

Optimal Voltage Control Using Event Triggered Communication

Sindri Magnusson
sindrim@seas.harvard.edu
Harvard University

Carlo Fischione
carlofi@kth.se
KTH Royal Institute of Technology

Na Li
nali@seas.harvard.edu
Harvard University

ABSTRACT

The integration of volatile renewable energy into distribution networks on a large scale will demand advanced voltage control algorithms. Communication will be an integral part of these algorithms, however, it is unclear what kind of communication protocols will be most effective for the task. Motivated by such questions, this paper investigates how voltage control can be achieved using event triggered communications. In particular, we consider online algorithms that require the network's buses to communicate only when their voltage is outside a feasible operation range. We prove the convergence of these algorithms to an optimal operating point at the rate $\mathcal{O}(1/t)$, assuming linearized power flows. We illustrate the performance of the algorithms on the full nonlinear AC power flow in simulations. Our results show that event-triggered protocols can significantly reduce the communication for smart grid control.

CCS CONCEPTS

• **Computer systems organization** → *Sensor networks*.

ACM Reference Format:

Sindri Magnusson, Carlo Fischione, and Na Li. 2019. Optimal Voltage Control Using Event Triggered Communication. In *e-Energy '19: Proceedings of the Tenth ACM International Conference on Future Energy Systems, June 25–28, 2019, Phoenix, AZ, USA*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3307772.3328303>

1 INTRODUCTION

Smart grids are meant to integrate large amounts of renewable energy resources into our distribution networks. For this to be possible, the voltage fluctuations from renewable resources must be countered with advanced control [8]. A prominent solution has been to use the injection of reactive powers from solar or PV-inverters to regulate the voltage fluctuations in the network. Much work has considered purely local controllers where each bus with an inverter regulates its reactive power injection based only on local voltage measurements, see [13, 16, 26] and references therein. However,

it has been shown that such purely local control algorithms are generally insufficient in voltage regulation [9, 16]. This means that the voltage regulation is generally not possible without some form of communication. This has sparked much research on distributed algorithms where the buses regulate their voltage by using some form of communication with other buses [6, 7, 17, 18, 22, 25].

In networked systems, event triggered control and communication algorithms have the potentiality to reduce the control and communication costs [11, 23]. These algorithms are based on performing a control action or communication only when the system demands it instead of at predetermined time instances. In fact, event triggered communication has been considered in the power system literature [12, 15, 24]. The work in [15, 24] considers the economic dispatch problem and the DC-OPF problem, respectively, and therefore does not deal with reactive power or voltage regulation. The work in [12] is more related to our work and considers event triggered algorithms for distributed reactive power sharing in micro grids. However, differently from [12], we consider algorithms that find an optimal operation point of the voltage regulation in terms of reactive power costs, instead of just stabilizing the system. Moreover, unlike [12], the algorithms developed in this paper enforce hard operation limits on the reactive power injections and also (in the limit) on the voltages, which must usually be respected in real-life systems.

The goal of this work is to develop algorithms for optimal voltage control using event triggered communication between the network's buses. The buses control their reactive power injection (within hard operation limits) based on local voltage measurements and possibly a communication of local Lagrangian multipliers that are associated with a violation of the local voltage constraints. The Lagrangian multiplier of each bus can intuitively be considered a price to the other nodes for violating the bus's voltage constraint. In the considered algorithms the network's buses only need to communicate at critical times when both a) the local voltage constraint is violated and b) when the difference between local and global version of the Lagrangian multipliers is significant enough. Hence, these algorithms have the potential to reduce the communication need significantly. We consider two communication structures: a) where each bus broadcasts its local Lagrangian multiple to the other buses and b) where each bus broadcasts its local Lagrangian multiple to a System Operator that then coordinates the algorithm (also in an event triggered fashion). We prove the convergence of the algorithm to an optimal operation point at the rate $\mathcal{O}(1/\sqrt{t})$ when the power flows are linearized. We also show how this convergence rate can be accelerated to $\mathcal{O}(1/t)$ by intelligent parameter tuning. We illustrate the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

e-Energy '19, June 25–28, 2019, Phoenix, AZ, USA

© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6671-7/19/06...\$15.00
<https://doi.org/10.1145/3307772.3328303>

performance of the algorithm on the full nonlinear AC power flow in simulations.

