An Optimal and Distributed Feedback Voltage Control under Limited Reactive Power

Guannan Qu, Na Li

***Abstract*—In this paper, we propose a distributed voltage control in power distribution networks through reactive power compensation. The proposed control can (i) operate in a distributed fashion where each bus makes its decision based on local voltage measurements and communication with neighbors, (ii) always satisfy the reactive power capacity constraint, (iii) stabilize the voltage magnitude in an acceptable range, and (iv) minimize an operational cost.**

***Index Terms*—Real-Time Voltage Control, Distribution Network**

I. Introduction

The primary purpose of voltage control is to maintain acceptable voltages at all buses along a distribution feeder under all possible operating conditions [1]–[7]. 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. To overcome the challenges, it has been proposed to utilize the computing, (local) sensing, and (local) communication capabilities of the inverters in the DERs to adjust the reactive power injected into the grid in order to maintain voltage stability [8], [9].

Various voltage control methods have been proposed [10]–[15]. One popular approach is using optimization methods. Typically, an Optimal Power Flow (OPF) problem with voltage constraints is formulated and then solved, either centralizedly (e.g. [10]–[13]) or distributedly (e.g. [16]–[19]). Here the "solving" process is purely computational (and usually involves communication if using distributed solvers). It is after the optimal solution is obtained that the reactive power is actually adjusted. Hence, this line of work is not feedback in nature - the algorithms assume knowledge of the disturbance (e.g. uncontrollable loads) and do not measure the system states (like voltage magnitudes). In this paper, we mainly focus on distributed *feedback* voltage control algorithms, in which each DER does not know the disturbance explicitly, but takes local measurements and adjusts its reactive power output based on the local measurements and local information, and perhaps with communication to its neighboring buses.

There indeed have been many efforts on developing feedback voltage control. One class is the traditional "Droop" control [14], [15], as advocated by IEEE 1547.8 Standard [20]. It monitors the local bus voltages and adjusts the reactive power injection accordingly. However, [3] shows that droop type algorithms are not able to maintain a feasible voltage profile under certain circumstances; ref. [6, Section V-A] shows droop control might experience stability and efficiency issues when the network is large. Therefore, other more sophisticated algorithms have been proposed, e.g. [3]–[7], [16], [21], [22]. [3, Algorithm 1] and [4] propose an integration type algorithm that reaches a feasible voltage profile. [3, Algorithm 2] and [5] utilize a dual ascent approach that minimizes a power loss related cost while reaching a feasible voltage profile. However these methods also have their limitations. For instance, [3]–[5] ignore the hard constraints on the reactive power injection capacity; [6], [7] does not meet the hard voltage constraint.[1] Though [21] meets both the voltage constraint and the hard reactive power limit, there is a lack of theoretic guarantee for convergence.

Besides the concerns on stability and voltage/reactive power constraints, another issue is the *optimality* of voltage control. Since there is an acceptable range for voltage and reactive power, there is flexibility on the operating point of voltage control. Some operating points will have a lower operational cost than others. For example, it is preferable for DERs to operate at certain power factor, which requires its reactive power injection to be close to a certain value. Though some existing methods, e.g., [3], [5]–[7], [16], [21], do optimize a particular objective, the objective can not be freely chosen by DERs and does not necessarily reflect the true cost of DERs. It will be appealing if the voltage control method not only stabilizes the voltage in the acceptable range, but also minimizes a cost that reflects the true operation cost.

Our Contribution: To overcome these challenges, we propose a distributed feedback voltage control that unifies the above algorithms in the sense that it can simultaneously (i) meet the voltage constraint asymptotically, (ii) satisfy the reactive power capacity constraint throughout, and (iii) minimize an operation cost that can be composed of a power loss related cost and reactive power operation costs. The algorithm takes the voltage measurements as inputs and determines the reactive power injection through local communication and computation. The communication graph is the same as the physical distribution network, meaning that each bus only needs to communicate with its 1-hop neighbors. The algorithm builds on the augmented Lagrangian multiplier theory [23, Sec. 3.2] and primal-dual gradient algorithms [24]–[26], [33]. We mathematically prove the performance of the algorithm in linear branch flow models [27] and numerically simulate the algorithm on a real distribution feeder.

We also mention that the use of communication in our

Guannan Qu and Na 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).

[1] Instead, [6], [7] incorporate a weighted voltage deviation as a soft penalty.

algorithm is inevitable. In fact, [21] shows that for a class of communication-free algorithms, there are scenarios in which those algorithms can *not* reach a feasible operating point (that satisfies both the voltage and the capacity constraint), despite the existence of a feasible operating point. The results in this paper are consistent with the performance limit in [21] and serve as an example as how to incorporate communication into algorithm design.

The remainder of the paper is organized as follows. In Section II we introduce the linearized power flow model that we use in the analysis, and present the problem formulation. In Section III we present our algorithm and also explain the rationale behind our algorithm design. In Section IV we analyze the convergence of our algorithm. Lastly, we provide simulations in Section V and conclude the paper in Section VI.

