_2^2 \le 4(\|Xa\|_2^2 + \|Xb\|_2^2) \le 4\sigma_{\max}(X)^2(\|a\|_2^2 + \|b\|_2^2) = 4\sigma_{\max}(X)^2$. Therefore, $\sigma_{\max}(A) \le 2\sigma_{\max}(X)$. Let p be the vector s.t. $\|p\|_2 = 1$ and $\|Xp\|_2 = \sigma_{\max}(X)$. Let $a = -b = p/\sqrt{2}$. Then $\|Au\|_2 = \|\sqrt{2}(Xp, Xp)^T\|_2 = 2\sigma_{\max}(X)$. Hence $\sigma_{\max}(A) = 2\sigma_{\max}(X)$. □

As in algorithm I, we require the following mild assumption to guarantee convergence:

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_d.*

Theorem 8. *In algorithm II (10) and (11), 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 (13) and the dual problem (14) respectively.*

Remark 4. *Different from Theorem 3 and [12], T_a contributes a linear term in the denominator of the maximum stepsize in Theorem 8. This means that heterogeneous update rates (asynchronicity) will shrink the stability region of algorithm II. This can be seen as a disadvantage of algorithm II.*

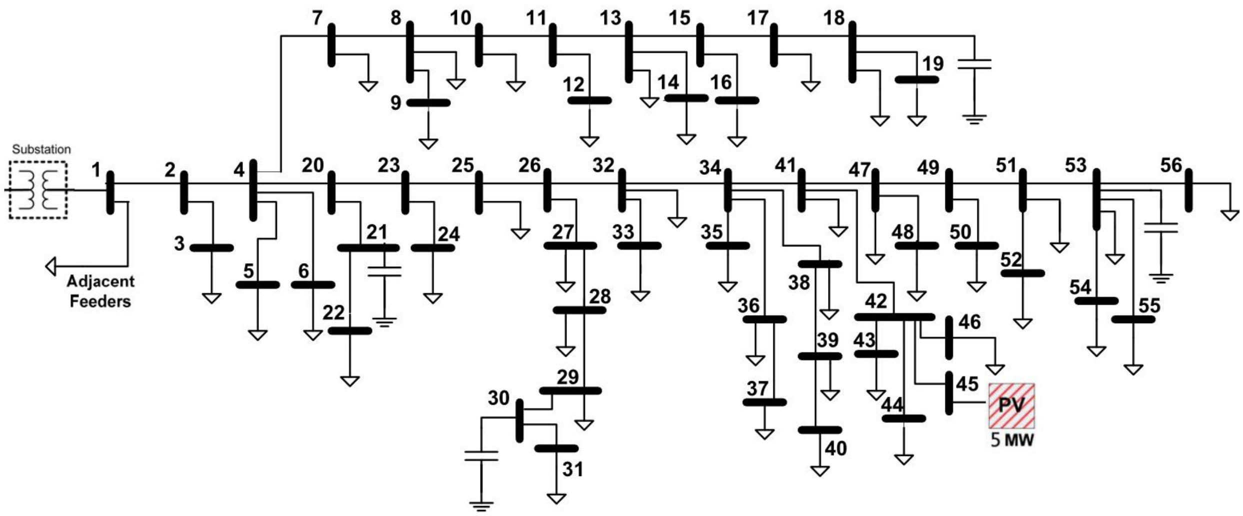

Figure 6: Schematic diagram of two SCE distribution systems.

VI. CASE STUDY

In this section we evaluate the two robust decentralized algorithms on a distribution circuit of South California Edison with a high penetration of photovoltaic (PV) generation [19]. Figure 6 shows a 56-bus distribution circuit. Note that Bus 1 indicates the substation, and there are 1 photovoltaic (PV) generators located on buses 45 and there are shunt capacitors located at bus 19, 21, 30, 53. See [19] for the network data including line impedance, peak MVA demand of 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 $[11.4\text{kV}, 12.6\text{kV}]$ 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 simulate the voltage control algorithms using the full nonlinear AC power flow model (1) [20].