1.1 Notation

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 sets. The superscript $(\cdot)^T$ stands for transpose. $\text{diag}(A_1, \dots, A_n)$ denotes the diagonal block matrix with A_1, \dots, A_n on the diagonal. We define the proximal operator of a function g as $\text{prox}_\gamma(g)(z) = \underset{y}{\text{argmin}} g(y) + \frac{1}{2\gamma} \|x - z\|$ (following the notation in [4]). $\|\cdot\|$ denotes the 2-norm for both vectors and matrices. That is $\|a\| = \sqrt{\sum_{i=1}^n a_i^2}$ and $\|A\| = \sqrt{\sum_{i,j=1}^n A_{ij}^2}$. The projection of $z \in \mathbb{R}$ to $[a, b] \subseteq \mathbb{R}$ is denoted by $[z]_a^b$.

2 SYSTEM MODEL AND PROBLEM FORMULATION

2.1 System Model

Consider a radial power distribution network with $N + 1$ buses represented by the set $\mathcal{N}_0 = \{0\} \cup \mathcal{N}$, where $\mathcal{N} = \{1, \dots, N\}$. 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 . The relationship between power and voltage can be given by the branch flow model for radial networks [1, 2] as

$$-p_i = P_{\sigma_i i} - r_{\sigma_i, i} |I_{\sigma_i, i}|^2 - \sum_{k:(i, k) \in \mathcal{E}} P_{ik}, \quad i \in \mathcal{N}, \quad (1a)$$

$$-q_i = Q_{\sigma_i i} - x_{\sigma_i, i} |I_{\sigma_i, i}|^2 - \sum_{k:(i, k) \in \mathcal{E}} Q_{ik}, \quad i \in \mathcal{N}, \quad (1b)$$

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

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

where σ_i is the parent of bus $i \in \mathcal{N}$. A subset $\mathcal{I} \subseteq \mathcal{N}$ of \mathcal{I} buses have a PV-inverter that can inject reactive power to the system to help regulate voltage fluctuations. Without loss of generality, we assume that $\mathcal{I} = \{1, \dots, I\}$. Denote the voltage and reactive power on the inverter buses by

$$v = [v_1, \dots, v_I]^T \in \mathbb{R}^I \quad \text{and} \quad q = [q_1, \dots, q_I]^T \in \mathbb{R}^I.$$

In this paper, we design algorithms for the full nonlinear power flow model in Eq. (1). However, in our algorithm development we use information from the linearized branch flow model [1], which has been shown to give good approximation of distribution circuits. We will also prove the convergence of our algorithm under the linear model. The linear model is

obtained by setting $|I_{ij}|^2 = 0$, which is generally a fair approximation since $|I_{ij}|^2$ is usually of much smaller magnitude than v_i , P_{ij} , and Q_{ij} [1]. By rearranging Eq. (1) we get that

$$v = Xq + d, \quad (2)$$

where $d = X^{\text{NI}}q^{\text{NI}} + Rp + \mathbf{1}v_0$, $q^{\text{NI}} = [q_{I+1}, \dots, q_N]^T$ is the reactive power at the non-inverter buses in $\mathcal{N} \setminus \mathcal{I}$, $p = [p_1, \dots, p_N]^T$ is the real power injection at each bus, and the matrices $X \in \mathbb{R}^{I \times I}$, $X^{\text{NI}} \in \mathbb{R}^{M \times N-I}$, and $R \in \mathbb{R}^{I \times N}$ are given as

$$X_{ij} = 2 \sum_{(h, k) \in \mathcal{P}_i \cap \mathcal{P}_j} x_{hk}, \quad X_{ij}^{\text{NI}} = 2 \sum_{(h, k) \in \mathcal{P}_i \cap \mathcal{P}_{M+j}} x_{hk}, \quad R_{ij} = 2 \sum_{(h, k) \in \mathcal{P}_i \cap \mathcal{P}_j} r_{hk},$$

where $\mathcal{P}_i \subseteq \mathcal{E}$ is the set of edges in the path from Bus 0 to i .