Notations. In this paper, letter q will be reserved for reactive power injections, letter p for active power injections, and v for squared voltage magnitudes. Notation $\mathcal{N}$ denotes the set of buses in the network, n denotes the number of buses (excluding the substation), letters i, j will be reserved for individual buses, and $\mathcal{N}_i$ for $i \in \mathcal{N}$ denotes the set of neighboring buses of i (including i itself). Notation $\|\cdot\|$ denotes Euclidean norm for vectors, and spectral norm for matrices. Notation $\mathbb{R}^N_{\geq 0}$ denotes the N-dimensional nonnegative orthant $\{x = [x_1, \ldots, x_n]^T : x_i \geq 0\}$. Notation $[x]^+$ means the projection of x onto the nonnegative orthant of an appropriate dimension. Notation $[x]^{\bar{x}}_{\underline{x}}$ for scalars x, $\bar{x}$, $\underline{x}$ where $\underline{x} \leq \bar{x}$ means projection of x onto the interval $[\underline{x}, \bar{x}]$. Notation $[y]^{\bar{y}}_{\underline{y}}$ for vectors y, $\underline{y}$, $\bar{y}$ of dimension N means projection of y on to the box constraint $\{z \in \mathbb{R}^N : z_i \in [\underline{y}_i, \bar{y}_i]\}$.

II. Preliminaries: Power Flow Model and Problem Formulation

A. Linearized Branch flow model for radial networks

Consider a radial distribution circuit that consists of a set $\mathcal{N} = \{0, 1, \ldots, n\}$ of buses and a set $\mathcal{E} \in \mathcal{N} \times \mathcal{N}$ of distribution lines connecting these buses. Bus 0 represents the substation and other buses in $\mathcal{N}$ represent branch buses. For each line $(i, j) \in \mathcal{E}$, 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 $i \in \mathcal{N}$, let V_i be the complex voltage and $s_i = p_i + \mathbf{i}q_i$ be the complex power injection. 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 radial distribution circuit [28], [29]:

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

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

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

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

where $\ell_{ij} := |I_{ij}|^2$, $v_i := |V_i|^2$. Eq. (1) defines a system of equations in the variables $(P, Q, \ell, v) := (P_{ij}, Q_{ij}, \ell_{ij}, v_i, (i, j) \in \mathcal{E}, i = 1, \ldots, n)$, which do not include phase angles of voltages and currents. Given (P, Q, ℓ, v) the phase angles can be uniquely determined for radial networks [28], [30].

If the power loss term ℓ_{ij} is set to be 0, the model can be approximated by the following linear model, known as Simplified Distflow [27].

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

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

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

From (2), we can derive that the voltage vector $v = [v_1, \ldots, v_n]^T$ and power injection vector $p = [p_1, \ldots, p_n]^T, q = [q_1, \ldots, q_n]^T$ 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 $\mathcal{P}_i \subset \mathcal{E}$ is the set of lines on the unique path from bus 0 to bus i. The detailed derivation is given in [31]. In [5], [31], it has been shown that when the resistances and reactances of the lines in the network are all positive, R, X are positive definite matrices and the inverse matrices have the same sparse structure of the distribution network. This means that, letting $Y := X^{-1}$, then $Y_{i,j} = 0$ for all $(i, j) \notin \mathcal{E}$.

B. Problem formulation

We separate q into two parts, $q = q^c + q^e$, where q^c denotes the reactive power injection governed by the Volt/Var control components and q^e denotes any other reactive power injection. Let $v^{par} := Rp + Xq^e + v_0$, then, $v = Xq^c + v^{par}$. Without causing any confusion, we will simply use q instead of q^c to denote the reactive power injected by the Volt/Var control devices. In the rest of the paper, for algorithm development and theoretical analysis we will use the linearized model to represent the physical power network: given control input q, the physical network determines voltage $v(q)$ given by,

$$v(q) = Xq + v^{par}. \quad (4)$$

For each i, we introduce scalar $\underline{q}_i$ and $\bar{q}_i$, the lower and upper reactive power capacity limit for the device at node i. We also introduce $f_i : \mathbb{R} \to \mathbb{R}$, the operating cost.[2] We assume f_i

[2]For example, it is preferable for DERs to operate at certain power factor, which requires its reactive power injection to be close to a certain value.

to be a μ-strongly convex and L-smooth function. This means that, for any $x, y \in \mathbb{R}$,

$$f_i(y) - f_i(x) \geq \langle \nabla f_i(x), y - x \rangle + \frac{\mu}{2}\|y - x\|^2$$
$$\|f_i(x) - f_i(y)\| \leq L\|x - y\|.$$

The Volt/Var control can be modeled as a control problem on a quasi-dynamical system. At time t, let the control action be $q(t)$, then it determines the voltage profile $v(t) := v(q(t))$ by (4). [3] Then given the voltage profile $v(t)$ and other available information, the controller determines a new reactive power injection $q(t+1)$ and the new $q(t+1)$ results in a new voltage profile $v(t+1) = v(q(t+1))$. Mathematically, the Volt/Var control problem is formulated as the following closed loop dynamical system,

$$v(t) = v(q(t)) = Xq(t) + v^{par}$$
$$q_i(t+1) = \text{Controller}_i(\text{information available to } i \text{ at time } t)$$

The objective of Volt/Var control is to design a controller that meets the following three requirements.

Requirement 1: Information. The controller at i shall only use information that is accessible locally or from neighboring buses in the network. This includes local decision variable $q_i(t)$, local voltage measurement $v_i(t)$, other local auxiliary (computational) variables and variables from neighboring buses.

Requirement 2: Hard Capacity constraint. Since there is always a limit on the available reactive power injection, for any t, this capacity constraint shall *not* be violated, i.e. $\underline{q}_i \leq q_i(t) \leq \bar{q}_i$.

