Voltage Control Using Limited Communication

Sindri Magnússon, Guannan Qu, Carlo Fischione, and Na Li

Abstract—In electricity distribution networks, the increasing penetration of renewable energy generation necessitates faster and more sophisticated voltage controls. Unfortunately, recent research shows that local voltage control fails in achieving the desired regulation, unless there is communication between the controllers. However, the communication infrastructure for distribution systems is less reliable and less ubiquitous as compared to that for the bulk transmission system. In this paper, we design distributed voltage control that uses limited communication. That is, only neighboring buses need to communicate a few bits between each other for each control step. We investigate how these controllers can achieve the desired asymptotic behavior of the voltage regulation and we provide upper bounds on the number of bits that are needed to ensure a predefined accuracy of the regulation. Finally, we illustrate the results by numerical simulations.

Index Terms—Distributed Optimization, Smart Grid, Voltage Regulation, Reactive Power, Limited Communication.

I. INTRODUCTION

There is an increasing penetration of distributed energy resources such as renewable energy in distribution networks. Unfortunately, such a penetration causes faster voltage fluctuations than what today's distribution networks can handle, see [1]. Therefore, to avoid overloading the distribution networks, the integration of renewable energy resources must be accompanied by faster and more sophisticated voltage regulation.

These challenges have motivated a growing research interest in voltage control, where fast voltage fluctuations are regulated through real-time reactive power injections to ensure that the voltage is maintained within an acceptable range. Such fast voltage control can be implemented in the emerging power devices such as inverters. The research efforts have focused on two main directions: local and distributed control strategies. In the local voltage control, control devices at each bus update the reactive power injections using only locally available information, such as local voltage measurements, see [2], [3], [4] and references therein. On the other hand, in distributed voltage control schemes, control devices at each bus determine the reactive power injection with additional information communicated from its neighboring buses in the distribution network, see [5], [6], [7], [8], [9]. Local control strategies have the obvious advantage over distributed ones in that they do not rely on communication. However, even though local control strategies perform well in some cases, they may

fail to ensure that the voltage is maintained within the accepted range in some cases, as proved by the impossibility result in [10]. Therefore, communication among the local controllers is needed to guarantee the performance of voltage regulation.

The communication capabilities of today's distribution networks generally suffer from low data rates, [11], [12]. To compensate for this deficiency, power system operators and industries are currently investing heavily in integrating the distribution networks with a sophisticated communication infrastructure. However, even with the promising capabilities of the future low latency networks, fast real-time control applications, like voltage control, rely on short packages that carry coarsely quantized information, [13]. Therefore, it is important to develop voltage control with very limited communication for early integration of renewable resources using today's grid limited communication capabilities and also for sustainable developments of the future smart grid.

Networked control systems with limited bandwidth communication have been well investigated [14], [15], [16]. However, standard algorithms used for stabilization of networked control systems are generally not applicable to the voltage regulation problem. Instead, most distributed voltage control algorithms in the literature are based on the tools of distributed optimization, e.g. [5], [6], [7], [8], [9]. Even though distributed optimization with limited bandwidth has received some attention [17], [18], [19], [20], none of these works fit the characteristics of the voltage control problem.

Other ways of limiting the communication of voltage control algorithms have been considered in the literature. The work in [21] proposes that the systems only communicates when there is an overvoltage in the system. Whereas the method in [22] proposes that buses communicate their physical state only whent he difference between their physical state and previously communicated state is too large. Unlike our work, these papers do not consider bandwidth limited communication.

In this paper, we study a distributed voltage control where only a few bits of communication between neighboring buses are needed during each control step. In particular, the voltage control device on each bus determines the reactive power injection based on its local voltage measurement and current reactive power injection, in addition to a few bits of information communicated from its physical neighbors. We show that the algorithm can regulate the voltages to an acceptable range, for any predefined accuracy, in a finite number of iterations. We also provide an upper bound on the number of communicated bits (in the worst case) that are needed to ensure a predefined accuracy of the desired voltage level. Though the theoretical analysis is based on a linearized power flow model that is applicable for radial distribution network developed in [23], we use the nonlinear power flow

model to numerically test the algorithm. Moreover, we test the performance under both static and dynamic operating conditions. Both the theoretical and numerical results confirm that reducing the communication to a few bits do not sacrifice the control performance. This opens up a large flexibility for implementing voltage control in practice. For instance it enables the use of power lines as the communication media for electricity distribution systems despite the limited bandwidth of power line communication [12].

Preliminary studies of this work appeared in [24]. However, in [24] most of the proofs are omitted (but they appear here) and all the numerical results especially with the nonlinear power flow model here are new. Moreover, the presentation and theoretical results here have been largely improved compared to [24].

A. Notation

Vectors and matrices are represented by boldface lower and upper case letters, respectively. The imaginary unit is denoted by $\mathbf{i}$, i.e., $\mathbf{i} = \sqrt{-1}$. The set of real, complex, and natural numbers are denoted by $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{N}$, respectively. The set of real n vectors and $n \times m$ matrices are denoted by $\mathbb{R}^n$ and $\mathbb{R}^{n \times m}$, respectively. Otherwise, we use calligraphy letters to represent represent sets. We let $\mathcal{S}^{n-1} = \{\mathbf{x} \in \mathbb{R}^n | 1 = ||\mathbf{x}||\}$ denote the unit sphere. The superscript $(\cdot)^{\mathrm{T}}$ stands for transpose. $\mathrm{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_n)$ denotes the diagonal block matrix with $\mathbf{A}_1, \ldots, \mathbf{A}_n$ on the diagonal. We let $|| \cdot ||$ denote the 2-norm. The distance between a point $\mathbf{x} \in \mathbb{R}^n$ and a set $\mathcal{X} \subseteq \mathbb{R}^n$ is given by $\mathrm{dist}(\mathbf{x}, \mathcal{X}) = \inf_{\mathbf{z} \in \mathbb{R}^n} ||\mathbf{x} - \mathbf{z}||$. For a matrix $\mathbf{A} \in \mathbb{R}^{n \times N}$, we let $\lambda_i(\mathbf{A})$, for $i = 1, \ldots, n$ denote its eigenvalues (not in any particular order) and $\lambda_{\max}(\mathbf{A}) = \max_{i=1,\ldots,n} |\lambda_i(\mathbf{A})|$. We let $\lceil \mathbf{x} \rceil_+$ denote the projection of $\mathbf{x} \in \mathbb{R}^n$ into the positive orthant $\mathbb{R}^n_+$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section we present the power flow model, Section II-A, and the voltage regulation problem, Section II-B.

A. System Model: Linearized Power Distribution Network

Consider a radial power distribution network with $N + 1$ buses represented by the set $\mathcal{N}_0 = \{0\} \cup \mathcal{N}$, where $\mathcal{N} = \{1, \ldots, N\}$. Bus 0 is a feeder bus and the buses in $\mathcal{N}$ are branch buses. Let $\mathcal{E} \subseteq \mathcal{N}_0 \times \mathcal{N}_0$ denote the set of directed flow lines, so if $(i, j) \in \mathcal{E}$ then i is the parent of j. For each i, let $s_i = p_i + \mathbf{i}q_i \in \mathbb{C}$, $V_i \in \mathbb{C}$, and $v_i \in \mathbb{R}_+$ denote the complex power injection, complex voltage, and squared voltage magnitude, respectively, at bus i. For each $(i, j) \in \mathcal{E}$, let $S_{ij} = P_{ij} + \mathbf{i}Q_{ij} \in \mathbb{C}$ and $z_{ij} = r_{ij} + \mathbf{i}x_{ij} \in \mathbb{C}$ denote the complex power flow and impedance in the line from bus i to bus j. To model the relationship between the variables, we use the linearized branch flow model from [23], which gives a good approximation in radial distribution networks.[1]

$$-p_i = P_{\sigma_i i} - \sum_{k:(i,k) \in \mathcal{E}} P_{ik}, \quad i \in \mathcal{N}, \tag{1a}$$

$$-q_i = Q_{\sigma_i i} - \sum_{k:(i,k) \in \mathcal{E}} Q_{ik}, \quad i \in \mathcal{N}, \tag{1b}$$

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

where σ_i is the parent of bus $i \in \mathcal{N}$, i.e., the unique $\sigma_i \in \mathcal{N}_0$ with $(\sigma_i, i) \in \mathcal{E}$. By rearranging Equation (1) we get that

$$\mathbf{v} = \mathbf{A}\mathbf{q}^{\mathrm{I}} + \mathbf{B}\mathbf{p}^{\mathrm{I}} + \mathbf{1}v_0, \tag{2}$$

where $\mathbf{v} = [v_1, \ldots, v_N]^{\mathrm{T}}$, $\mathbf{q}^{\mathrm{I}} = [q_1, \ldots, q_N]^{\mathrm{T}}$, $\mathbf{p}^{\mathrm{I}} = [p_1, \ldots, p_N]^{\mathrm{T}}$,

$$\mathbf{A}_{ij} = 2 \sum_{(h,k) \in \mathcal{P}_i \cap \mathcal{P}_j} x_{hk}, \quad \text{and} \quad \mathbf{B}_{ij} = 2 \sum_{(h,k) \in \mathcal{P}_i \cap \mathcal{P}_j} r_{hk},$$

where $\mathcal{P}_i \subseteq \mathcal{E}$ is the set of edges in the path from bus 0 to bus i. We use the following result in the algorithm development.

Proposition 1. $\mathbf{A}$ *is a positive definite matrix whose inverse has the following structure*

$$a_{ij} := [\mathbf{A}^{-1}]_{ij} = \begin{cases} \dfrac{1}{2}\left(x_{\sigma_i i}^{-1} + \displaystyle\sum_{k:(i,k) \in \mathcal{E}} x_{ik}^{-1}\right) & \text{if } i{=}j, \\[2ex] -\dfrac{1}{2}x_{ij}^{-1} & \text{if } (i,j){\in}\mathcal{E} \\ & \text{or } (j,i){\in}\mathcal{E}, \\[1ex] 0 & \text{otherwise.} \end{cases} \tag{3}$$

Proof: It is proved in [2, Lemma 1] that $\mathbf{A}$ is positive definite. Direct calculations show that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. ∎

We now introduce the Voltage Regulation Problem.