2.2 Optimal Voltage Control

When the real power injection p has been decided, then voltage control is needed to tune the reactive power injection q on the inverter buses to adjust their voltages v within some feasible range $v \in [\underline{v}, \bar{v}]$. The reactive power injection can be adjusted within some interval $q \in [q, \bar{q}]$.¹ From the previous section, we can define the output voltage v when we input the real and reactive powers p and q into the circuit as follows:

$$v(q) := \text{Solution to Eq. (1) for } p \text{ and } q \text{ given.} \quad (3)$$

Ideally, we wish to find the optimal reactive power by solving an optimization problem:

$$\begin{aligned} & \underset{q_1, \dots, q_n}{\text{minimize}} && \sum_{i \in \mathcal{I}} C_i(q_i) := \frac{a_i}{2} q_i^2 + b_i q_i + c_i \\ & \text{subject to} && \underline{v} \leq v(q) \leq \bar{v} \\ & && \underline{q} \leq q \leq \bar{q}. \end{aligned} \quad (4)$$

where $C_i(q_i)$ is the quadratic cost of injecting q_i units of reactive power at bus i . We are often only interested in solving the feasibility problem, i.e., the problem of finding any reactive powers q that enforce the voltages inside the feasible range $\underline{v} \leq v(q) \leq \bar{v}$. In this case we simply put $a_i = 1$ and $b_i = 0$. However, there are situations where some feasible reactive powers are more desirable than others, e.g., in PV generators we generally want the apparent power to be small so it is desirable to minimize q_i^2 . Moreover, in dynamic environments where p is changing quickly we might want to reduce the fluctuations in q_i over time by proper choice of $C_i(\cdot)$. The goal of this paper is to establish algorithms that solve the problem using event-triggered communication.

2.3 Event-Triggered Algorithms

We study how the optimization problem can be solved with distributed feedback control algorithms. We could consider local algorithms where each bus $i \in \mathcal{I}$ updates its reactive power based only on local voltage measurements²

$$\text{Measurement:} \quad v_i(t) = v_i(q(t)) \quad (5a)$$

$$\text{Control:} \quad q_i(t+1) = K_i(v_i(t)), \quad (5b)$$

¹The reactive power can be decomposed into $q = q^{\text{Adj.}} + q^{\text{Con.}}$ where $q^{\text{Adj.}}$ is adjustable and $q^{\text{Con.}}$ is fixed.

²With slight abuse of notation we denote $v_i(t) := v_i(q(t))$.

where $K_i(\cdot)$ is some local controller. However, the impossibility result in [9] shows that such local controllers can in general not even find a feasible solution to (4). Therefore, we must include some form of communication in the control law in Eq. (5b). Such communication can be modeled as follows:

$$q_i(t+1) = K_i(v_i(t), \text{Comm}_i(t)), \quad (6)$$

where $\text{Comm}_i(t)$ is information that bus i has received from other buses until iteration t .

In this paper, we study how the communication at each bus can be limited to only critical events, e.g., when the bus's voltage is violated. In particular, we let the communicated information $\text{Comm}_i(t)$ be *event-driven*. That means that each bus $i \in \mathcal{I}$ communicates a local Lagrangian multiplier related the local voltage constraint $[\underline{v}_i, \bar{v}_i]$ only when the constraint is violated and when there has been significant enough change in the Lagrangian multiplier. This triggering event (that $v_i(t) \notin [\underline{v}_i, \bar{v}_i]$) is a natural choice since the buses need the help from the other buses to maintain their voltage in the feasible range $[\underline{v}_i, \bar{v}_i]$, while the reactive powers can be maintained locally.

REMARK 1. *We consider both algorithms were a) the inverter buses broadcast information to the other inverter buses and b) where the coordination is managed by a system operator (still in event-trigger fashion). We note that the former communication scheme a) is reasonable in current systems where reactive controllers are usually sparse and the set \mathcal{I} is small. Or in micro grids and even in larger grids if the number of communications can be kept sufficiently low by the event-triggered mechanism. However, in general large networks communication scheme b) is more suitable.*

3 ALGORITHM: EVENT TRIGGERED OPTIMAL VOLTAGE CONTROL

In this section we describe our event triggered voltage control algorithms for solving Problem (4). In Section 3.1 we consider a distributed event triggered algorithm where the nodes broadcast relevant information when needed. In Section 3.2 we consider a variant of the algorithm where the solution is coordinated by a system operator which both receives and sends information to the buses in an event triggered fashion.