Requirement 3: Optimality. Under any system operating condition v^{par}, the controller drives the distribution system to the optimal point of the following optimization problem,

$$\min_{q_i} \quad f(q) \triangleq \sum_{i=1}^{n} f_i(q_i) + \frac{d}{2} q^T X q \tag{5a}$$
$$s.t. \quad \underline{v}_i \leq v_i(q) \leq \bar{v}_i \tag{5b}$$
$$\underline{q}_i \leq q_i \leq \bar{q}_i \tag{5c}$$

In the optimization problem, the cost function (5a) is composed of the sum of the operating costs f_i, as well as a network level cost $\frac{1}{2}q^T X q$, with a weighting parameter $d > 0$ balancing the two costs. The cost $\frac{1}{2}q^T X q$ is related to the network loss [5]. Constraint (5b) means that the voltage must lie within the upper and lower limit, and constraint (5c) is the reactive power capacity limit. We have the following minor assumption on problem (5).

Assumption 1. *There exists a strictly feasible solution q for problem* (5)*, meaning that the feasible solution satisfies constraints (5b,5c) with strict inequalities.*

Remark 1. *For easy exposition and without loss of generality, we assume that there is a control component at each bus i. For the buses where there is no control components, we can set $\underline{q}_i = \bar{q}_i = 0$.*

Remark 2. *The linear model (4) is only for algorithm development and theoretical analysis. Since our controller is a feedback controller, it can be implemented for real-world systems. We use the nonlinear power flow model to test our proposed controller numerically in Section V.*

III. Algorithm Development

In this section, we formally introduce our algorithm, Optimal Distributed Feedback Voltage Control (OptDist-VC). For each bus i, we introduce auxiliary variables, $\hat{q}_i, \xi_i, \bar{\lambda}_i, \underline{\lambda}_i$. At each iteration t, node i measures the local voltages $v_i(t)$, passes the variables $\hat{q}_i(t)$, $\xi_i(t)$ and $f_i'(\hat{q}_i(t))$ to its neigboring buses in the network, and then computes variables $\hat{q}_i(t+1)$, $q_i(t+1)$, $\xi_i(t+1)$, $\bar{\lambda}_i(t+1)$, $\underline{\lambda}_i(t+1)$ and injects the reactive power $q_i(t+1)$ according to the following equations,

OptDist-VC:

$$\hat{q}_i(t+1) = \hat{q}_i(t) - \alpha\Big\{\bar{\lambda}_i(t) - \underline{\lambda}_i(t) + d\hat{q}_i(t) + \sum_{j\in\mathcal{N}_i} Y_{ij}[f_j'(\hat{q}_j(t)) + G_j(\xi_j(t), \hat{q}_j(t))]\Big\} \tag{6a}$$
$$q_i(t+1) = [\hat{q}_i(t+1)]_{\underline{q}_i}^{\bar{q}_i} \tag{6b}$$
$$\xi_i(t+1) = \xi_i(t) + \beta H_i(\xi_i(t), \hat{q}_i(t)) \tag{6c}$$
$$\bar{\lambda}_i(t+1) = [\bar{\lambda}_i(t) + \gamma(v_i(t) - \bar{v}_i)]^+ \tag{6d}$$
$$\underline{\lambda}_i(t+1) = [\underline{\lambda}_i(t) + \gamma(\underline{v}_i - v_i(t))]^+ \tag{6e}$$

where $[\cdot]^+$ means projection onto the nonnegative orthant of an appropriate dimension; $[\cdot]_{\underline{q}_i}^{\bar{q}_i}$ means projection onto the $[\underline{q}_i, \bar{q}_i]$ box constraint. Quantity α, β, γ and c are positive scalar parameters. Quantity $G_i(\xi_i, q_i)$ and $H_i(\xi_i, q_i)$ are defined by,

$$G_i(\xi_i, q_i) = \begin{cases} \xi_i + c(q_i - \underline{q}_i) & q_i + \frac{\xi_i}{c} < \underline{q}_i \\ 0 & \underline{q}_i \leq q_i + \frac{\xi_i}{c} \leq \bar{q}_i \\ \xi_i + c(q_i - \bar{q}_i) & q_i + \frac{\xi_i}{c} > \bar{q}_i \end{cases} \tag{7}$$

$$H_i(\xi_i, q_i) = \begin{cases} q_i - \underline{q}_i & q_i + \frac{\xi_i}{c} < \underline{q}_i \\ -\frac{\xi_i}{c} & \underline{q}_i \leq q_i + \frac{\xi_i}{c} \leq \bar{q}_i \\ q_i - \bar{q}_i & q_i + \frac{\xi_i}{c} > \bar{q}_i \end{cases}. \tag{8}$$

It can be readily seen that our algorithm OptDist-VC meets Requirement 1 since update equation (6) only needs local and neighboring information. Moreover, (6b) implies that $q_i(t)$ always lies within $[\underline{q}_i, \bar{q}_i]$ and hence Requirement 2 is met. At last, the following Theorem 1 shows that $q(t)$ will converge to the optimal solution of (5) and hence Requirement 3 is satisfied. *In conclusion, OptDist-VC meets all the three design requirements.*

Theorem 1. *In OptDist-VC (6), for any $c > 0$, when α, β, γ are small enough and satisfy mild conditions,*[4] *$q(t)$ will converge to the unique optimizer of (5).*

[3] Without causing any confusion, we abuse the notation $v(\cdot)$ to denote both the network model (4) mapping q to the voltage profile $v(q)$, and the voltage profile at a certain time step $v(t)$.

[4] For the detailed conditions on the step sizes, please see Remark 3 or the technical report of this paper [32, Appendx-B and Appendix-C].

OptDist-VC (6) is based on a primal-dual gradient algorithm [24]–[26], [33] for an augmented Lagrangian [23], in which $\hat{q}_i(t)$ is the primal variable, $\xi_i(t), \bar{\lambda}_i(t), \underline{\lambda}_i(t)$ are the dual variables, and α, β, γ are the step sizes. In Section III-A, we will describe in detail the rationale behind our algorithm design.