B. Voltage Regulation Problem

Suppose that the real power injection $\mathbf{p}^{\mathrm{I}}$ at each bus has been decided. Write the reactive power injection as $\mathbf{q}^{\mathrm{I}} = \mathbf{q} + \mathbf{q}^{\mathrm{U}}$, where $\mathbf{q}$ is the adjustable reactive power that can be used for voltage regulation and $\mathbf{q}^{\mathrm{U}}$ denotes other reactive power injection that cannot be changed by the voltage control devices. Then the goal of the voltage regulation problem is to find feasible voltages $\mathbf{v}$ and adjustable reactive powers $\mathbf{q}$ so that the physical relationship (2) holds and that $\mathbf{v}$ and $\mathbf{q}$ are inside some feasible operation range $[\mathbf{v}^{\min}, \mathbf{v}^{\max}]$ and $[\mathbf{q}^{\min}, \mathbf{q}^{\max}]$. Formally, the voltage regulation problem is to find the reactive power injection $\mathbf{q}$ so that,

$$\mathbf{v}(\mathbf{q}) = \mathbf{A}\mathbf{q} + \mathbf{d}, \tag{4a}$$

$$\mathbf{v}^{\min} \leq \mathbf{v}(\mathbf{q}) \leq \mathbf{v}^{\max} \tag{4b}$$

$$\mathbf{q}^{\min} \leq \mathbf{q} \leq \mathbf{q}^{\max} \tag{4c}$$

---

[1] The results also directly apply to the linearized power flow model in [10]. The linearized model is for the purpose of deriving the control algorithm and analyzing the theoretical performance. The designed control algorithms can be applied to the nonlinear distribution networks. In Section V, we use the nonlinear power flow model to numerically test the algorithm performance.

where $\mathbf{d} := \mathbf{A}\mathbf{q}^{\mathrm{U}} + \mathbf{B}\mathbf{p}^{\mathrm{I}} + \mathbf{1}v_0$. Throughout the paper we implicitly assume that the voltage regulation problem has a solution.

Assumption 1. *There exists* $\mathbf{q} \in \mathbb{R}^N$ *satisfying Equation* (4).

Our problem of voltage control is to design a distributed control for finding the feasible reactive power injections and voltages that satisfy equation (4) that works under any operating condition denoted by $\mathbf{d}$. In particular, where each bus controls its reactive power based only *on local information and a limited number of bits communicated from neighboring buses*. That is, each bus $i \in \mathcal{N}$ updates its reactive power injection according to

$$\mathbf{q}_i(t+1) = K_i(\texttt{Local\_Info}_i(t), \texttt{Comm}_i(t)), \quad (5)$$

where t is the iteration index and $K_i(\cdot)$ is the control law at bus i. The function $K_i(\cdot)$ depends on the local information

$$\texttt{Local\_Info}_i(t) = (\mathbf{q}_i(0), \ldots, \mathbf{q}_i(t), \mathbf{v}_i(0), \ldots, \mathbf{v}_i(t)),$$

and the communicated information that bus i has received from its neighbors

$$\texttt{Comm}_i(t) = ((\mathbf{b}_j(0))_{j \in \mathcal{N}_i}, \ldots, (\mathbf{b}_j(t))_{j \in \mathcal{N}_i}),$$

where $\mathcal{N}_i = \{j \in \mathcal{N} | (i,j) \in \mathcal{E} \text{ or } (i,j) \in \mathcal{E}\}$ and $\mathbf{b}_j(t)$ is the information that bus j communicates to its neighbors at iteration t. We study how such control laws can be achieved when $\mathbf{b}_i(t)$ contains a limited number of bits.

Note that the problem can generally not be solved without communication, i.e., with $\texttt{Comm}_i(t)$ empty for all i and t, as proved by the impossibility result in [10]. Some authors have proposed distributed solution algorithms where only neighbors in the power networks communicate, see [5], [6], [7], [8]. This is challenging in practice because the communication infrastructure for distribution systems is less reliable and less ubiquitous as compared to that for the bulk transmission system. To compensate for that, we study limited communication voltage control algorithms where controllers only communicate a few bits to their neighbors. These algorithms are based on combining the ideas from the voltage control algorithms in [7], [8] and the limited communication gradient methods in [19]. However, the algorithms and results in [19] are not directly applicable here because the communication structure is different and some of the assumptions in [19] do not hold for the voltage regulation problem.

III. VOLTAGE CONTROL WITH LIMITED COMMUNICATION

We present our limited communication algorithm for the voltage control problem in Section III-A. We highlight the convergence properties of the algorithm in Section III-B.

A. Algorithm

The following algorithm is an instance of the control algorithm in Equation (5) where each bus communicates only 2-bits per iteration.

VC-LB Algorithm: Voltage Control - Limited Bandwidth

a) **Initialization:** Set $t = 0$, and for each bus $i \in \mathcal{N}$ set

$$\boldsymbol{\lambda}_i(0) = (\boldsymbol{\lambda}_i^{\min}(0), \boldsymbol{\lambda}_i^{\max}(0)) = (0, 0)$$
$$\boldsymbol{\mu}_i(0) = (\boldsymbol{\mu}_i^{\min}(0), \boldsymbol{\mu}_i^{\max}(0)) = (0, 0).$$

where the variables $\boldsymbol{\lambda}_i^{\min}(t)$, $\boldsymbol{\lambda}_i^{\max}(t)$, $\boldsymbol{\mu}_i^{\min}(t)$, and $\boldsymbol{\mu}_i^{\max}(t)$ are dual variables related to violating the bounds in equations (4c) and (4b), see Section IV for details.

b) **Local Computation:** Each bus $i \in \mathcal{N}$ computes its next reactive power injection as follows

$$\mathbf{q}_i(t+1) = \boldsymbol{\lambda}_i^{\min}(t) - \boldsymbol{\lambda}_i^{\max}(t) + \sum_{j \in \mathcal{N}_i} a_{ij}(\boldsymbol{\mu}_j^{\min}(t) - \boldsymbol{\mu}_j^{\max}(t)). \quad (6)$$

Bus i can then also compute the communicated signal

$$\mathbf{b}_i(t+1) = \texttt{sign} \left[\begin{array}{c} \mathbf{q}_i(t+1) - \mathbf{q}_i^{\max} \\ \mathbf{q}_i^{\min} - \mathbf{q}_i(t+1) \end{array} \right]. \quad (7)$$

c) **Local Control:** Each bus $i \in \mathcal{N}$ injects the reactive power $\mathbf{q}_i(t+1)$ into the power network.

d) **Local Measurement:** Each bus $i \in \mathcal{N}$ measures the voltage magnitude $\mathbf{v}_i(\mathbf{q}(t+1))$, given by the physical relationship (4a).

e) **Communication:** Each bus $i \in \mathcal{N}$ communicates $\mathbf{b}_i(t+1)$ to each of its neighbours $j \in \mathcal{N}_i$ using a two bits.

f) **Local Computation:** Each bus $i \in \mathcal{N}$ updates its dual variables

$$\boldsymbol{\lambda}_i(t+1) = \left\lceil \boldsymbol{\lambda}_i(t) + \alpha \left[\begin{array}{c} \mathbf{v}_i(\mathbf{q}(t+1)) - \mathbf{v}_i^{\max} \\ \mathbf{v}_i^{\min} - \mathbf{v}_i(\mathbf{q}(t+1)) \end{array} \right] \right\rceil_+, \quad (8)$$

$$\boldsymbol{\mu}_i(t+1) = \lceil \boldsymbol{\mu}_i(t) + \beta \mathbf{b}_i(t) \rceil_+. \quad (9)$$

Bus i also updates a local copy of $\boldsymbol{\mu}_j(t+1)$ for each neighbor $j \in \mathcal{N}_i$ using Equation (9).

g) **Update Iteration Index:** $t = t + 1$ and go to step **b)**.

This algorithm satisfies the structure of Equation (5). This can be seen by noting that $\boldsymbol{\lambda}_i(t)$ is a function of $\texttt{Local\_Info}_i(t)$ via Equation (8) and $\boldsymbol{\mu}_i(t)$ is a function of $\texttt{Local\_Info}_i(t)$ and $\texttt{Comm}_i(t)$ via Equations (7) and (9). The algorithm is easy to implement since the **Local Computation** step is based on few elementary operation. Moreover, for the **Communication** step each bus only needs to communicate 2 bits of information since $\mathbf{b}_i(t)$ can only take one of the $2^2 = 4$ values $(-1, -1)$, $(1, 1)$, $(1, -1)$, and $(-1, 1)$. The parameters $\alpha, \beta > 0$ are step-sizes and are discussed further in the next section.

Remark 1. *We have considered the extreme case when only 2-bits are communicated per iterations to make the algorithm analysis more manageable. It is possible that, using more bits, can improve performance and studying such trade-offs would be an interesting extension to this work. However, from both the theoretical results and numerical simulations, 2-bits communication voltage control can achieve a comparable performance as the non-quantized counterpart for our problem.*

B. Main Convergence Results

We now study the converge of the **VC-LB Algorithm** to a solution to the voltage regulation problem. We show that in finite number of iterations the reactive power $\mathbf{q}(t)$ and voltage $\mathbf{v}(\mathbf{q}(t))$ satisfy Equation (4) approximately and exactly under mild additional assumptions. We measure the feasibility of the reactive power $\mathbf{q} \in \mathbb{R}^N$ (and the associated voltage $\mathbf{v}(\mathbf{q}) \in \mathbb{R}^N$) as follows

$$\texttt{fes}(\mathbf{q}) = \texttt{dist}\big((\mathbf{q}, \mathbf{v}(\mathbf{q})), \mathcal{Q} \times \mathcal{V}\big), \tag{10}$$

where $\mathcal{Q} = [\mathbf{q}^{\min}, \mathbf{q}^{\max}]$ and $\mathcal{V} = [\mathbf{v}^{\min}, \mathbf{v}^{\max}]$.

The following theorem (proved in Section IV-B) establishes that the **VC-LB Algorithm** can solve the problem up to any $\epsilon > 0$ accuracy in finite number of iterations if the step-sizes are chosen appropriately.

Theorem 1. *(Approximate Solution) Let $\epsilon \in (0, 1]$ and choose the step-sizes $\alpha, \beta > 0$ such that[2]*

$$\alpha = \frac{1}{L} \quad and \quad \beta = \frac{\epsilon}{4N^{3/2}L} \tag{11}$$

where

$$L = \max_{i=1,\ldots,N} 2\left(\lambda_i(\mathbf{A}) + \frac{1}{\lambda_i(\mathbf{A})}\right).$$

Then there exists $T \in \mathbb{N}$ such that $\texttt{fes}(\mathbf{q}(T)) \leq \epsilon$ where T is upper bounded by

$$T \leq \left\lceil \frac{16N^3 LQ\lambda_{\max}(\mathbf{A})}{\epsilon^2} \right\rceil, \tag{12}$$

where $Q = \max\left\{(\mathbf{q}_i^{\min})^2, (\mathbf{q}_i^{\max})^2 \; : i = 1, \ldots, N\right\}$.

Theorem 1 shows that the **VC-LB Algorithm** can solve the voltage regulation problem to any precision in finite number of iterations and using finite number of communicated bits. In particular, the theorem shows that any $\epsilon > 0$ accuracy can be reached by communicating $\mathcal{O}\left(1/\epsilon^2\right)$ bits. This is comparable to convergence rate without quantization. This follows from the fact that the **VC-LB Algorithm** is a quantized version of a dual gradient method where the primal variable is $\mathbf{q}$ and the dual gradient is Lipschitz continuous, See Section IV. Under these assumptions the feasibility of the primal iterates $\mathbf{q}(t)$ of dual gradient methods (with constant step-size) is $\mathcal{O}(1/\epsilon^2)$, see, for example [25]. We note that to compute L requires the knowledge of $\mathbf{A}$. However, since $\mathbf{A}$ is generally not changing L needs to be computed once and can then be used every time when the algorithm runs.