We run the 2 algorithms under a moderate load, large PV generation setting. This setting suggests that part of the initial voltages are higher than the voltage upper bound. The simulation parameters and results are summarized in Figure 7 and Figure 8. From the 2 figures, it can be shown that the buses update at different clock rates. Besides, due to the delays, they even make incorrect updates (e.g., in Figure 8, at $t = 20$, all the voltages are already within the acceptable range, but the buses somehow drive the voltages outside of the range). However, despite the heterogeneous update clock rates and delays, the algorithm still converges asymptotically, which verifies the proposed algorithm's effectiveness.

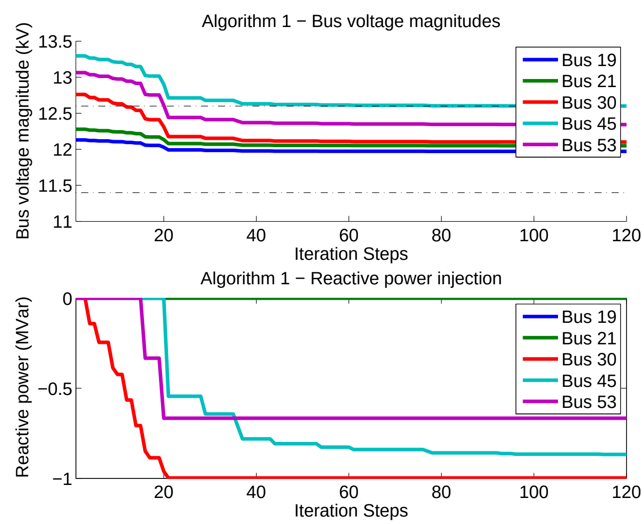

Figure 7: Simulation result of algorithm I. The upper figure shows the voltage profile, and the lower figure shows reactive power injection. Simulation parameter: $T_a = 25$, $T_d = 15$, $\epsilon = 8$.

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

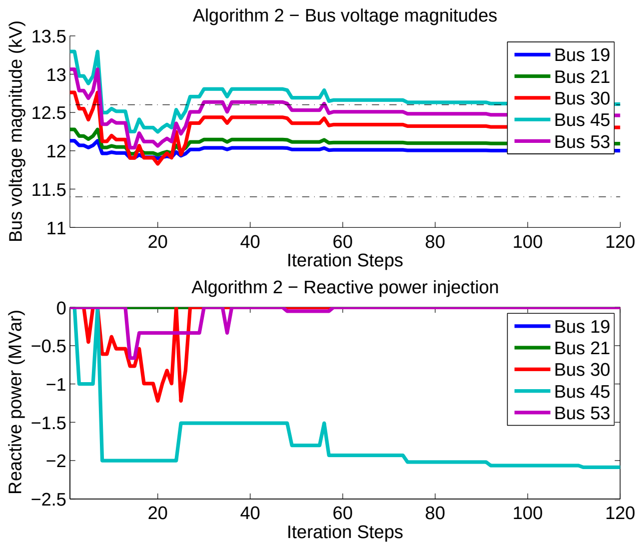

Figure 8: Simulation result of algorithm I. The upper figure shows the voltage profile, and the lower figure shows reactive power injection. Simulation parameter: $T_a = 25$, $T_d = 15$, $\epsilon = 5$.

Two problems remain unsolved. First, in this paper we propose two algorithms, but the relationship between the two algorithms remain understudied. Second, in this paper we assume the reactive injections are unbounded, while in real systems, they are bounded due to the capacity of control devices. If we incorporate the bounds into the algorithms in this paper, the algorithms will fail to converge under some circumstances. Actually, it can be shown that no decentralized algorithms can work if the reactive injection are bounded. This suggests that inter bus communication is necessary for voltage control with reactive power bounds. However, it is still an ongoing research on how to design control rules with inter-bus communication that guarantee convergence with reactive power bounds.

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] "P1547.8–recommended practice for establishing methods and procedures that provide supplemental support for implementation strategies for expanded use of ieee standard 1547," *Approved & Proposed IEEE Smart Grid Standards*.

[6] 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.

[7] 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.

[8] 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.

[9] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," 2013.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] M. Farivar and S. Low, "Branch flow model: Relaxations and convexification," *arXiv:1204.4865v2*, 2012.

[15] 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.

[16] 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.

[17] S. H. Low and D. E. Lapsley, "Optimization flow controli: basic algorithm and convergence," *IEEE/ACM Transactions on Networking (TON)*, vol. 7, no. 6, pp. 861–874, 1999.

[18] D. P. Bertsekas, "Nonlinear programming," 1999.