Remark 3. *In our theoretic proof in [32, Appendx-B and Appendix-C], our requirement on α,β are such that, $\frac{\alpha}{\beta} \geq \frac{8\|X\|}{\mu c}$, $\alpha \cdot \frac{\alpha}{\beta} = \frac{\|X\|^2}{16\ell\mu^2\kappa_Y^2\kappa_f^3}$ where constants ℓ, κ_Y, κ_f are given by $\ell = L + c + d\|X\|$, $\kappa_Y = \|Y\|\|X\|$, $\kappa_f = \frac{\ell}{\min(c,\mu)}$. After α, β are selected, we need γ satisfying $\gamma < \min(\frac{\mu}{4\|X\|^2}, \frac{\|X\|^2(1-\rho(\alpha,\beta))^2}{\mu C_2(\alpha,\beta)^2})$, where $\rho(\alpha,\beta) = 1 - \frac{\beta^2\|X\|^2}{64\ell^2(\alpha\mu)^2\kappa_Y^2\kappa_f^3}$, and $C_2(\alpha,\beta) = 2\|X\|\sqrt{\frac{\|Y\|}{\ell}}\sqrt{\Big[4\ell\frac{\|X\|^3}{\mu^2} + 4\frac{\alpha}{\beta}\ell(1+\frac{L+d\|X\|}{\mu})^2\|X\|^2\Big]}$. The condition basically says that α, β, γ should be small, and moreover the ratio $\frac{\alpha}{\beta}$ shouldn't be too large or too small, and should be slightly larger than $\frac{8\|X\|}{\mu c}$. However, we remark that our theoretic step size requirement is conservative. In simulations, step sizes α, β, γ obtained through trial and error that are much larger than the theoretic bounds can also guarantee convergence.*

A. Rationale Behind Algorithm Design

We introduce Lagragian multipliers $\lambda = [\underline{\lambda}^T, \bar{\lambda}^T]^T \in \mathbb{R}^{2n}$ for optimization problem (5), in which $\underline{\lambda} = [\underline{\lambda}_1, \ldots, \underline{\lambda}_n]^T$ and $\underline{\lambda}_i$ corresponds to the lower limit in the voltage constraint (5b); $\bar{\lambda} = [\bar{\lambda}_1, \ldots, \bar{\lambda}_n]^T$ and $\bar{\lambda}_i$ corresponds to the upper limit in (5b). We also introduce multiplier $\xi = [\xi_1, \ldots, \xi_n]^T$ for the reactive power constraint (5c), and ξ_i is for both the lower limit and the upper limit in the capacity constraint (5c). Next, we introduce the augmented Lagrangian for the optimization problem,

$$\mathcal{L}(q, \xi, \lambda) = f(q) + \underline{\lambda}^T(\underline{v} - v(q)) + \bar{\lambda}^T(v(q) - \bar{v}) + K(\xi, q) \tag{9}$$

where $K(\xi, q) = \sum_{i=1}^n K_i(\xi_i, q_i)$, and $K_i(\xi_i, q_i)$ is a quadratic penalty function defined to be,

$$\begin{aligned} K_i(\xi_i, q_i) &= \min_{u \in [q_i - \bar{q}_i, q_i - \underline{q}_i]} \xi_i u + \frac{c}{2}u^2 \\ &= \begin{cases} \xi_i(q_i - \underline{q}_i) + \frac{c}{2}(q_i - \underline{q}_i)^2 & q_i + \frac{\xi_i}{c} < \underline{q}_i \\ -\frac{\xi_i^2}{2c} & \underline{q}_i \leq q_i + \frac{\xi_i}{c} \leq \bar{q}_i \\ \xi_i(q_i - \bar{q}_i) + \frac{c}{2}(q_i - \bar{q}_i)^2 & q_i + \frac{\xi_i}{c} > \bar{q}_i \end{cases} \end{aligned} \tag{10}$$

We note here that the $G_i(\cdot,\cdot)$ in (7) and $H_i(\cdot,\cdot)$ in (8) are in fact partial derivatives of $K_i(\cdot,\cdot)$, $\frac{\partial K_i(\xi_i,q_i)}{\partial q_i} = G_i(\xi_i, q_i)$, $\frac{\partial K_i(\xi_i,q_i)}{\partial \xi_i} = H_i(\xi_i, q_i)$.

In (9), $\underline{\lambda}^T(\underline{v} - v(q)) + \bar{\lambda}^T(v(q) - \bar{v})$ is the standard term in Lagrangian multiplier theory that penalizes violation of the voltage constraint, while the $K_i(\xi_i, q_i)$ term is a special quadratic penalty function that penalizes violation of both the upper limit and the lower limit of constraint (5c). For details of such quadratic penalty functions, we refer the readers to [23, Section 3.2], [33, Appendix-G]. In short, similar to the standard Lagrangian case, the max-min problem

$$\max_{\lambda \in \mathbb{R}^{2n}_{\geq 0}, \xi \in \mathbb{R}^n} \min_{q \in \mathbb{R}^n} \mathcal{L}(q, \xi, \lambda) \tag{11}$$

is equivalent to the original optimization problem (5) (cf. Lemma 2). The reason we use the augmented Lagrangian instead of the standard Lagrangian is that the augmented Lagrangian has better second order characteristics, which leads to better dynamic performance when applying primal dual gradient algorithms to $\mathcal{L}$.