In practice, we often like to run the algorithm over an extended period of time. Therefore, it is desirable that once the algorithm converges to some $\epsilon > 0$ accuracy then it does not oscillate much away from that accuracy. We show this numerically in Section V. However, it is hard to prove this analytically due to technical difficulties explained in Section IV-C. Nevertheless, the following theorem (proved in

Section IV-C) shows that there exist step-sizes that ensure that any $\epsilon > 0$ accuracy is held for all sufficiently large t.

Theorem 2. *Suppose that the voltage regulation problem is strictly feasible, i.e., there exists some $\mathbf{q} \in \mathbb{R}^N$ such that[3]*

$$\mathbf{q} \in (\mathbf{q}^{\min}, \mathbf{q}^{\max}) \quad and \quad \mathbf{v}(\mathbf{q}) \in (\mathbf{v}^{\min}, \mathbf{v}^{\max}).$$

Then for any $\epsilon > 0$ there exist step-sizes $\alpha, \beta > 0$ and $T \in \mathbb{N}$ such that

$$\texttt{fes}(\mathbf{q}(t)) \leq \epsilon, \quad for \; all \;\; t \geq T. \tag{13}$$

We can also use the **VC-LB Algorithm** to solve the voltage regulation problem exactly provided a mild additional assumption. In particular, if we use the **VC-LB Algorithm** to solve the restricted voltage regulation problem of finding reactive power $\mathbf{q}$ that solves

$$\mathbf{q} \in [\mathbf{q}^{\min} + \mathbf{1}\rho, \mathbf{q}^{\max} - \mathbf{1}\rho] \tag{14a}$$
$$\mathbf{v}(\mathbf{q}) \in [\mathbf{v}^{\min} + \mathbf{1}\rho, \mathbf{v}^{\max} - \mathbf{1}\rho], \tag{14b}$$

for some $\rho > 0$. Then any $\epsilon = \rho$ accurate solution to the voltage regulation problem in Equation (14) is an exact solution to the original problem. Hence, we can apply Theorem 1 to find an exact solution in finite number of iterations where each bus communicates finite number of bits.

Corollary 1 (Exact Solution). *Let $\rho > 0$ be given and consider the **VC-LB Algorithm** for solving the restricted the voltage regulation problem in Equation (14).[4] Moreover, suppose restricted the voltage regulation problem in Equation (14) is feasible. Choose the step-sizes $\alpha, \beta > 0$ such that*

$$\alpha = \frac{1}{L} \quad and \quad \beta = \frac{\rho}{4N^{3/2}L}.$$

Then there exists $T \in \mathbb{N}$ such that $\texttt{fes}(\mathbf{q}(T)) \leq 0$ where T is upper bounded by

$$T \leq \left\lceil \frac{16N^3 LQ\lambda_{\max}(\mathbf{A})}{\rho^2} \right\rceil. \tag{15}$$

The parameter ρ should be chosen so that the intervals in Equation (14) are not empty. Also note that increasing ρ restricts the set of feasible voltages $\mathbf{v}(\mathbf{q})) \in [\mathbf{v}^{\min} + \mathbf{1}\rho, \mathbf{v}^{\max} - \mathbf{1}\rho]$, which must be non empty for the theorem to hold.

C. Implementation without Violating Capacity Constraint

As shown in Section III-A and III-B, while the **VC-LB** algorithm can guarantee that the reactive power $\mathbf{q}(t)$ will reach the capacity constraint $[\mathbf{q}^{\min}, \mathbf{q}^{\max}]$ (approximately), $\mathbf{q}(t)$ may still violate these limits during the transient. This can be problematic since many physical control devices do not allow violating the capacity constraint for an extended amount of time. One way to reduce this violation is to use a smaller step size α. By doing this, the voltage constraint multiplier $\boldsymbol{\lambda}(t)$ will update slower compared to the capacity constraint

---

[2] Similar convergence results can be ensured for any step-size $\alpha, \beta > 0$ giving in Proposition (2)in Section IV-B, in which case the upper bound in Equation (12) is replaced by (28). However, we present the results here using these particular step-sizes to get an explicit form on the upper bound in (12) in terms of available parameters.

[3] Note that this is a stronger requirement that given by Assumption 1
[4] The change in the **VC-LB Algorithm** is that in Equation (7) we replace $\mathbf{q}_i^{\max}$ and $\mathbf{q}_i^{\min}$ by $\mathbf{q}_i^{\max} - \rho$ and $\mathbf{q}_i^{\min} + \rho$, respectively, and in Equation (8) we replace $\mathbf{v}_i^{\max}$ and $\mathbf{v}_i^{\min}$ by $\mathbf{v}_i^{\max} - \rho$ and $\mathbf{v}_i^{\min} + \rho$, respectively.

multiplier $\boldsymbol{\mu}(t)$, hence putting more priority on enforcing the capacity constraint. We illustrate this in simulations, see Section V.

We now propose an alternative implementation of the **VC-LB** algorithm that meets the reactive capacity constraint during the transients inspired by [26], [27]. At time t, instead of implementing $\mathbf{q}_i(t)$ that may violate the capacity constraint, we implement $\mathbf{q}_i^{phy}(t)$, defined to be the projection of $\mathbf{q}_i(t)$ onto the capacity constraint set $[\mathbf{q}_i^{\min}, \mathbf{q}_i^{\max}]$. Mathematically, we implement $\mathbf{q}_i^{phy}(t) = \max(\min(\mathbf{q}_i(t), \mathbf{q}_i^{\max}), \mathbf{q}_i^{\min})$. Here the superscript in variable $\mathbf{q}_i^{phy}(t)$ refers to "physical", which is to emphasize that $\mathbf{q}_i^{phy}(t)$ is the amount of reactive power this is *physically* implemented. Clearly, $\mathbf{q}_i^{phy}(t)$ will always satisfy the capacity constraint. We formally introduce the revised VC-LB below.

VC-LB with Projection (VC-LB-P)

a) **Initialization:** The same as **VC-LB** Algorithm.
b) **Local Computation:** The same as **VC-LB** Algorithm.
c) **Local Control:** Each bus $i \in \mathcal{N}$ injects the reactive power $\mathbf{q}_i^{phy}(t+1) = \max(\min(\mathbf{q}_i(t+1), \mathbf{q}_i^{\max}), \mathbf{q}_i^{\min})$ into the power network.
d) **Local Measurement:** Each bus $i \in \mathcal{N}$ measures the voltage magnitude $\mathbf{v}_i(\mathbf{q}^{phy}(t+1))$, given by the physical relationship (4a).
e) **Communication:** The same as VC-LB Algorithm.
f) **Local Computation:** Each bus $i \in \mathcal{N}$ updates its dual variables

$$\boldsymbol{\lambda}_i(t+1) = \left\lceil \boldsymbol{\lambda}_i(t) + \alpha \left[\begin{array}{c} \mathbf{v}_i(\mathbf{q}^{phy}(t+1)) - \mathbf{v}_i^{\max} \\ \mathbf{v}_i^{\min} - \mathbf{v}_i(\mathbf{q}^{phy}(t+1)) \end{array} \right] \right\rfloor_+, \tag{16}$$

$$\boldsymbol{\mu}_i(t+1) = \lceil \boldsymbol{\mu}_i(t) + \beta \mathbf{b}_i(t) \rceil_+. \tag{17}$$

Bus i also updates a local copy of $\boldsymbol{\mu}_j(t+1)$ for each neighbor $j \in \mathcal{N}_i$ using Equation (9).
g) **Update Iteration Index:** $t = t + 1$ and go to step **b)**.

The difference between **VC-LB-P** with the original **VC-LB** lies in step c), d), f). In step c), we implement $\mathbf{q}_i^{phy}(t+1)$ instead of $\mathbf{q}_i(t+1)$. As a result, in step d) the measured voltage becomes $\mathbf{v}_i(\mathbf{q}^{phy}(t+1))$ instead of $\mathbf{v}_i(\mathbf{q}(t+1))$, and correspondingly in step f), update equation (16) uses $\mathbf{v}_i(\mathbf{q}^{phy}(t+1))$ instead of $\mathbf{v}_i(\mathbf{q}(t+1))$.

We now explain the rationale behind the new implementation **VC-LB-P**. Note that in **VC-LB-P**, the update equation for Lagrangian multiplier $\boldsymbol{\mu}_i(t)$ and variable $\mathbf{q}_i(t)$ is kept the same as **VC-LB**. This means that $\boldsymbol{\mu}_i(t)$ still reflects the capacity constraint violation of $\mathbf{q}_i(t)$, so $\mathbf{q}_i(t)$ will still meet the capacity constraint asymptotically, and hence $\mathbf{q}_i(t) - \mathbf{q}_i^{phy}(t) \to 0$, i.e., the projection step will not substantially change the convergence results established in Section III-B. We note that similar projection steps have been done in [26], [27] and [26] has theoretically shown that under the projection step, a modified primal-dual algorithm can still converge. Though we could not rigorously prove the convergence of VC-LB-P in this paper, we will conduct extensive simulation to verify the convergence of VC-LB-P.

IV. CONVERGENCE ANALYSIS

The goal of this section is to prove the convergence results from the previous section. We give preliminary results in Section IV-A and then prove Theorems 1 and 2, respectively, in Sections IV-B and IV-C. Readers that not interested in these proofs can go straight to Section V with no loss of information.

A. Preliminaries: Duality Theory

We prove that the algorithm converges to an approximate solution to the following optimization problem [5]

$$\begin{aligned} \underset{\mathbf{q}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{q}^{\mathrm{T}}\mathbf{A}\mathbf{q} \\ \text{subject to} \quad & \mathbf{v}^{\min} \leq \mathbf{v}(\mathbf{q}) \leq \mathbf{v}^{\max} \\ & \mathbf{q}^{\min} \leq \mathbf{q} \leq \mathbf{q}^{\max}. \end{aligned} \tag{18}$$

Problem (18) is convex because of Proposition 1. Its dual is

$$\begin{aligned} \underset{\mathbf{z}=(\boldsymbol{\lambda},\boldsymbol{\mu})}{\text{maximize}} \quad & D(\mathbf{z}) \\ \text{subject to} \quad & \mathbf{z} \geq 0, \end{aligned} \tag{19}$$

where $\boldsymbol{\lambda} = (\boldsymbol{\lambda}^{\min}, \boldsymbol{\lambda}^{\max})$ and $\boldsymbol{\mu} = (\boldsymbol{\mu}^{\min}, \boldsymbol{\mu}^{\max})$ are the dual variables associated to the constraints $\mathbf{v}^{\min} \leq \mathbf{v}(\mathbf{q}) \leq \mathbf{v}^{\max}$ and $\mathbf{q}^{\min} \leq \mathbf{q} \leq \mathbf{q}^{\max}$, respectively, and $D : \mathbb{R}^{4N} \to \mathbb{R}$ is the dual function, see Chapter 5 in [28] for the details. The dual gradient is

$$\nabla D(\mathbf{z}) = \left[\begin{array}{c} \nabla^{\boldsymbol{\lambda}} D(\mathbf{z}) \\ \nabla^{\boldsymbol{\mu}} D(\mathbf{z}) \end{array} \right] \tag{20}$$

where

$$\nabla^{\boldsymbol{\lambda}} D(\mathbf{z}) = \left[\begin{array}{c} \mathbf{v}^{\min} - \mathbf{v}(\mathbf{q}(\mathbf{z})) \\ \mathbf{v}(\mathbf{q}(\mathbf{z})) - \mathbf{v}^{\max} \end{array} \right], \quad \nabla^{\boldsymbol{\mu}} D(\mathbf{z}) = \left[\begin{array}{c} \mathbf{q}^{\min} - \mathbf{q}(\mathbf{z}) \\ \mathbf{q}(\mathbf{z}) - \mathbf{q}^{\max} \end{array} \right],$$

and

$$\mathbf{q}(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \boldsymbol{\lambda}^{\min} - \boldsymbol{\lambda}^{\max} + \mathbf{A}^{-1}\boldsymbol{\mu}^{\min} - \mathbf{A}^{-1}\boldsymbol{\mu}^{\max}. \tag{21}$$