[19] 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.

[20] 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.

[21] W. Rudin, *Principles of mathematical analysis*. McGraw-Hill New York, 1964, vol. 3.

Appendix

Define $\lambda(t) = [\bar{\lambda}(t)^T, \underline{\lambda}(t)^T]^T$, $\bar{\pi}(t) = \bar{\lambda}(t+1) - \bar{\lambda}(t)$, $\underline{\pi}(t) = \underline{\lambda}(t+1) - \underline{\lambda}(t)$, 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 $[\bar{\mu}(t)]_i = v_i(t - t_i) - \bar{v}_i$ and $[\underline{\mu}(t)]_i = \underline{v}_i - v_i(t - t_i)$ for $i = 1, ..., N$.

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 (10), apparently we have $[\mu(t)]_l[\pi(t)]_l \geq \frac{1}{\epsilon}[\pi(t)]_l^2$. If the update rule is (11), we have $[\lambda(t+1)]_l = [[\lambda(t)]_l + \epsilon[\mu(t)]_l]^+$. Apply Projection Theorem and therefore

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

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. □

Lemma 10. *The following inequalities hold:*

(i)

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

(ii)

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

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

$$\begin{aligned} |[\nabla D(\lambda(t)) - \mu(t)]_i| &= |X_i^T(\underline{\lambda}(t)-\bar{\lambda}(t)) - X_i^T q(t-\tau_i(t))| \\ &\le \|X_i\|\|\underline{\lambda}(t)-\bar{\lambda}(t)-q(t-\tau_i(t))\| \\ &\le \|X_i\| \sum_{t'=t-2T_d-T_a}^{t-1} (\|\underline{\pi}(t')\| + \|\bar{\pi}(t')\|) \\ &\le \sqrt{2}\|X_i\| \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\| \end{aligned} \tag{15}$$

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

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

where ζ_j is a result of the implementation delay and the non-update time steps caused by (10). Define n-dimensional vector $\hat{\underline{\pi}}(t')$ whose jth component is $\underline{\pi}_j(t')$ when $t-\tau_i(t)-\zeta_j \le t' \le t-1$ and 0 for other t'. Similarly define $\hat{\bar{\pi}}(t')$ whose jth component is $\bar{\pi}_j(t')$ when $t-\tau_i(t)-\zeta_j \le t' \le t-1$ and 0 for other t'. By the definition we have $\|\hat{\underline{\pi}}(t')\| \le \|\underline{\pi}(t')\|$ and $\|\hat{\bar{\pi}}(t')\| \le \|\bar{\pi}(t')\|$. Observe that $\tau_i(t)+\zeta_j$ is upper bounded by $2T_d+T_a$ by Assumption 1 and Assumption 2, we have

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

which leads to the second last inequality of (15).

Sum (15) over i and $n+i$ where $i=1,...,n$, we have

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

(ii) By some similar arguments in the derivation of the second last inequality in (15), we have

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

□

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

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

Proof. Apply Proposition 7, Lemma 9(i) and Lemma 10 to the second order Taylor expansion of $D(\lambda(t+1))$, we have

$$\begin{aligned} D(\lambda(t+1)) &= D(\lambda(t)) + \nabla D(\lambda(t))^T \pi(t) \\ &\quad + \frac{1}{2}\pi^T(t)\nabla^2 D\pi(t) \\ &= D(\lambda(t)) + (\nabla D(\lambda(t)) - \mu(t))^T \pi(t) \\ &\quad + \mu(t)^T\pi(t) + \frac{1}{2}\pi^T(t)\nabla^2 D\pi(t) \\ &\ge D(\lambda(t)) - \|\nabla D(\lambda(t)) - \mu(t)\|\|\pi(t)\| \\ &\quad + (\frac{1}{\epsilon} - \sigma_{max}(X))\|\pi(t)\|^2 \\ &\ge D(\lambda(t)) - 2\|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\|\|\pi(t)\| \\ &\quad + (\frac{1}{\epsilon} - \sigma_{max}(X))\|\pi(t)\|^2 \\ &\ge D(\lambda(t)) + (\frac{1}{\epsilon} - \sigma_{max}(X))\|\pi(t)\|^2 \\ &\quad - \|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} (\|\pi(t)\|^2 + \|\pi(t')\|^2) \end{aligned} \tag{16}$$