We then write down the standard primal dual gradient algorithm [24]–[26] for solving the max-min problem (11),

$$\begin{aligned} \hat{q}_i(t+1) &= \hat{q}_i(t) - \alpha\frac{\partial \mathcal{L}(\hat{q}(t), \xi(t), \lambda(t))}{\partial q_i} \\ &= \hat{q}_i(t) - \alpha\Big[f_i'(\hat{q}_i(t)) + G_i(\xi_i(t), \hat{q}_i(t)) \\ &\quad + \sum_{j=1}^n X_{ij}(\bar{\lambda}_j(t) - \underline{\lambda}_j(t) + d\hat{q}_j(t))\Big] \quad (12a) \\ \xi_i(t+1) &= \xi_i(t) + \beta\frac{\partial \mathcal{L}(\hat{q}(t), \xi(t), \lambda(t)}{\partial \xi_i} \\ &= \xi_i(t) + \beta H_i(\xi_i(t), \hat{q}_i(t)) \quad (12b) \\ \bar{\lambda}_i(t+1) &= [\bar{\lambda}_i(t) + \gamma\frac{\partial \mathcal{L}(\hat{q}(t), \xi(t), \lambda(t))}{\partial \bar{\lambda}_i}]^+ \\ &= [\bar{\lambda}_i(t) + \gamma(v_i(\hat{q}(t)) - \bar{v}_i)]^+ \quad (12c) \\ \underline{\lambda}_i(t+1) &= [\underline{\lambda}_i(t) + \gamma\frac{\partial \mathcal{L}(\hat{q}(t), \xi(t), \lambda(t))}{\partial \underline{\lambda}_i}]^+ \\ &= [\underline{\lambda}_i(t) + \gamma(\underline{v}_i - v_i(\hat{q}(t)))]^+ \quad (12d) \end{aligned}$$

Though literature has shown the convergence of primal-dual gradient algorithms with properly chosen step sizes and under some conditions [24]–[26], [33], the algorithm in (12) does not meet our design requirements. Firstly, step (12a) requires information across the network to implement, violating Requirement 1. Secondly, the $\hat{q}_i(t)$ in step (12a) might violate the capacity constraint, violating Requirement 2.

We now propose two modifications to (12) to meet the design requirements and the two modifications together change (12) into OptDist-VC (6).

Modification (a). We now modify (12a) such that each bus only needs local and neighbor's information to update. Eq. (12a) is a gradient update for the q coordinates of $\mathcal{L}$, and the gradient is given by,

$$\begin{aligned} &\nabla_q \mathcal{L}(q, \xi(t), \lambda(t)) \\ &= \nabla f(q) + X(\bar{\lambda}(t) - \underline{\lambda}(t)) + G(\xi(t), q). \end{aligned} \tag{13}$$

where $\nabla f(q) = [f_1'(q_1), \ldots, f_n'(q_n)]^T + dXq$, and $G(\xi, q) = \nabla_q K(\xi, q) = [G_1(\xi_1, q_1), \ldots, G_n(\xi_n, q_n))]^T$. Because of the sparse structure of $Y := X^{-1}$, the scaled gradient $Y\nabla_q \mathcal{L}(q, \xi(t), \lambda(t))$ is given by,

$$\begin{aligned} [Y\nabla_q \mathcal{L}(q, \xi(t), \lambda(t))]_i = \Big\{&\bar{\lambda}_i(t) - \underline{\lambda}_i(t) + dq_i \\ &+ \sum_{j \in \mathcal{N}_i} Y_{ij}[f_j'(q_j) + G_j(\xi_j(t), q_j)]\Big\} \end{aligned}$$

To calculate the i'th element of the scaled gradient $Y\nabla\mathcal{L}_q(q,\xi(t),\lambda(t))$, agent i only needs local information $(\bar{\lambda}_i(t),\underline{\lambda}_i(t))$ and information from neighbors $(f_j'(q_j),G_j(\xi_j,q_j)$ where $j\in\mathcal{N}_i)$. Moreover, since Y is positive definite, $Y\nabla_q\mathcal{L}(q,\xi(t),\lambda(t))$ is still a descent direction for $\mathcal{L}$ (in the q coordinates) and hence using the scaled gradient in the primal dual gradient algorithm still has convergence guarantee [25], [26]. Therefore, we change (12a) into the following "scaled" gradient update, which gives rise to step (6a) in OptDist-VC.

$$\begin{aligned}\hat{q}_i(t+1) &= \hat{q}_i(t) - \alpha[Y\nabla_q\mathcal{L}(\hat{q}(t),\xi(t),\lambda(t))]_i \\ &= \hat{q}_i(t) - \alpha\Big\{\bar{\lambda}_i(t) - \underline{\lambda}_i(t) + d\hat{q}_i(t) \\ &\quad + \sum_{j\in\mathcal{N}_i} Y_{ij}[f_j'(\hat{q}_j(t)) + G_j(\xi_j(t),\hat{q}_j(t))]\Big\}. \quad (14)\end{aligned}$$

Modification (b). To fix the problem that $\hat{q}_i(t)$ may violate the capacity constraint, we do not actually implement $\hat{q}_i(t)$, but instead implement $q_i(t) = [\hat{q}_i(t)]_{\underline{q}_i}^{\bar{q}_i}$, the projection of $\hat{q}_i(t)$ onto the capacity constraint. This gives rise to (6b) in our algorithm. Another issue is that update (12c) (12d) uses $v_i(\hat{q}(t))$, which is not the measured voltage since the implemented reactive power is not $\hat{q}(t)$. Therefore, we replace the $v_i(\hat{q}(t))$ in (12c) (12d) with $v_i(q(t))$ and get,