In Lemma 1 we show that $\nabla D(\cdot)$ is L-Lipschitz continuous, with L given in the lemma. Therefore, the gradient decent method

$$\boldsymbol{\lambda}(t+1) = \lceil \boldsymbol{\lambda}(t) + \gamma \nabla^{\boldsymbol{\lambda}} D(\boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \rceil_+ \tag{22a}$$

$$\boldsymbol{\mu}(t+1) = \lceil \boldsymbol{\mu}(t) + \gamma \nabla^{\boldsymbol{\mu}} D(\boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \rceil_+ \tag{22b}$$

converges to the set of optimal dual variables for appropriate step-size γ [29, Chapter 2]. The work in [7], [8], [10] shows how the dual gradient iterations in Equations (22a) and (22b) can be implemented in a distributed manner among the buses so that only neighbouring buses need to communicate per iteration. However, these algorithms communicate real numbers, which is challenging in practice as communication among controllers is generally constrained to low data rates. To compensate for that, we have presented the **VC-LB Algorithm** in Section III, which can be equivalently be written as follows:

$$\boldsymbol{\lambda}(t+1) = \lceil \boldsymbol{\lambda}(t) + \alpha \nabla^{\boldsymbol{\lambda}} D(\boldsymbol{\lambda}(t)) \rceil_+, \tag{23a}$$

$$\boldsymbol{\mu}(t+1) = \lceil \boldsymbol{\mu}(t) + \beta \,\texttt{sign}(\nabla^{\boldsymbol{\mu}} D(\boldsymbol{\mu}(t))) \rceil_+, \tag{23b}$$

---

[5]The cost can be interpenetrated as a network loss, see [7], [8].

where $\alpha, \beta > 0$ are step-sizes and the primal variables are updated according to (21). We now prove Theorems 1 and 2 by considering the **VC-LB Algorithm** on the form in Equation (23).

B. Proof of Theorem 1

We now prove Theorem 1. The main step of the proof is illustrated in the following result (proved in Appendix B).[6]

Proposition 2. *Suppose $\epsilon > 0$ and $\mathbf{z} = (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^{4N}$ are such that $V(\mathbf{z}) > \epsilon$ where*[7]

$$V(\mathbf{z}) = ||\mathbf{z} - \lceil \mathbf{z} + \nabla D(\mathbf{z})\rceil_+||. \tag{24}$$

Choose the step-sizes $\alpha, \beta > 0$

$$\alpha < \frac{2}{L} \tag{25a}$$

$$\beta < \min\left\{\frac{\epsilon}{2N^{3/2}L}, \sqrt{\frac{\alpha(1 - L\alpha/2)\epsilon^2}{2NL}}\right\}. \tag{25b}$$

Then for

$$\bar{\mathbf{z}} = (\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) = \left[\begin{array}{cc}\boldsymbol{\lambda} & + & \alpha\nabla^{\boldsymbol{\lambda}}D(\boldsymbol{\lambda}) \\ \boldsymbol{\mu} & & \beta\,sign(\nabla^{\boldsymbol{\mu}}D(\boldsymbol{\mu}))\end{array}\right]$$

following holds

$$D(\lceil\bar{\mathbf{z}}\rceil^+) \geq D(\mathbf{z}) + \delta(\alpha, \beta) \tag{26}$$

where

$$\delta(\alpha, \beta) = \min\left\{\left(\alpha - \frac{L}{2}\alpha^2\right)\frac{\epsilon^2}{2} - NL\beta^2, \frac{\epsilon\beta LN}{2N^{3/2}L} - \beta^2 LN\right\}$$
$$> 0 \tag{27}$$

where L is an Lipschitz constant on $\nabla D(\cdot)$.

From Lemma 2 in Appendix A we have that $\mathrm{fes}(\mathbf{q}(\mathbf{z})) \leq V(\mathbf{z})$. Hence, Proposition 2 shows that for all $\epsilon > 0$, we can choose step-sizes $\alpha, \beta > 0$ so that if $\mathrm{fes}(\mathbf{q}(\mathbf{z})) > \epsilon$ then the dual objective function value is improved by taking a step of the algorithm in Equations (23a) and (23b). Using this intuition, we have the following result.

Proposition 3. *Consider the VC-LB Algorithm and take some $\epsilon > 0$. Choose the step-sizes $\alpha, \beta > 0$ as in Equations (25a) and (25b). Then there exists $T \in \mathbb{N}$ such that $\mathrm{fes}(\mathbf{q}(T)) \leq V(\mathbf{z}) \leq \epsilon$ where T is upper bounded by*

$$T \leq \left\lceil \frac{D^\star - D(\mathbf{z}(0))}{\delta(\alpha, \beta)} \right\rceil. \tag{28}$$

Proof: Suppose that $V(\mathbf{z}(t)) > \epsilon$, for $t = 0, \ldots, T_0 - 1$, where $T_0 := \lceil (D^\star - D(\mathbf{z}(0)))/\delta(\alpha, \beta)\rceil$. Then by Proposition 2 we have that $D(\mathbf{z}(t)) \geq D(\mathbf{z}(t-1)) + \delta(\alpha, \beta)$ for $t = 0, \ldots, T_0$ or

$$0 \leq D^\star - D(\mathbf{z}(T_0)) \leq D^\star - D(T_0 - 1) - \delta(\alpha, \beta)$$
$$\leq D^\star - D(0) - T_0\delta(\alpha, \beta) \leq 0,$$

---

[6] Note that Proposition 2 is similar to Lemma 4 in [19]. However, in Lemma 4 in [19] the gradient is assumed to be bounded. Moreover, unlike in [19] the dual algorithm in this paper is an hybrid between the non-quantizes gradient step in Equation (23a) and the quantized gradient step in (23b). Therefore, the results in [19] do not apply here.

[7] Note that $V(\mathbf{z}) = 0$ is equivalent to the primal/dual variables $(\mathbf{q}(\mathbf{z}), \mathbf{z})$ satisfying the KKT conditions for optimality, see Lemma 2 in [19].

where the last inequality comes by that $T_0 \geq (D^\star - D(\mathbf{z}(0)))/\delta(\alpha, \beta)$. Hence, $\mathbf{z}(T_0)$ is an optimal solution to the dual problem in Equation (19) and $\mathbf{q}(\mathbf{z}(T_0))$ is an optimal solution to the primal problem in Equation 18 implying $V(\mathbf{z}(T_0)) = 0 \leq \epsilon$ from Lemma 2. ∎

Theorem 1 follows from Proposition 3. Direct calculation shows that the step-sizes in Equation (11) satisfy the condition given in Equations (25a) and (25b). Therefore, we just have to show for α and β in Equation (11) following holds

$$\frac{D^\star - D(\mathbf{z}(0))}{\delta(\alpha, \beta)} \leq \frac{16N^3 LQ\lambda_{\max}(\mathbf{A})}{\epsilon^2} \tag{29}$$

to conclude the proof. Equation (29) follows directly from the following two identity (proved below):

$$\delta(\alpha, \beta) = \frac{3\epsilon^2}{16N^2 L}, \quad \text{and} \quad D^\star - D(\mathbf{z}(0)) \leq \lambda_{\max}(\mathbf{A})NQ,$$

for all $\hat{\mathbf{q}} \in [\mathbf{q}^{\min}, \mathbf{q}^{\max}]$ such that $\mathbf{v}(\hat{\mathbf{q}}) \in [\mathbf{v}^{\min}, \mathbf{v}^{\max}]$. The equality is obtained by plugging α and β from Equation (11) into Equation (27). To prove the inequality we note that $D^\star = (\mathbf{q}^\star)^{\mathrm{T}}\mathbf{A}\mathbf{q}^\star$ because of the strong duality and $D(\mathbf{z}(0)) = \min_{\mathbf{q} \in \mathbb{R}^N} \mathbf{q}^{\mathrm{T}}\mathbf{A}\mathbf{q} = 0$, since $\mathbf{z}(0) = \mathbf{0}$. Yielding

$$D^\star - D(\mathbf{z}(0)) \leq (\mathbf{q}^\star)^{\mathrm{T}}\mathbf{A}\mathbf{q}^\star \leq \lambda_{\max}(\mathbf{A})||\mathbf{q}^\star||^2 \leq \lambda_{\max}(\mathbf{A})NQ.$$

C. Proof of Theorem 2

We now prove Theorem 2. The result is based on the following result (due to limited space the proof is given in the extended version of this paper found in [30]).