Sum (16) over t, we have

$$\begin{aligned} D(\lambda(T+1)) &\ge D(\lambda(0)) + (\frac{1}{\epsilon} - \sigma_{max}(X)) \sum_{t=0}^{T} \|\pi(t)\|^2 \\ &\quad - \|X\|_F \sum_{t=0}^{T} \sum_{t'=t-2T_d-T_a}^{t-1} (\|\pi(t)\|^2 + \|\pi(t')\|^2) \\ &\ge D(\lambda(0)) + (\frac{1}{\epsilon} - \sigma_{max}(X)) \sum_{t=0}^{T} \|\pi(t)\|^2 \\ &\quad - \|X\|_F (2T_d+T_a) \sum_{t=0}^{T} \|\pi(t)\|^2 \\ &\quad - \|X\|_F \sum_{t=0}^{T} \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\|^2 \\ &\ge D(\lambda(0)) + (\frac{1}{\epsilon} - A) \sum_{t=0}^{T} \|\pi(t)\|^2 \end{aligned} \tag{17}$$

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

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

Then $D(\lambda(t))$ is lower bounded. Since $D(\lambda(t))$ is upper bounded (assume the primal problem is feasible), by (17), $\sum_{t=0}^{T} \|\pi(t)\|^2$ is also upper bounded. Because $\sum_{t=0}^{T} \|\pi(t)\|^2$ is a series consisting of non-negative terms, we must have

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

□

We now prove Theorem 8.

Proof. We first show there must exist one accumulation point of the sequence $\{\lambda(t)\}$. According to (17), $D(\lambda(t))$ is lower bounded by $D(\lambda(0))$. Since the sum of $(v_i^{par} - \bar{v}_i)$ and $(\underline{v}_i - v_i^{par})$ is negative, the set $\{\lambda \geq 0 | D(\lambda) \geq D(\lambda(0))\}$ is compact, then the existence of accumulation points follows from Weierstrass Theorem [21].

We next show every accumulation point λ^* of $\{\lambda(t)\}$ maximizes the dual problem. Let subsequence $\lambda(t_k)$ converges to λ^* with $t_{k+1} - t_k \geq T, \forall k$. Define $t_{i,k} = \max\{t \in T_i | t \leq t_k\}$. By Assumption 1, $\{t_{i,k}\}_{k=1}^{\infty}$ is a strictly increasing sequence and hence $t_{i,k} \to \infty$ as $k \to \infty$. With Lemma 10, $\|\nabla D(\lambda(t)) - \mu(t)\| \leq 2\|X\|_F \sum_{t'=t-2T_d-T_a}^{t-1} \|\pi(t')\| \to 0$. Therefore, $\lim_k \bar{\mu}_i(t_k) = \lim_k \partial D(\lambda(t_k))/\partial \bar{\lambda}_i = \partial D(\lambda^*)/\partial \bar{\lambda}_i$. By Lemma 9 (ii), $\|v(t+1) - v(t)\| \to 0$, hence $\|\mu(t_k) - \mu(t_{i,k})\| \leq \sum_{t'=t_k-T_d-T_a}^{t_k-1} \|v(t'+1) - v(t')\| \to 0$. Therefore $\lim_k \bar{\mu}_i(t_{i,k}) = \lim_k \bar{\mu}_i(t_k) = \partial D(\lambda^*)/\partial \bar{\lambda}_i$. Also since $\bar{\lambda}_i(t_k) = \bar{\lambda}_i(t_{i,k})$,

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

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 [18], 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 (14). Let λ' be a maximizer of $D(\lambda)$, then $\underline{\lambda}' - \bar{\lambda}'$ is the maximizer 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 [18], $\forall i$, $\bar{\lambda}_i'([X(\underline{\lambda}' - \bar{\lambda}')]_i + v_i^{par} - \bar{v}_i) = 0$ 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, λ' is unique, i.e. the dual problem (14) has a unique maximizer.

At last, since any accumulation point of $\{\lambda(t)\}$ is the maximizer of (14), the accumulation point of $\{\lambda(t)\}$ is unique. According to [21] this means $\lambda(t)$ will converge to the maximizer of the dual problem (14). Also, by (10), (11) and Assumption 2, $q(t)$ will converge to the minimizer of the primal problem (13).

□