$$\begin{aligned}\bar{\lambda}_i(t+1) &= [\bar{\lambda}_i(t) + \gamma\frac{\partial\mathcal{L}(q(t),\xi(t),\lambda(t))}{\partial\bar{\lambda}_i}]^+ \\ &= [\bar{\lambda}_i(t) + \gamma(v_i(q(t)) - \bar{v}_i)]^+ \\ \underline{\lambda}_i(t+1) &= [\underline{\lambda}_i(t) + \gamma\frac{\partial\mathcal{L}(q(t),\xi(t),\lambda(t))}{\partial\underline{\lambda}_i}]^+ \\ &= [\underline{\lambda}_i(t) + \gamma(\underline{v}_i - v_i(q(t)))]^+\end{aligned}$$

which uses the measured voltage $v_i(q(t)) = v_i(t)$ and gives rises to (6d) (6e). We emphasize that after the change, the update for $\underline{\lambda}_i(t),\bar{\lambda}_i(t)$ does not use the true gradient of $\mathcal{L}$, $\frac{\partial\mathcal{L}(\hat{q}(t),\xi(t),\lambda(t))}{\partial\lambda}$ any more; but uses a gradient that is evaluated at a different point, $\frac{\partial\mathcal{L}(q(t),\xi(t),\lambda(t))}{\partial\lambda}$. We will show in Section IV that despite of the inconsistency, we will have $q(t)-\hat{q}(t)\to 0$ and the algorithm still converges.

IV. Algorithm Analysis

We first analyze the Lagrangian (9), and show that the original problem (5) is indeed equivalent to the max-min problem (11). For notational simplicity, we define $\tilde{X} = [X,-X]^T$, $v_b = [-(v^{par})^T+\bar{v}^T, -\underline{v}^T+(v^{par})^T]^T$ and rewrite Lagrangian (9) as,

$$\mathcal{L}(q,\xi,\lambda) = f(q) + \lambda^T(\tilde{X}q - v_b) + K(\xi,q). \quad (15)$$

We define $\mathcal{D}(\xi,\lambda) = \min_{q\in\mathbb{R}^n}\mathcal{L}(q,\xi,\lambda)$ and $\mathcal{S}(\lambda) = \max_{\xi\in\mathbb{R}^n}\mathcal{D}(\xi,\lambda)$. We introduce the following Lemma regarding $\mathcal{L},\mathcal{D},\mathcal{S}$, the proof of which can be found in the technical report of this paper [32, Appendix-A].

Lemma 2. *The following statements hold.*

(a) *For every λ, there is a unique pair $q^*(\lambda)$ and $\xi^*(\lambda)$ that satisfies $\nabla_q\mathcal{L}(q^*(\lambda),\xi^*(\lambda),\lambda) = \nabla_\xi\mathcal{L}(q^*(\lambda),\xi^*(\lambda),\lambda) = 0$. Moreover, $q_i^*(\lambda)\in[\underline{q}_i,\bar{q}_i]$.*

(b) *$\xi^*(\lambda)$ and $q^*(\lambda)$ is Lipschitz in λ.*

$$\|q^*(\lambda) - q^*(\lambda')\| \le \frac{\|\tilde{X}\|}{\mu}\|\lambda-\lambda'\| \quad (16)$$

$$\|\xi^*(\lambda) - \xi^*(\lambda)\| \le (1+\frac{L+d\|X\|}{\mu})\|\tilde{X}\|\|\lambda-\lambda'\| \quad (17)$$

(c) *$\mathcal{S}(\lambda)$ is a concave function, and $\nabla\mathcal{S}(\lambda) = \tilde{X}q^*(\lambda) - v_b$, and $\mathcal{S}$ is $\frac{\|\tilde{X}\|^2}{\mu}$-smooth.*

(d) *$\mathcal{S}(\lambda)$ is upper bounded over $\lambda\in\mathbb{R}^{2n}_{\ge 0}$ and moreover, for any real number $a\in\mathbb{R}$, level set $Level_a = \{\lambda\in\mathbb{R}^{2n}_{\ge 0} : \mathcal{S}(\lambda)\ge a\}$ is bounded.*

(e) *Let λ^* be any solution of $\max_{\lambda\in\mathbb{R}^{2n}_{\ge 0}}\mathcal{S}(\lambda)$, then $q^*(\lambda^*)$ is the unique solution of optimization problem (5).*

Lemma 2 (e) shows that finding the optimal solution of (5) is equivalent to solving the max-min problem (11). We now show our algorithm OptDist-VC (6) solves the max-min problem (11).

As mentioned in Section III-A, OptDist-VC (6) is not the standard primal dual algorithm because in update (6d) (6e), the gradient of $\mathcal{L}$ is not evaluated at $\hat{q}(t)$, but at $q(t)$ instead. For this reason, we regard the update of $\lambda(t)$ as the gradient projection algorithm for $S(\lambda)$ where the gradient contains some "error". For notational simplicity, we abuse the notation $q^*(\cdot),\xi^*(\cdot)$ and define $q^*(t) := q^*(\lambda(t)),\xi^*(t) := \xi^*(\lambda(t))$. Then, the gradient of $\mathcal{S}$ at $\lambda(t)$ is (by Lemma 2 (c))

$$\nabla\mathcal{S}(\lambda(t)) = \tilde{X}q^*(t) - v_b.$$

In the meanwhile, notice the update for $\lambda(t)$ is

$$\lambda(t+1) = [\lambda(t) + \gamma(\tilde{X}q(t) - v_b)]^+$$

So the update for $\lambda(t)$ is "inexact" projected gradient method for function $\mathcal{S}$, and the inexact gradient is $\tilde{X}q(t)-v_b$. Therefore, the gradient error is $\epsilon(t) = \|\tilde{X}(q(t)-q^*(t))\|$. Notice that by Lemma 2 (a), $q^*(t)$ lies within the capacity constraint and therefore by projection theorem, $\|q(t)-q^*(t)\| \le \|\hat{q}(t)-q^*(t)\|$. Hence the gradient error satisfies