Proposition 4. *Suppose that the voltage regulation problem is strictly feasible, i.e., there exists some $\mathbf{q} \in \mathbb{R}^N$ such that*

$$\mathbf{q} \in (\mathbf{q}^{\min}, \mathbf{q}^{\max}) \quad \text{and} \quad \mathbf{v}(\mathbf{q}) \in (\mathbf{v}^{\min}, \mathbf{v}^{\max}).$$

Define the set $\mathcal{Z}(\epsilon) = \{\mathbf{z} \in \mathbb{R}_+^{4N}|V(\mathbf{z}) \leq \epsilon\}$ and the function $F(\epsilon) = \inf\{D(\mathbf{z})|\mathbf{z} \in \mathcal{Z}(\epsilon)\}$. Then following holds:
a) There exists $\kappa > 0$ such that the set $\mathcal{Z}(\epsilon)$ is bounded and $F(\epsilon) > -\infty$ for all $\epsilon \leq \kappa$.
b) We have $\lim_{\epsilon \to 0^+} F(\epsilon) = f^\star$.
c) For any $\epsilon > 0$, step-sizes $\alpha, \beta > 0$ chosen as in Equations (25) and $T \in \mathbb{N}$ such that $V(\mathbf{z}(T)) \leq \epsilon$ we have

$$D(\mathbf{z}(t)) \geq F(\epsilon) - \frac{L}{2}\beta^2, \text{ for all } t \geq T. \tag{30}$$

From Proposition 3 and Proposition 4 we have that for any $\eta > 0$ there exit step-sizes $\alpha, \beta > 0$ and $T \in \mathbb{N}$ such that $D^\star - D(\mathbf{z}(t)) \leq \eta$ holds for all $t \geq T$. From Lemma 5 in Appendix A, there exists $\eta > 0$ such that if $D^\star - D(\mathbf{z}) \leq \eta$ then $\mathrm{dist}(\mathbf{z}, \mathcal{Z}^\star) \leq \epsilon/\phi$, where $\phi = ||\mathbf{G}||$ and

$$\mathbf{G} = \left(\begin{array}{cccc}\mathbf{I} & -\mathbf{I} & \mathbf{A}^{-1} & -\mathbf{A}^{-1} \\ \mathbf{A} & -\mathbf{A} & \mathbf{I} & -\mathbf{I}\end{array}\right).$$

Therefore, there exists $\alpha, \beta > 0$ and $T \in \mathbb{N}$ such that $\mathrm{dist}(\mathbf{z}(t), \mathcal{Z}^\star) \leq \epsilon/\phi$ for all $t \geq T$. We also have for $\mathbf{z} \in \mathbb{R}^{4N}$, by setting $\mathbf{z}^\star = \min_{\bar{\mathbf{z}} \in \mathcal{Z}^\star} ||\mathbf{z} - \bar{\mathbf{z}}||$, that

$$\mathrm{fes}(\mathbf{q}(\mathbf{z})) \leq ||(\mathbf{q}(\mathbf{z}), \mathbf{v}(\mathbf{q}(\mathbf{z}))) - (\mathbf{q}(\mathbf{z}^\star), \mathbf{v}(\mathbf{q}(\mathbf{z}^\star)))||,$$
$$= ||\mathbf{G}(\mathbf{z} - \mathbf{z}^\star)|| \leq \phi||\mathbf{z} - \mathbf{z}^\star|| = \phi\mathrm{dist}(\mathbf{z}, \mathcal{Z}^\star),$$

where we have used that $(\mathbf{q}(\mathbf{z}), \mathbf{v}(\mathbf{q}(\mathbf{z}))) = \mathbf{G}\mathbf{z} + (\mathbf{0}, \mathbf{d})$. Therefore, $\mathrm{fes}(\mathbf{q}(\mathbf{z}(t))) = \mathrm{dist}(\mathbf{z}(t), \mathcal{Z}^\star) \leq \epsilon$ for all $t \geq T$.

V. Numerical Results using Non Linear Power Flow

We test our algorithm on the 56-bus radial distribution network in [31], as shown in Figure 1. Bus 1 is the feeder bus, and there are PVs installed at bus 33, 40, 45, 55. All quantities, when units are not explicitly given, are in the per unit (p.u) system. The nominal value of voltage is 12kV. Throughout the simulation we set $v^{\min} = 11.4$kV, $v^{\max} = 12.6$kV ($\pm 5\%$ of the nominal value), and for each bus $i \in \mathcal{N}$, we set $\mathbf{q}_i^U = 0$, $\mathbf{q}_i^{\min} = -0.5$MW, and $\mathbf{q}_i^{\max} = 0.5$MW. The power flow is calculated using MATPOWER [32],[8] which uses the full nonlinear power flow instead of the linearized power flow in our analysis. We demonstrate the proposed algorithm in static and dynamic voltage control scenarios.

A. Static Voltage Control

We first test the case in which the real power injection $\mathbf{p}_i$ is fixed and is from the load data in [31], except we have scaled up the load from bus 7 to bus 19 to make the problem more challenging. The parameters of the algorithm is set as $\alpha = 0.2$, $\beta = 10^{-5}$ and $\rho = 0$. Fig. 2 illustrates the voltage profile and the reactive power injection trajectory under **VC-LB**. The results show that the voltage converge to the feasible range within roughly 400 iterations, or 800 bits of communication. Similarly, the reactive power converges to the feasible range within roughly 1200 iterations, or 2400 bits of communication.

Fig. 3 demonstrates the convergence of **VC-LB** under different step-sizes α, β and parameter ρ (recall ρ is introduced to find an exact solution, see Corollary 1). It plots the feasibility measure fes($\mathbf{q}(t)$) in Equation (10) as a function of iteration counts. First of all, Fig. 3 shows that when $\rho = 0$, the algorithm converges to an approximately feasible point, within certain accuracy. Further, there is a trade-off between how high the accuracy is and how many communication bits are needed to achieve that accuracy. Secondly, for the $\rho > 0$ case, Fig. 3 demonstrates our algorithm is indeed driving the system towards an exact solution, consistent with Corollary 1.

We comment that the step-sizes in Equation (11) are conservative. For the parameters of this network, we can calculate $L \approx 7 \times 10^3$, so the step-sizes in Equation (11) are very small ($\alpha \approx 10^{-4}$, $\beta \approx 6 \times 10^{-7}$ using $\epsilon = 1$), leading to very slow convergence, as illustrated in the blue curve in Fig. 3. On the contrary, the results in Fig. 2 and Fig. 3 have shown that in practice fast convergence can be obtained for step-sizes that are much larger than (11). This is because to obtain the theoretical guarantees we must account for every potential worst-case behavior and also make some relaxations to make the mathematical derivations tractable. This is a typical trade-off between theoretical convergence guarantees and the convergence in practice, as discussed for example in [28].

We note that in Fig. 2, though the reactive power injection meets the capacity constraint asymptotically, it violates the

---

[8]All simulations are run on MATLAB R2016b on Macbook Pro 2015 Model with 2.7 GHz Intel Core i5.

---

constraint during the transient. As discussed in Section III-C, we now test the two methods to deal with this. Firstly, we test the **VC-LB** algorithm with a small α, which as discussed in Section III-C, will put more priority on enforcing capacity constraint. We use the same simulation setting and step sizes as Fig. 2, except that α is reduced to 0.08 ($\beta = 10^{-5}$, $\rho = 0$ are kept same). Results are shown in Fig. 4. Compared to Fig. 2, the capacity constraint violation is now only minor in Fig. 4. Interestingly, we note that compared to Fig. 2, the voltage profile in Fig. 4 converges slower, which makes sense since in Fig. 4 we have put more priority on enforcing capacity constraint as opposed to the voltage constraint. Secondly, we test the variant of **VC-LB** algorithm, the **VC-LB-P** algorithm in Section III-C. Simulation settings are the same as Fig. 2 and we use the same step sizes $\alpha = 0.2, \beta = 10^{-5}, \rho = 0$. The voltage profile and the reactive power injection profile are presented in Fig. 5. Fig. 5 confirms that the **VC-LB-P** algorithm still converges and in the meanwhile the reactive power injection does not violate the capacity constraint at any time.

B. Dynamic Voltage Control

In practice voltage regulation algorithms must respond quickly to fluctuating electric behaviors of the consumers of the network. To this end, we test the **VC-LB-P** Algorithm in a dynamic environment where the real power injections change over the course of the algorithm. Specifically, we let the real power injections fluctuate (randomly) to a new set of values once in a while, and within each interval between two fluctuations, the algorithm **VC-LB-P Algorithm** is allowed to run 500 iterations (communicate 1000 bits) to respond to the fluctuation. The value of the fluctuating real power injection is determined through multiplying the static value of the real power injection (the value used in the previous subsection) by a random scalar drawn uniformly from $[0.75, 1.25]$. The parameters of **VC-LB-P** Algorithm is set as $\alpha = 0.2$, $\beta = 10^{-5}$, and $\rho = 0$, the same as that of Fig. 5.

The voltage profile and the reactive power injection of a selected set of buses over 8 intervals (4000 iterations) are shown in Fig. 6. The figure shows that in all cases the **VC-LB-P** Algorithm regulates the voltage profile within the 500-iteration (1000-bit) limit and in the meanwhile not violating the reactive power capacity constraint at any time. We further note that the time it takes to conduct a 1-bit communication, when using protocols for extremely low latency communications [13] (at the cost of limited data-rate), can be made extremely small. This suggests that the **VC-LB-P** Algorithm in this paper can be used to provide fast voltage regulation in future smart grids.

VI. Conclusion

This paper studied distributed voltage control algorithms where only a few bits of communication between neighboring buses are needed. The convergence of these algorithms was studied and their practical applicability illustrated in simulations. Future work is to 1) study the trade-offs between the data rate and control performance, 2) characterize the
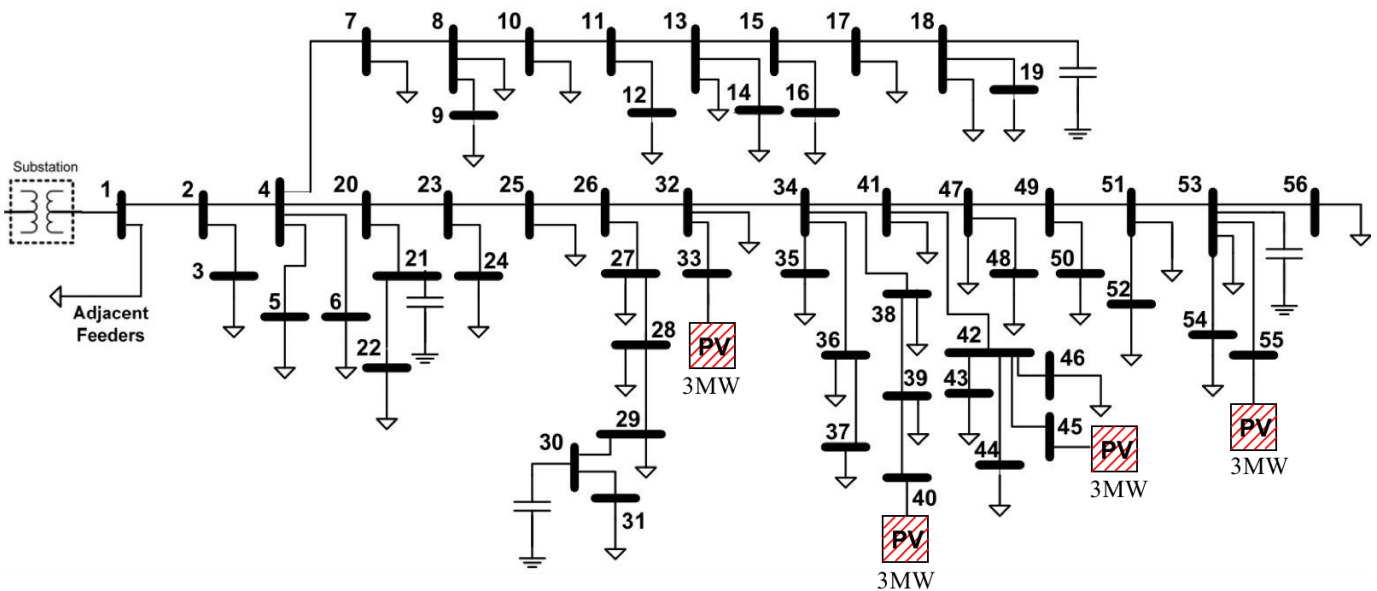
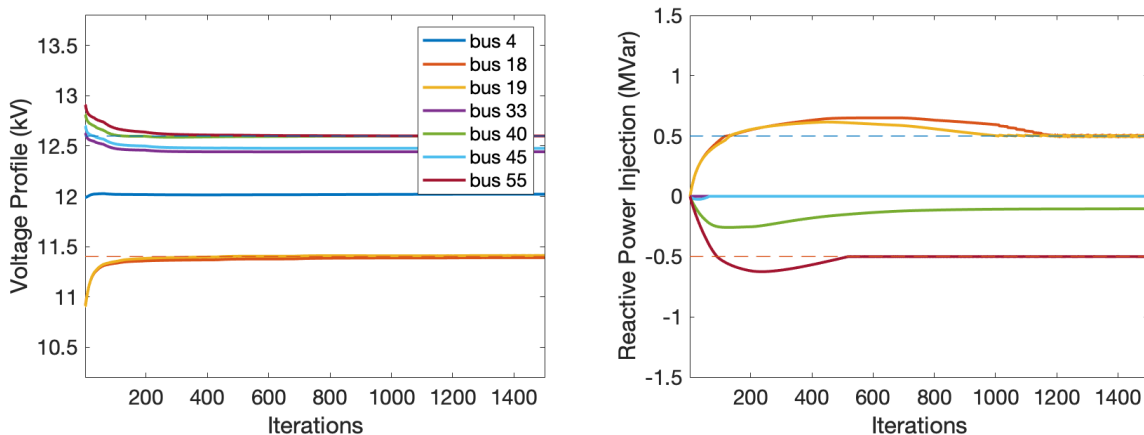Fig. 1: The 56 bus network used in our simulation.