3.1 Distributed Algorithm

The following event triggered algorithm solves the optimization problem in Eq. (4). We discuss its main steps and properties below.

DIST-ETVC Algorithm: Distributed Event Triggered Optimal Voltage Control

STEP 1 Initialization: Set $t = 0$, and $\lambda_i(0) = \lambda_i^{\text{Local}}(0) = \underline{\lambda}_i^{\text{Local}}(0) = \bar{\lambda}_i^{\text{Local}}(0) = \lambda_i^j(0) = 0$ for all $i, j \in \mathcal{I}$.

STEP 2 Local Control: Each bus $i \in \mathcal{I}$ injects into the grid the reactive power

$$q_i(t) = \left[\frac{1}{a_i} \left(\sum_{j \in \mathcal{I}} X_{ij} \lambda_i^j(t) - b_i \right) \right]_{q_i}^{\bar{q}_i} \quad (7)$$

where a_i and b_i come from the cost in Eq. (4).

STEP 3 Local Measurement and Computation: Each bus $i \in \mathcal{I}$ measures the voltage magnitude

$$v_i(t) = v_i(q(t))$$

and then updates $\underline{\lambda}_i$ and $\bar{\lambda}_i$ as follows

$$\underline{\lambda}_i^{\text{Local}}(t+1) = \lceil \underline{\lambda}_i^{\text{Local}}(t) + \gamma(\underline{v}_i - v_i(t)) \rceil_+ \quad (8a)$$

$$\bar{\lambda}_i^{\text{Local}}(t+1) = \lceil \bar{\lambda}_i^{\text{Local}}(t) + \gamma(v_i(t) - \bar{v}_i) \rceil_+, \quad (8b)$$

$$\lambda_i^{\text{Local}}(t+1) = \underline{\lambda}_i^{\text{Local}}(t+1) - \bar{\lambda}_i^{\text{Local}}(t+1) \quad (8c)$$

where $\gamma > 0$ is a step-size parameter.

STEP 4 Event Triggered Communication: Each bus $i \in \mathcal{I}$ decides whether it should communicate as follows.

• **If either**

$$(a) \lambda_i^{\text{Local}}(t+1) = 0 \text{ and } \lambda_i^{\text{Local}}(t) = 0 \text{ or}$$

$$(b) |\lambda_i^{\text{Local}}(t+1) - \lambda_i(t)| \leq \text{ET}(t) :$$

Then do not communicate and set

$$\lambda_i(t+1) = \lambda_i(t).$$

• **Otherwise:** Set

$$\lambda_i(t+1) = \lambda_i^{\text{Local}}(t+1)$$

and broadcast $\lambda_i(t+1)$ to the other nodes.

Then each inverter bus $i \in \mathcal{I}$ can update its local version of λ as follows

$$\lambda_i^j(t+1) = \begin{cases} \lambda_j(t) & \text{if no communication is} \\ & \text{received from bus } j \\ \lambda_j(t+1) & \text{otherwise.} \end{cases}$$

STEP 5 Set $t = t + 1$ and go back to **STEP 2**.