$$\epsilon(t) \le \|\tilde{X}\|\|\hat{q}(t)-q^*(t)\|. \quad (18)$$

Based on the above analysis, we divide the rest of the proof into two steps: bounding the gradient error (i.e. bounding $\|\hat{q}(t)-q^*(t)\|$); and then finishing the proof of convergence of our algorithm despite the gradient error. Due to space limit, we cannot provide the full proof here, and the full proof is included in the report of this paper [32, Appendix-B and Appendix-C]. In short, we will prove the gradient error is bounded by [32, Appendix-B],

$$\epsilon(t) \le C_1\rho^t + C_2\sum_{k=0}^{t-1}\rho^{t-1-k}\|\lambda(k+1)-\lambda(k)\|$$

for some $C_1,C_2 > 0$ and $\rho\in(0,1)$. Inexact gradient descent with this type of error has already be shown to have convergence guarantee, according to some related studies (e.g. [34, Sec. IV-D]). For completeness we give a full proof of convergence in the report [32, Appendix-C].

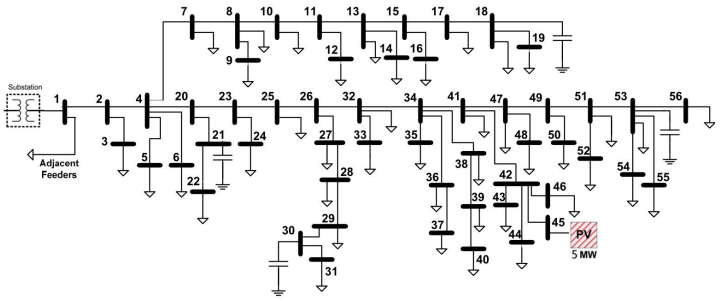

Figure 1: Schematic diagram of two SCE distribution systems.

V. CASE STUDY

We evaluate our algorithm OptDist-VC on a distribution circuit of South California Edison with a high penetration of photovoltaic (PV) generation [35]. Figure 1 shows a 56-bus distribution circuit. Note that Bus 1 indicates the substation, and there is a photovoltaic (PV) generator located at bus 45 and there are shunt capacitors located at bus 19, 21, 30, 53. See [35] 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 all the buses and those control components can supply or consume at most 1 MVar reactive power (i.e. $\bar{q}_i = 1MW, \underline{q}_i = -1MW$). 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. The cost functions f_i are randomly generated quadratic functions, and parameter $d = 0.1$. 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) [36].

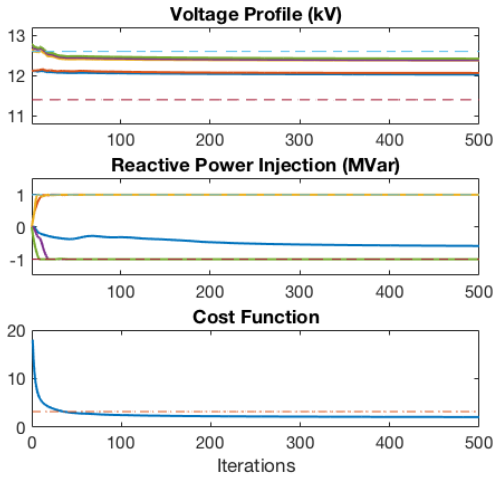

Figure 2: Case I: moderate loads, heavy PV generation. Blue: bus 4; red: bus 19; orange: bus 41; purple: bus 42; green: bus 45.

We simulate three different cases. In case I, the PV generator is generating a large amount of power but the loads are moderate, resulting in high voltages at some buses (Figure 2); in case II, the PV generator is generating a large amount of power and some buses are heavily loaded, resulting in high voltages at some buses and low voltages at other buses (Figure 3). The results of case I and case II are summarized in Figure 2 and Figure 3. It shows that OptDist-VC can bring the voltage to the acceptable range and in the meanwhile not violating the capacity constraint and minimizing the cost function. We note that in case I and case II, OptDist-VC can bring the cost function value slightly below the optimal solution of (5) (shown as dashed line in the lower plot of Fig. 2 and Fig.

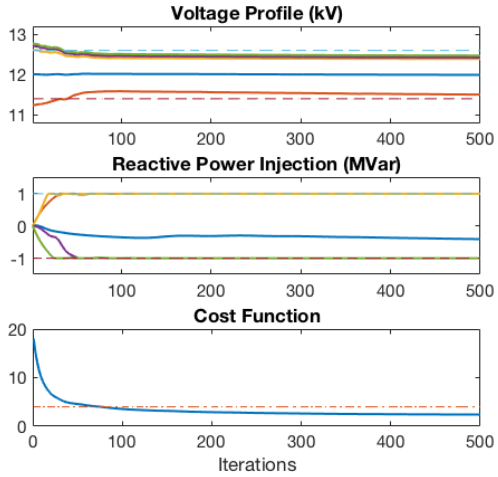

Figure 3: Case II: heavy loads, heavy PV generation. Blue: bus 4; Red: Bus 19; Orange: bus 41; Purple: bus 42; Green: Bus 45.