Fig. 2: Convergence of the $\mathbf{v}(t)$ and $\mathbf{q}(t)$ in the **VC-LB Algorithm**. Parameters are $\alpha = 0.2, \beta = 10^{-5}, \rho = 0$. Notice that during each iteration, each node uses 2-bits to communicate.

online performance under dynamic operating scenarios, and 3) implement the algorithm on a real-world testbed.

REFERENCES

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

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

[3] N. Li, G. Qu, and M. Dahleh, "Real-time decentralized voltage control in distribution networks," in *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*, Sept 2014, pp. 582–588.

[4] 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, Sept 2016.

[5] B. Zhang, A. Y. S. Lam, A. D. Domnguez-Garca, 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, July 2015.

[6] 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.

[7] S. Bolognani and S. Zampieri, "A distributed control strategy for reactive power compensation in smart microgrids," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2818–2833, Nov 2013.

[8] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, "Distributed reactive power feedback control for voltage regulation and loss minimization," *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 966–981, April 2015.

[9] H. Jan Liu, W. Shi, and H. Zhu, "Hybrid voltage control in distribution networks under limited communication rates," *arXiv preprint arXiv:1707.07094*, 2017.

[10] G. Cavraro, S. Bolognani, R. Carli, and S. Zampieri, "The value of communication in the voltage regulation problem," in *2016 54th IEEE Conference on Decision and Control (CDC)*, Des 2016.

[11] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on smart grid communication infrastructures: Motivations, requirements and challenges," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 5–20, First 2013.

[12] S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: The role of power line communications in the smart grid," *Proceedings*
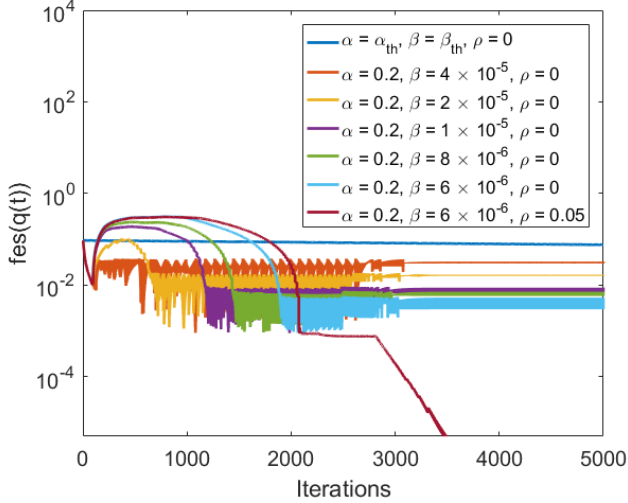
Fig. 3: $\mathrm{fes}(\mathbf{q}(t))$ in Equation (10) (in per unit system) under different values of α, β and ρ. α_{th} and β_{th} are the theoretic value of step sizes given in Equation (11) with $\epsilon = 1$ and their values are $\alpha_{\mathrm{th}} = 1.4093 \times 10^{-4}$, $\beta_{\mathrm{th}} = 6.4061 \times 10^{-7}$ under per unit system. The $\alpha = 0.2, \beta = 1 \times 10^{-5}, \rho = 0$ case has the final values of $\mathrm{fes}(\mathbf{q}(t))$ dropping below 1×10^{-2}, and the corresponding trajectory of $\mathbf{v}(t)$ and $\mathbf{q}(t)$ are in Fig. 2.

of the IEEE, vol. 99, no. 6, pp. 998–1027, 2011.

[13] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, Sept 2016.

[14] R. W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE transactions on Automatic Control*, vol. 45, no. 7, pp. 1279–1289, 2000.

[15] G. N Nair, F. Fagnani, S. Zampieri, and R. J Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[16] L. Zhang, H. Gao, and O. Kaynak, "Network-induced constraints in networked control systemsa survey," *IEEE transactions on industrial informatics*, vol. 9, no. 1, pp. 403–416, 2013.

[17] M. G. Rabbat and R. D. Nowak, "Quantized incremental algorithms for distributed optimization," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 4, pp. 798–808, April 2005.

[18] Angelia Nedic, Alex Olshevsky, Asuman Ozdaglar, and John N Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE, 2008, pp. 4177–4184.

[19] S. Magnússon, C. Enyioha, N. Li, C. Fischione, and V. Tarokh, "Convergence of limited communications gradient methods," *IEEE Transactions on Automatic Control*, vol. 63, no. 5, May 2018.

[20] S. Magnsson, C. Enyioha, N. Li, C. Fischione, and V. Tarokh, "Communication complexity of dual decomposition methods for distributed resource allocation optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 717–732, Aug 2018.

[21] Frédéric Olivier, Petros Aristidou, Damien Ernst, and Thierry Van Cutsem, "Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 926–936, 2016.

[22] Yuan Fan, Guoqiang Hu, and Magnus Egerstedt, "Distributed reactive power sharing control for microgrids with event-triggered communication," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 1, pp. 118–128, 2017.

[23] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr 1989.

[24] S. Magnússon, C. Fischione, and N. Li, "Voltage control using limited communication," in *20th IFAC World Congress*, 2017, pp. 1 – 6.

[25] A. Beck, A. Nedic, A. Ozdaglar, and M. Teboulle, "An o(1/k) gradient method for network resource allocation problems," *Control of Network Systems, IEEE Transactions on*, vol. 1, no. 1, pp. 64–73, March 2014.

[26] Na Li Guannan Qu, "An optimal and distributed feedback voltage control under limited reactive power," *arXiv, Cornell University Library*, 2018, [Online]. Available: https://arxiv.org/pdf/1810.11121.pdf.

[27] Guido Cavraro, Saverio Bolognani, Ruggero Carli, and Sandro Zampieri, "The value of communication in the voltage regulation problem," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 5781–5786.

[28] D. P. Bertsekas, *Nonlinear Programming: 2nd Edition*, Athena Scientific, 1999.

[29] Y. Nesterov, *Introductory Lectures on Convex Optimization*, Springer, 2004.

[30] S. Magnússon, G. Qu, N. Li, and C. Fischione, "Voltage control using limited communication," 2018, Avilable online: https://www.dropbox.com/s/uo8tw6x0k4hwdyf/voltage_V4.pdf?dl=0.

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

[32] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2011.

[33] A. Nedic and A. Ozdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods," *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1757–1780, 2009.

APPENDIX A
IMPORTANT LEMMAS

Lemma 1. *We have $\nabla D(\mathbf{z}) = \mathbf{M}\mathbf{z} + \mathbf{r}$, where*

$$\mathbf{M} = \begin{pmatrix} -\mathbf{A} & \mathbf{A} & -\mathbf{I} & \mathbf{I} \\ \mathbf{A} & -\mathbf{A} & \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} & -\mathbf{A}^{-1} & \mathbf{A}^{-1} \\ \mathbf{I} & -\mathbf{I} & \mathbf{A}^{-1} & -\mathbf{A}^{-1} \end{pmatrix}. \quad (31)$$

The eigenvalues of $\mathbf{M}$ are

$$\lambda_i(\mathbf{M}) = -2\left(\lambda_i(\mathbf{A}) + \frac{1}{\lambda_i(\mathbf{A})}\right), \quad for\ i = 1, \ldots, n \quad (32)$$

and $\lambda_i(\mathbf{M}) = 0$ for $i = n+1, \ldots, 4n$.

Proof: The fact that $\nabla D(\mathbf{z}) = \mathbf{M}\mathbf{z} + \mathbf{r}$ follows from Equations (20) and (21). To find the eigenvalues of $\mathbf{M}$, let $\mathbf{w}_i \in \mathbb{R}^n$ be an eigenvector of the matrix $\mathbf{A}$ associated to the eiginvalue $\lambda_i(\mathbf{A})$. Then direct calculations show that $\mathbf{M}\mathbf{v}_i = \lambda_i(\mathbf{M})\mathbf{v}$, with $\lambda_i(\mathbf{M})$ defined in Equation (32), where $\mathbf{v}_i = (\lambda_i(\mathbf{A})\mathbf{w}_i, -\lambda_i(\mathbf{A})\mathbf{w}_i, \mathbf{w}_i, -\mathbf{w}_i)$.