The algorithm is a variant of a dual gradient/decomposition method where $\underline{\lambda}_i^{\text{Local}}$ and $\bar{\lambda}_i^{\text{Local}}$ are the dual variables related to the constraints $\underline{v}_i \leq v_i(q)$ and $v_i(q) \leq \bar{v}_i$, respectively, as we show in Section 5. The local dual variable $\lambda_i^{\text{Local}} = \underline{\lambda}_i^{\text{Local}} - \bar{\lambda}_i^{\text{Local}}$ at bus i can be thought of as a price for violating bus's i voltage constraint. In particular, when $\lambda_i^{\text{Local}} > 0$ then bus i needs help from the other buses to decrease its voltage $v_i(t)$ and when $\lambda_i^{\text{Local}} < 0$ then bus i needs help to increase its voltage. The variable λ_i denotes the most recent value of λ_i^{Local} that bus i has communicated, and $\lambda_i^j = \lambda_i$ is a local copy node j has of λ_i . The variables λ_i and λ_i^{Local} are generally not the same since the nodes do not communicate at every iteration, see discussion below. In **STEP 1** the algorithm is initialized, we initialize the variables at zero but other values could be considered as well. The local reactive power control in **Step 2** is based on the dual variables and is obtained by solving a local sub-optimization-problem, which is a typical part of dual gradient methods, see next two sections for the details. The local λ -update in **Step 3** is a standard gradient update for the dual problem of the primal problem in Eq. (4). This step requires no communication since $v_i(t)$ can be measured after the buses inject their reactive power $q_i(t)$ into the system.

The main difference from standard dual gradient methods is in the communication in **Step 4**. The communication is not carried out at every iteration, instead it is triggered by the dual variables. In particular, bus i does not communicate when either condition (a) or (b) take place but communicates otherwise. In case (a), when $\bar{\lambda}_i^{\text{local}}(t+1) = 0$ then neither constraint $v_i \leq v_i(q)$ nor $v_i(q) \leq \bar{v}_i$ is violated. Therefore, the node does not need help from the other nodes to optimize its local objective function $\frac{a_i}{2}q_i^2 + b_iq_i + c_i$. This means that if the algorithm converges to the optimal solution q^* and $v^* = v(q^*)$, then bus i needs to communicate only finite number of times provided that v_i^* is strictly inside $[v_i, \bar{v}_i]$. In general, $\lambda_i^{\text{local}}(t+1)$ changes from being 0 only if the voltage limit $v_i(t) \in [v_i, \bar{v}_i]$ is violated (see Eq. (8)). Therefore, the first communication of each bus happens the first time its voltage limit is violated. For example, if at bus i the voltage limit is satisfied during the first T iterations ($v_i(t) \in [v_i, \bar{v}_i]$ for $t = 0, \dots, T-1$) then it does not communicate during the first T iterations. If bus's i voltage limit is then violated at iteration $T+1$ (say $v_i(t) > \bar{v}_i$) then $\bar{\lambda}_i^{\text{local}}$ becomes larger than 0 and bus i might need to communicate. After that, bus i might need to communicate while its voltage limit is violated and $v_i(t) > \bar{v}_i$. When the voltage $v_i(t)$ at bus i enters the feasible range $[v_i, \bar{v}_i]$ then $\bar{\lambda}_i^{\text{local}}(t)$ starts decreasing (see Eq. (8b)) but it might take few iterations for $\bar{\lambda}_i^{\text{local}}$ to reach 0 again, at which point bus i stops communicating. However, if the optimal voltage at node i , v_i^* , is on the boundary of $[v_i, \bar{v}_i]$ then $\bar{\lambda}_i^{\text{local}}$ might not converge to 0 and we will need condition (b) in **Step 4** to stop the communication at bus i .

The communication of each node is further restricted by condition (b) in **Step 4**. The condition says that bus i does not communicate unless the difference between the local variable $\bar{\lambda}_i^{\text{local}}$ and its global counterpart λ_i is significant enough. In particular, bus i does not communicate if the difference $|\lambda_i^{\text{local}}(t+1) - \lambda_i(t)|$ is upper bounded by the function $\text{ET}(t)$. The function $\text{ET} : \mathbb{N} \rightarrow \mathbb{R}_+$ can be chosen arbitrarily. For simplicity of presentation we assume $\text{ET}(t)$ is the same for all the buses. We can set $\text{ET}(t) = 0$ and still keep all the event-trigger benefits of condition (a). Another option is to let $\text{ET}(t)$ be a constant function, i.e., $\text{ET}(t) = \kappa$. This would mean that even the buses that have their optimal voltage v_i^* on the boundary of $[v_i, \bar{v}_i]$ (e.g. $v_i^* = \bar{v}_i$) can converge to approximate solutions in finite number of iterations. This is reasonable in practice, we are usually satisfied with obtaining a suboptimal solution in finite number of iterations, especially in dynamic environments. A third option is to let $\text{ET}(t)$ be a decreasing function, so that we allow more flexibility in the initial iterations but enforce more accuracy as the algorithm proceeds. We show in the next section that the algorithm converges to an optimal or approximately optimal solution in all three cases.