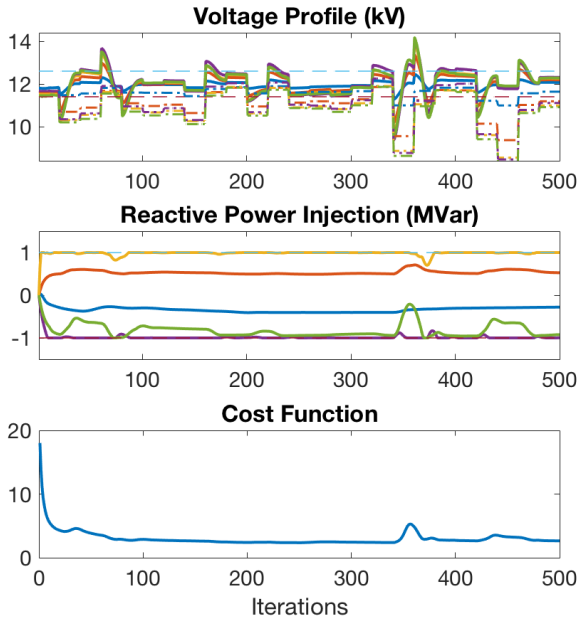

Figure 4: Case III: fluctuating load and PV generation. Blue: bus 4; red: bus 25; orange: bus 34; purple: bus 45; green: bus 53. In the upper plot, the dashed lines are the voltage profile if no voltage control algorithm is used.

3). This is because (5) uses a linearized model, while our simulation uses the nonlinear model. Nonetheless, it is clear from Figure 2 and Figure 3 that OptDist-VC is indeed bringing the cost function down.

In addition to the above two cases, we consider a third case where the load and the PV generation are fluctuating for every other 20 controller iterations, and the amount of fluctuation is drawn from a uniform distribution. The fluctuating speed is 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 OptDist-VC under a more realistic setting. For comparison, in this case we also simulate the network voltage profile when no voltage control algorithm is used. The simulation results are given in Figure 4. It shows that, after every change in the load and PV generation, although the voltage constraint is violated, our algorithm OptDist-VC can quickly bring the voltage into the acceptable range and in the mean while, not violating the capacity constraint, bringing the cost function down and maintaining it at a very low value.

VI. Conclusion

This paper proposes a Distributed Feedback Voltage Control Algorithm OptDist-VC that can (i) meet the voltage constraint asymptotically, (ii) satisfy the reactive power capacity constraint throughout, and (iii) minimize a cost that can be composed of a power loss related cost and reactive power operation costs. Future work includes extending the approach to jointly control active and reactive power.

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] N. Li, G. Qu, and M. Dahleh, "Real-time decentralized voltage control in distribution networks," in *Allerton*, 2014.

[4] 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.

[5] 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.

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

[7] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Fast localized voltage regulation in single-phase distribution grids," in *Smart Grid Communications (SmartGridComm), 2015 IEEE International Conference on*. IEEE, 2015, pp. 725–730.

[8] 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.

[9] 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.

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

[11] N. Li, L. Chen, and S. H. Low, "Exact convex relaxation of opf for radial networks using branch flow model," in *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*. IEEE, 2012, pp. 7–12.

[12] H.-G. Yeh, D. F. Gayme, and S. H. Low, "Adaptive var control for distribution circuits with photovoltaic generators," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1656–1663, 2012.

[13] S. Deshmukh, B. Natarajan, and A. Pahwa, "Voltage/var control in distribution networks via reactive power injection through distributed generators," *IEEE Transactions on smart grid*, vol. 3, no. 3, pp. 1226–1234, 2012.

[14] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *IEEE 52nd Conference on Decision and Control*, 2013.

[15] P. Jahangiri and D. C. Aliprantis, "Distributed volt/var control by pv inverters," *IEEE Transactions on power systems*, vol. 28, no. 3, pp. 3429–3439, 2013.

[16] B. Zhang, A. Y. Lam, A. D. Domínguez-García, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1714–1726, 2015.

[17] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Message passing for dynamic network energy management," *arXiv preprint arXiv:1204.1106*, 2012.

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

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

[20] "IEEE draft recommended practice for establishing methods and procedures that provide supplemental support for implementation strategies for expanded use of IEEE standard 1547," *IEEE P1547.8/D8, July 2014*, pp. 1–176, Nov 2014.

[21] G. Cavraro, S. Bolognani, R. Carli, and S. Zampieri, "The value of communication in the voltage regulation problem," in *IEEE Conference on Decision and Control (CDC)*, 2016.

[22] B. A. Robbins, C. N. Hadjicostis, and A. D. Domínguez-García, "A two-stage distributed architecture for voltage control in power distribution systems," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1470–1482, 2013.

[23] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[24] A. Nedić and A. Ozdaglar, "Subgradient methods for saddle-point problems," *Journal of optimization theory and applications*, vol. 142, no. 1, pp. 205–228, 2009.

[25] A. Cherukuri, E. Mallada, and J. Cortés, "Asymptotic convergence of constrained primal–dual dynamics," *Systems & Control Letters*, vol. 87, pp. 10–15, 2016.

[26] D. Feijer and F. Paganini, "Stability of primal–dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, no. 12, pp. 1974–1981, 2010.

[27] 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.

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

[29] 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.

[30] S. H. Low, "Convex relaxation of optimal power flowpart i: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.

[31] 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.

[32] N. L. Guannan Qu. (2017) An optimal and distributed feedback voltage control under limited reactive power. [Online]. Available: http://scholar.harvard.edu/files/gqu/files/pscc2018full.pdf

[33] G. Qu and N. Li, "On the exponential stability of primal-dual gradient dynamics," *arXiv preprint arXiv:1803.01825*, 2018.

[34] ——, "Harnessing smoothness to accelerate distributed optimization," *IEEE Transactions on Control of Network Systems*, 2017.

[35] 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.

[36] 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.