To prove that 0 is a an eigenvalue of $\mathbf{M}$ with multiplicity 3n, i.e., $\lambda_i(\mathbf{M}) = 0$ for $i = n+1, \ldots, 4n$, we note that $\mathbf{M}$ is a rank n matrix and hence has a $3n$ dimensional null space. This can be seen by noting that the rows 2, 3, and 4 of the block matrix $\mathbf{M}$, see Equation (31), can be obtained from elementary row operations (Gauss elimination) on the first row of $\mathbf{M}$. In particular, $\mathbf{R}_2 = -\mathbf{R}_1$, $\mathbf{R}_3 = \mathbf{A}\mathbf{R}_1$, $\mathbf{R}_4 = -\mathbf{A}\mathbf{R}_1$, where $\mathbf{R}_i$ is the i-th row of the block representation of $\mathbf{M}$ given in Equation (31). ∎

Lemma 2. *For $V(\cdot)$ defined in Equation (24) we have:*

a) *A feasible dual variable $\mathbf{z} \in \mathbb{R}_+^{4N}$ is an optimal solution to the Dual Problem (19) if and only if $V(\mathbf{z}) = 0$.*

b) *For all $\mathbf{z} \in \mathbb{R}_+^{4N}$ $\mathrm{fes}(\mathbf{q}(\mathbf{z})) \leq V(\mathbf{z})$, where $\mathbf{q}(\mathbf{z})$ is defined in Equation (21).*

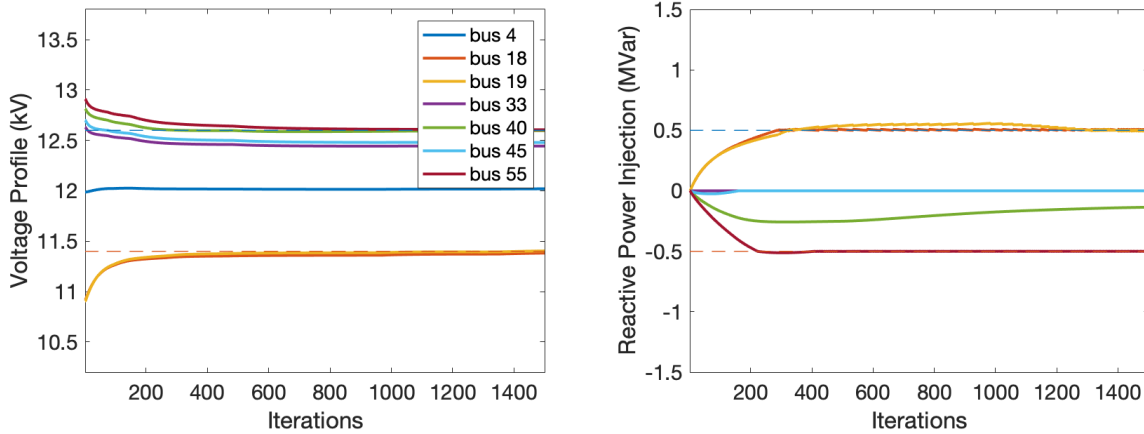*Proof:* a) Follows from [19, Lemma 3].

Fig. 4: Convergence of the $\mathbf{v}(t)$ and $\mathbf{q}(t)$ in the **VC-LB** Algorithm. Parameters are $\alpha = 0.08, \beta = 10^{-5}, \rho = 0$.

Fig. 5: Convergence of the $\mathbf{v}(t)$ and $\mathbf{q}^{phy}(t)$ in the **VC-LB-P** Algorithm. Parameters are $\alpha = 0.2, \beta = 10^{-5}, \rho = 0$. Note that the reactive power injection never exceeds the upper and lower limit.
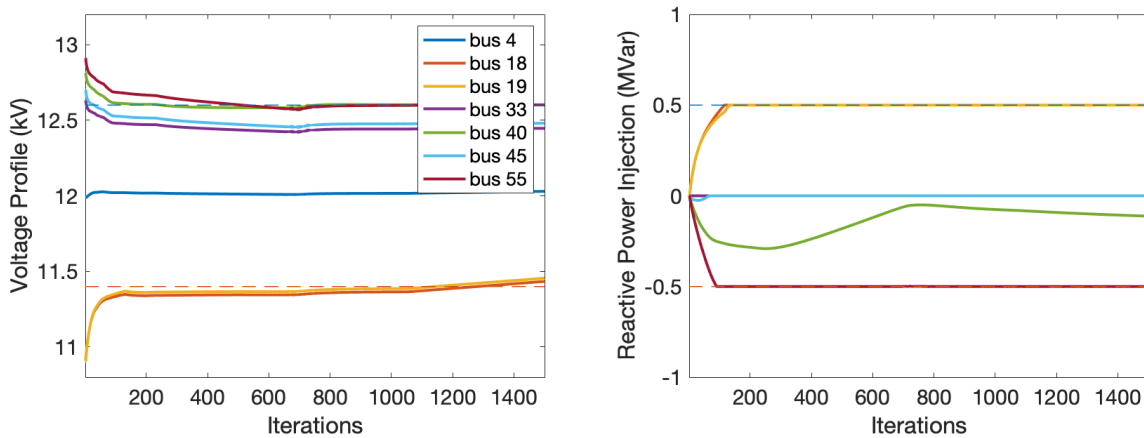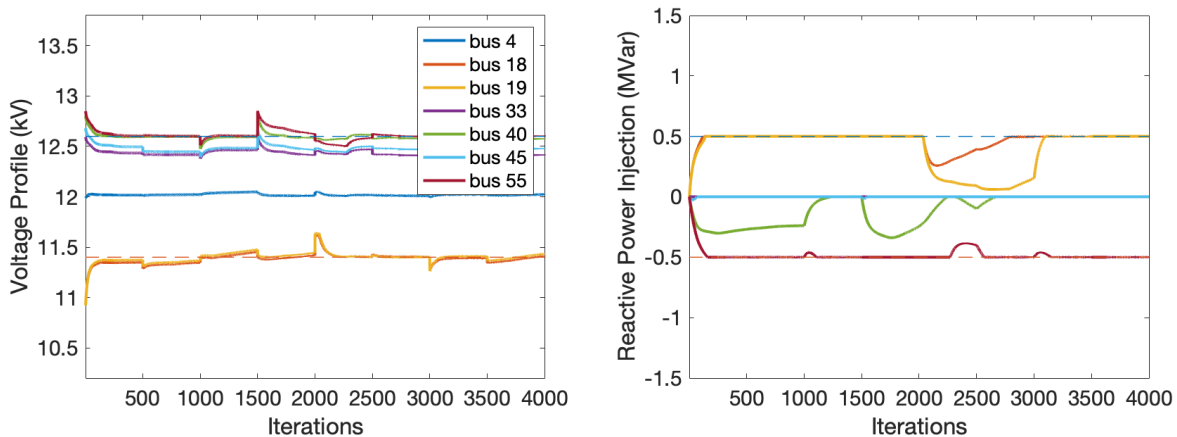

Fig. 6: Dynamic voltage control: convergence of $\mathbf{v}(t)$ and $\mathbf{q}^{phy}(t)$ in the **VC-LB-P Algorithm**. Parameters are $\alpha = 0.2, \beta = 10^{-5}, \rho = 0$.

b) We have that

$$\texttt{fes}(\mathbf{q}(\mathbf{z}))=\left|\left|\left\lceil\begin{bmatrix}\mathbf{v}^{\min}-\mathbf{v}(\mathbf{q}(\mathbf{z}))\\\mathbf{v}(\mathbf{q}(\mathbf{z}))-\mathbf{v}^{\max}\\\mathbf{q}^{\min}-\mathbf{q}(\mathbf{z})\\\mathbf{q}(\mathbf{z})-\mathbf{q}^{\max}\end{bmatrix}\right\rceil_{+}\right|\right|=||\lceil\nabla D(\mathbf{z})\rceil_{+}||,$$

where the later equality comes from Equation (20). Using that

$$\lceil\nabla_i D(\mathbf{z})\rceil_+ \le |\mathbf{z}_i - \lceil\mathbf{z}_i + \nabla_i D(\mathbf{z})\rceil_+|$$

for $i = 1,\ldots,4N$, we get that

$$\begin{aligned}||\lceil\nabla D(\mathbf{z})\rceil_+||^2 &= \sum_{i=1}^{4N}\lceil\nabla_i D(\mathbf{z})\rceil_+^2\\&\le \sum_{i=1}^{4N}|\mathbf{z}_i - \lceil\mathbf{z}_i + \nabla_i D(\mathbf{z})\rceil_+|^2 = V(\mathbf{z})^2,\end{aligned}$$

which proves the result. ∎

Lemma 3. *For all $\beta \in [0,1]$, $z \in \mathbb{R}$ and $x, \alpha_1, \alpha_2 \in \mathbb{R}_+$ with $\alpha_1 \le |x - \lceil x + z\rceil_+|$ following holds*

$$\beta|x - \lceil x + z\rceil^+| \le |x - \lceil x + \beta z\rceil^+| \tag{33}$$
$$\alpha_1 = |x - \lceil x + \alpha_1 sign(z)\rceil_+| \tag{34}$$
$$0 \le z(\lceil x + \alpha_2 z\rceil^+ - x). \tag{35}$$

Proof: The proof follows similar steps as the proof of [19, Lemma 9]. In particular, the inequalities (33) and (34) are easily checked by using that for $x \in \mathbb{R}_+$ and $z \in \mathbb{R}$ we have $|x - \lceil x + z\rceil^+| = |z|$ if $x + z \ge 0$ and $|x - \lceil x + z\rceil^+| = x$ if $x + z \le 0$. The Inequality (35) comes by rearrange the inequality $\texttt{sign}(z)\lceil x + \alpha_1 z\rceil_+ \le \texttt{sign}(z)x$ and multiply $|z|$ on both sides. ∎

Lemma 4. *Suppose that the voltage regulation problem is strictly feasible, i.e., there exists some $\mathbf{q} \in \mathbb{R}^N$ such that $\mathbf{q} \in (\mathbf{q}^{\min}, \mathbf{q}^{\max})$ and $\mathbf{v}(\mathbf{q}) \in (\mathbf{v}^{\min}, \mathbf{v}^{\max})$. Then the set $\mathcal{Z}^\star$ of optimizers of the dual problem given in Equation 19 is bounded.*

Proof: Follows directly from Lemma 1 in [33]. ∎

Lemma 5. *For $\epsilon > 0$, there exists $\eta > 0$ such that if $D^\star - D(\mathbf{z}) \le \eta$ then $\texttt{dist}(\mathbf{z}, \mathcal{Z}^\star) \le \epsilon$.*

Proof: We prove this by contradiction. Suppose that no such $\eta > 0$ exists. Then we can generate a sequence $\mathbf{z}_k \in \mathbb{R}_+$ such that $\lim_{k\to\infty} D(\mathbf{z}_k) = D^\star$ and $\texttt{dist}(\mathbf{z}_k, \mathcal{Z}^\star) > \epsilon$ for all $k \in \mathbb{N}$. The sequence $\mathbf{z}_k$ is bounded because $\mathcal{Z}^\star$ is bounded, see Lemma 4. Therefore, the level sets $\mathcal{L}(\kappa) = \{\mathbf{z} \in \mathbb{R}_+^{4N} | D(\mathbf{z}) \ge D^\star - \kappa\}$ are also bounded, see [28, Proposition B.9]. As a result, $\mathbf{z}_k$ has a convergent subsequence $\mathbf{z}_{k_r}$ with $\lim_{r\to\infty} \mathbf{z}_{k_r} = \bar{\mathbf{z}} \in \mathbb{R}_+$. Then $D(\bar{\mathbf{z}}) = \lim_{r\to\infty} D(\mathbf{z}_{k_r}) = D^\star$, so $\bar{\mathbf{z}} \in \mathcal{Z}^\star$. This contradicts the fact that $\texttt{dist}(\mathbf{z}_k, \mathcal{Z}^\star) > \epsilon$ for all $k \in \mathbb{N}$, since $\texttt{dist}(\mathbf{z}_{k_r}, \mathcal{Z}^\star) \le ||\mathbf{z}_{k_r} - \bar{\mathbf{z}}|| \le \epsilon$ for sufficiently large r, which yields the result. ∎