The **DIST-ETVC Algorithm** is fully distributed in the sense that there is no central authority that takes part in coordinating the algorithm. Nevertheless, when the buses communicate they need to broadcast information to the other inverters in the network. This is reasonable in networks with small to medium amount of inverters, i.e., when the

set \mathcal{I} is of manageable size. However, in networks with large amounts of inverters the information and management of the network is usually done by a centralized system operator. This motivates us to consider a decentralized variant of the **DIST-ETVC Algorithm** where the system operator manages the information flow, considered in the next subsection. An additional complication that can occur is communication collision, if multiple inverters communicate during the same iteration. We assume that the underlying communication protocol does not cause such conflicts, e.g., each node is allocated an orthogonal communication channel.

3.2 Decentralized Algorithm

We now consider a variant of the **DIST-ETVC Algorithm** where the information flow is managed by a system operator.

DEC-ETVC Algorithm: Decentralized Event Triggered Optimal Voltage Control

STEP 1 Initialization: Set $t = 0$, and $\lambda_i(0) = \lambda_i^{\text{local}}(0) = \lambda_i^{\text{local}}(0) = \lambda_i^{\text{local}}(0) = z_i(0) = z_i^{\text{local}}(0) = 0$ for all $i \in \mathcal{I}$.

STEP 2 Local Control: Each bus $i \in \mathcal{I}$ injects into the grid the reactive power

$$q_i(t) = \left[\frac{1}{a_i} (z_i^{\text{local}}(t) - b_i) \right]_{\underline{q}_i}^{\bar{q}_i}$$

where a_i and b_i come from the cost in Eq. (4).

STEP 3 Local Measurement and Computation: Same as **Step 3** in the **DIST-ETVC Algorithm**.

STEP 4 Event Triggered Communication: Is same as **Step 4** in the **DIST-ETVC Algorithm** except when bus i communicates then it sends $\lambda_i(t)$ to the system operator.

STEP 5 System Operators Update - Computation and Event-Triggered Communication: The system operator updates z as follows

$$z(t+1) = X\lambda(t+1). \quad (9)$$

For all $i \in \mathcal{I}$ the system operator decides if it communicates the new value $z_i(t+1)$ to bus i as follows.

- If $|z_i(t+1) - z_i^{\text{local}}(t)| \leq \text{ET}(t)$: then do nothing and bus i will update its local variable as

$$z_i^{\text{local}}(t+1) = z_i^{\text{local}}(t).$$

- Otherwise: Set

$$z_i^{\text{local}}(t+1) = z_i(t+1)$$

and communicate the new value z_i^{local} to bus i .

STEP 6 Set $t = t + 1$ and go back to **STEP 2**.

The main difference from the **DIST-ETVC Algorithm** is **Step 5** performed by the system operator and small modifications to **Step 2** and **Step 4** in the **DIST-ETVC Algorithm**. Moreover, we have introduced the new variables z and z^{local} . The variable z aggregates the information $X\lambda$

that the buses need to update their reactive power q_i in Eq. (7) of the **DIST-ETVC Algorithm**. This information is maintained by the system operator and updated according to Eq. (9) in **Step 5** of the **DEC-ETVC Algorithm**. The variable z_i^{Local} denotes the most recent information bus i has received about the value of z_i . The system operator sends the new information $z_i(t+1)$ to bus i only if it has changed significantly enough compared to z_i^{Local} , where the significant change is measured by the function $\text{ET}(t)$. As before, we assume that $\text{ET}(t)$ is the same for all the buses and the system operator to simplify the presentation. Similarly as in the **DIST-ETVC Algorithm** $\text{ET}(t)$ can be constant or decreasing function, we will explore both choices in simulations in Section 6. Moreover, we assume, similarly as for the **DIST-ETVC Algorithm**, that there are no communication conflicts even though multiple buses communicate during the same iteration or if the SO communicates to multiple buses during the same iteration.

The **DEC-ETVC Algorithm** nicely decouples the local and global information between the buses and the system operator. In particular, unlike in the **DIST-ETVC Algorithm**, in the **DEC-ETVC Algorithm** the buses do not need to know any global network information, such as the parameters of the matrix X . Such global information is maintained by the system operator in the update of $z(t+1)$ in Eq. (9). Still, the system operator does not need to access any local information from the buses, such as the local cost function $\frac{a_i}{2}q_i^2 + b_iq_i + c_i$ or bounds $\underline{q}_i, \bar{q}_i, \underline{v}_i$, and \bar{v}_i . We illustrate the main difference between the communication structures of the two algorithms in Figure 1.

4 MAIN CONVERGENCE RESULTS

In this section we highlight the convergence properties of the algorithm. In particular, we provide our main convergence results in Section 4.1. We then discuss how the convergence can be improved in Section 4.2. We focus our attention in this section on the **DIST-ETVC Algorithm**, and leave rigorous convergence analysis of the **DEC-ETVC Algorithm** as future work.

4.1 Main Results

We now show that the **DIST-ETVC Algorithm** converges to the optimal solution to Problem (3).

THEOREM 1 (CONSTANT STEP-SIZE). *Suppose that the relationship $v(q)$, between the voltage and reactive power, follows the linear Eq. (2) and that*

$$\gamma \in \left(0, \frac{1}{L}\right)$$

and let q^* be the optimal solution to Problem (3), $a_{\min} = \min\{a_1, \dots, a_I\}$, and $\lambda^* \in \mathbb{R}_+^{2n}$ be a constant given in Section 5.2. Then we have the following results:

a) If $\text{ET}(t) = 0$ then

$$\|q(t) - q^*\| \leq \frac{\|\lambda^*\|}{\sqrt{a_{\min}\gamma}} \frac{1}{\sqrt{t}}, \quad \text{for all } t \in \mathbb{N}.$$

b) If $\text{ET}(t) = \kappa$ then

$$\|\bar{q}(t) - q^*\| \leq \frac{\|\lambda^*\|}{\sqrt{a_{\min}\gamma}} \frac{1}{\sqrt{t}} + \frac{2\|X\|\sqrt{I}}{\sqrt{a_{\min}}} \sqrt{\kappa}, \quad \text{for all } t \in \mathbb{N},$$

where

$$\bar{q}(t) = \frac{1}{t} \sum_{\tau=1}^t q(\tau). \quad (10)$$

c) For general $\text{ET}(t)$ we have

$$\|\bar{q}(t) - q^*\| \leq \frac{1}{\sqrt{a_{\min}\gamma}} \frac{1}{\sqrt{t}} (\|\lambda^*\| + 2A(t)) \quad \text{for all } t \in \mathbb{N},$$

where

$$A(t) = \frac{4\|X\|^2\sqrt{I}}{\gamma a_{\min}} \sum_{\tau=1}^t \text{ET}(\tau). \quad (11)$$

PROOF. See Section 5.2. \square

The theorem establishes the convergence of the **DIST-ETVC Algorithm** for three different choices of the function $\text{ET}(t)$. We note that these results are theoretical upper bounds, establishing what behavior we can expect in the worst case. We illustrate in simulations that the algorithm converges faster in practice. Moreover, to yield the results we have considered the linear voltage Eq. (1) instead of the full nonlinear Eq. (2). We note that because of the high nonlinearities of Eq. (2) it is very difficult to get any theoretical guarantees by using the full model. However, by proving the convergence under the linear model we provide theoretical foundations to the algorithm, which are then complemented by evaluating the algorithm in simulations under the full nonlinear model in Section 6.