APPENDIX B
PROOF OF PROPOSITION 2

The dual function $D(\cdot)$ is concave and with L-Lipschitz continues gradient so we have [29, Equation (2.1.6)]

$$D(\lceil\bar{\mathbf{z}}\rceil^+) \ge D(\mathbf{z}) + \langle\nabla D(\mathbf{z}), \lceil\bar{\mathbf{z}}\rceil^+ - \mathbf{z}\rangle - \frac{L}{2}||\lceil\bar{\mathbf{z}}\rceil^+ - \mathbf{z}||^2,$$

$$\begin{aligned}&=D(\mathbf{z}) + \langle\nabla^{\boldsymbol{\lambda}}D(\mathbf{z}), \lceil\bar{\boldsymbol{\lambda}}\rceil_+ - \boldsymbol{\lambda}\rangle - \frac{L}{2}||\lceil\bar{\boldsymbol{\lambda}}\rceil_+ - \boldsymbol{\lambda}||^2\\&\quad + \langle\nabla^{\boldsymbol{\mu}}D(\mathbf{z}), \lceil\bar{\boldsymbol{\mu}}\rceil_+ - \boldsymbol{\mu}\rangle - \frac{L}{2}||\lceil\bar{\boldsymbol{\mu}}\rceil_+ - \boldsymbol{\mu}||^2\\&\ge D(\mathbf{z}) + \alpha\left(1 - \frac{L}{2}\alpha\right)||\boldsymbol{\lambda} - \lceil\boldsymbol{\lambda} + \nabla^{\boldsymbol{\lambda}}D(\mathbf{z})\rceil_+||^2\\&\quad + \langle\nabla^{\boldsymbol{\mu}}D(\mathbf{z}), \lceil\bar{\boldsymbol{\mu}}\rceil_+ - \boldsymbol{\mu}\rangle - \frac{L}{2}||\lceil\bar{\boldsymbol{\mu}}\rceil_+ - \boldsymbol{\mu}||^2, \quad (36)\end{aligned}$$

where the final inequality is proved below. From Lemma 2 we have $V(\mathbf{z}) \ge \texttt{fes}(\mathbf{q}(\mathbf{z})) > \epsilon$, yielding

$$\epsilon^2 \le V(\mathbf{z})^2 = ||\boldsymbol{\lambda} - \lceil\boldsymbol{\lambda} + \nabla^{\boldsymbol{\lambda}}D(\boldsymbol{\lambda})\rceil_+||^2 + ||\boldsymbol{\mu} - \lceil\boldsymbol{\mu} + \nabla^{\boldsymbol{\mu}}D(\boldsymbol{\mu})\rceil_+||^2.$$

Consider the two cases separately when (a)

$$||\boldsymbol{\mu} - \lceil\boldsymbol{\mu} + \nabla^{\boldsymbol{\mu}}D(\boldsymbol{\mu})\rceil_+||^2 \ge \frac{\epsilon^2}{2}$$

and (b)

$$||\boldsymbol{\mu} - \lceil\boldsymbol{\mu} + \nabla^{\boldsymbol{\mu}}D(\boldsymbol{\mu})\rceil_+||^2 < \frac{\epsilon^2}{2} \quad \text{and}$$
$$||\boldsymbol{\lambda} - \lceil\boldsymbol{\lambda} + \nabla^{\boldsymbol{\lambda}}D(\boldsymbol{\lambda})\rceil_+||^2 \ge \frac{\epsilon^2}{2}.$$

In case (a) we have (proved below)

$$D(\lceil\bar{\mathbf{z}}\rceil^+) \ge D(\mathbf{z}) + \left(\frac{\epsilon}{2N^{3/2}L} - \beta\right)\beta LN, \quad (37)$$

and case (b) we have (proved below)

$$D(\lceil\bar{\mathbf{z}}\rceil^+) \ge D(\mathbf{z}) + \alpha\left(1 - \frac{L}{2}\alpha\right)\frac{\epsilon^2}{2} - NL\beta^2. \quad (38)$$

Then Equation (26) is obtained by Equations (37) and (38).

Proof of Equation (36): We have

$$\begin{aligned}\langle\nabla^{\boldsymbol{\lambda}}D(\mathbf{z}), \lceil\bar{\boldsymbol{\lambda}}\rceil_+ - \boldsymbol{\lambda}\rangle &= \sum_{i=1}^{2N}|\nabla_i^{\boldsymbol{\lambda}}D(\mathbf{z})||\lceil\boldsymbol{\lambda}_i + \alpha\nabla_i^{\boldsymbol{\lambda}}D(\mathbf{z})\rceil_+ - \boldsymbol{\lambda}_i|\\&\ge \sum_{i=1}^{2N}\frac{1}{\alpha}(\lceil\boldsymbol{\lambda}_i + \alpha\nabla_i^{\boldsymbol{\lambda}}D(\mathbf{z})\rceil_+ - \boldsymbol{\lambda}_i)^2 \quad (39)\end{aligned}$$

where the equality comes by using that each term of the sum is positive from Equation (35) and the inequality comes by using the non-expansiveness of the projection to get

$$|\nabla_i^{\boldsymbol{\lambda}}D(\mathbf{z})| \ge \frac{1}{\alpha}\left|\lceil\boldsymbol{\lambda}_i + \alpha\nabla_i^{\boldsymbol{\lambda}}D(\mathbf{z})\rceil_+ - \boldsymbol{\lambda}_i\right|.$$

We also have

$$||\lceil\bar{\boldsymbol{\lambda}}\rceil_+ - \boldsymbol{\lambda}||^2 = \sum_{i=1}^{2N}(\lceil\boldsymbol{\lambda}_i + \alpha\nabla_i^{\boldsymbol{\lambda}}D(\mathbf{z})\rceil_+ - \boldsymbol{\lambda}_i)^2. \quad (40)$$

By combining (39) and (40) we get

$$\langle \nabla^{\boldsymbol\lambda} D(\mathbf{z}), \lceil \bar{\boldsymbol\lambda} \rceil_+ - \boldsymbol\lambda \rangle - \frac{L}{2} ||\lceil \bar{\boldsymbol\lambda} \rceil_+ - \boldsymbol\lambda||^2$$

$$\geq \left(\frac{1}{\alpha} - \frac{L}{2} \right) \sum_{i=1}^{2N} (\lceil \boldsymbol\lambda_i + \alpha \nabla^{\boldsymbol\lambda}_i D(\mathbf{z}) \rceil_+ - \boldsymbol\lambda_i)^2$$

$$\geq \alpha^2 \left(\frac{1}{\alpha} - \frac{L}{2} \right) \sum_{i=1}^{2N} (\lceil \boldsymbol\lambda_i + \nabla^{\boldsymbol\lambda}_i D(\mathbf{z}) \rceil_+ - \boldsymbol\lambda_i)^2$$

$$\geq \alpha \left(1 - \frac{L}{2}\alpha \right) ||\boldsymbol\lambda - \lceil \boldsymbol\lambda + \nabla^{\boldsymbol\lambda} D(\mathbf{z}) \rceil_+||^2,$$

where the second inequality comes by using (33) and the fact that $\alpha < 2/L < 1$ (see Lemma 1).

Proof of Equation (37)**:** Suppose that

$$||\boldsymbol\mu - \lceil \boldsymbol\mu + \nabla^{\boldsymbol\mu} D(\boldsymbol\mu) \rceil_+||^2 \geq \epsilon^2/2.$$

Then we have

$$\langle \nabla^{\boldsymbol\mu} D(\mathbf{z}), \lceil \bar{\boldsymbol\mu} \rceil_+ - \boldsymbol\mu \rangle$$

$$= \sum_{i=1}^{2N} |\nabla^{\boldsymbol\mu}_i D(\mathbf{z})| |\lceil \boldsymbol\mu_i + \beta \mathtt{sign}(\nabla^{\boldsymbol\mu}_i D(\mathbf{z})) \rceil_+ - \boldsymbol\mu_i|$$

$$\geq |\boldsymbol\mu_j - \lceil \boldsymbol\mu_j + \nabla^{\boldsymbol\mu}_j D(\mathbf{z}) \rceil_+| |\boldsymbol\mu_j - \lceil \boldsymbol\mu_j + \beta \mathtt{sign}(\nabla^{\boldsymbol\mu}_j D(\mathbf{z})) \rceil_+|$$

$$\geq \frac{\epsilon}{2\sqrt{N}} \beta \tag{41}$$

where $j = \mathrm{argmax}_{i=1,\cdots,2N} |\boldsymbol\mu_i - \lceil \boldsymbol\mu_i + \nabla^{\boldsymbol\mu}_i D(\mathbf{z}) \rceil_+|$, the equality comes by using that each term of the sum is positive from Equation (35), the first inequality from the non-expansiveness of the projection [28, Proposition B.11], and the final inequality comes from the fact that $||\boldsymbol\mu - \lceil \boldsymbol\mu + \nabla^{\boldsymbol\mu} D(\mathbf{z}) \rceil_+||_\infty \geq \epsilon/(\sqrt{2}\sqrt{2N})$, since $||\boldsymbol\mu - \lceil \boldsymbol\mu + \nabla^{\boldsymbol\mu} D(\mathbf{z}) \rceil_+|| \geq \epsilon/\sqrt{2}$, and Equation (34) together with the fact that $\beta \leq \epsilon/(2\sqrt{N}) \leq |\boldsymbol\mu_j - \lceil \boldsymbol\mu_j + \nabla^{\boldsymbol\mu}_j D(\mathbf{z}) \rceil_+|$. We also have

$$||\boldsymbol\mu - \lceil \bar{\boldsymbol\mu} \rceil_+||^2 = \sum_{i=1}^{2N} |\boldsymbol\mu_i - \lceil \boldsymbol\mu_i + \beta \mathtt{sign}(\nabla^{\boldsymbol\mu}_i D(\mathbf{z})) \rceil_+|^2$$

$$\leq 2N\beta^2. \tag{42}$$

Combining Equations (41) and (42) and rearranging yields

$$\langle \nabla^{\boldsymbol\mu} D(\mathbf{z}), \lceil \bar{\boldsymbol\mu} \rceil_+ - \boldsymbol\mu \rangle - \frac{L}{2} ||\lceil \bar{\boldsymbol\mu} \rceil_+ - \boldsymbol\mu||^2 \geq \left(\frac{\epsilon}{2N^{3/2}L} - \beta \right) \beta LN,$$

which proves Equation (37).

Proof of Equation (38)**:** The result follows from that

$$\langle \nabla^{\boldsymbol\mu} D(\mathbf{z}), \lceil \bar{\boldsymbol\mu} \rceil_+ - \boldsymbol\mu \rangle - \frac{L}{2} ||\lceil \bar{\boldsymbol\mu} \rceil_+ - \boldsymbol\mu||^2 \geq -NL\beta^2, \tag{43}$$

where inequality comes by that

$$||\boldsymbol\mu - \lceil \bar{\boldsymbol\mu} \rceil_+||^2 = \sum_{i=1}^{2N} (\boldsymbol\mu - \lceil \bar{\boldsymbol\mu} \rceil_+)^2 \leq 2N\beta^2.$$

and $\langle \nabla^{\boldsymbol\mu} D(\mathbf{z}), \lceil \bar{\boldsymbol\mu} \rceil_+ - \boldsymbol\mu \rangle \geq 0$, from Equation (35).