Part a) of the theorem shows that when $\text{ET}(t) = 0$ then the algorithm's iterates $q(t)$ converge at the rate $\mathcal{O}(1/\sqrt{t})$ to the optimal solution q^* . That means that we need at most K/ϵ^2 iterations, where $K = \|\lambda^*\|^2/(a_{\min}\gamma)$, to obtain an $\epsilon > 0$ accurate solution, i.e., to find q such that $\epsilon > \|q - q^*\|$. Part b) of the result establishes the convergence when $\text{ET}(t) = \kappa > 0$, i.e., when $\text{ET}(t)$ is a constant function. In this case we show that the average value $\bar{q}(t)$ of the iterates converges to the optimal value q^* , see Eq. (10). In particular, $\bar{q}(t)$ converges at the rate $\mathcal{O}(1/\sqrt{t})$ to an approximate solution, which gets more accurate as κ decreases. This is natural, since when we reach an κ -accurate solution then the buses do not communicate meaning that they cannot improve their estimates further. Such approximate solutions are reasonable in practice, since we can only run limited number of iterations in real-world application and hence must generally suffer some inaccuracies. We note that even though the theorem establishes only the converge of $\bar{q}(t)$, we show numerically in Section (6) that $q(t)$ has similar convergence behavior.

Part 3) of the theorem shows that we can ensure an asymptotic convergence of $\bar{q}(t)$ to the true optimal value q^* when $\text{ET}(t) > 0$, as long as $\text{ET}(t)$ converges to 0 fast enough. For example, if $\text{ET}(t)$ are summable

$$\sum_{t=0}^{\infty} \text{ET}(t) < \infty,$$

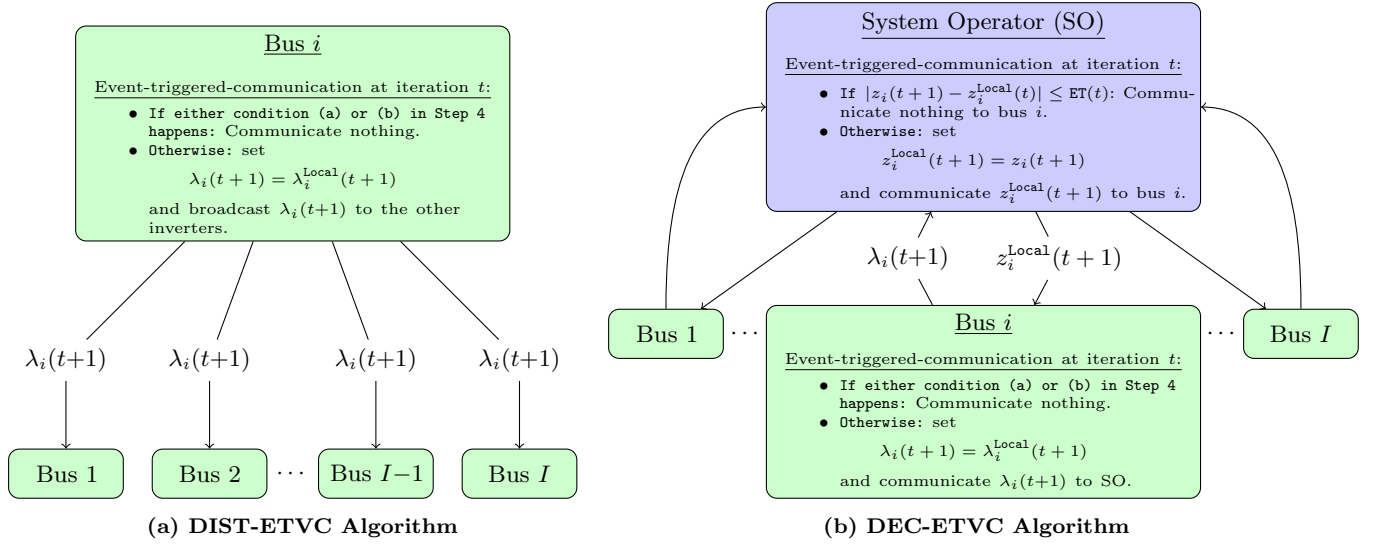


Figure 1: The communication in both algorithms. In the **DIST-ETVC Algorithm** the inverter buses broadcast the value of the dual variable $\lambda_i(t+1)$ when a local triggering event takes place. In the **DEC-ETVC Algorithm** the buses communicate $\lambda_i(t+1)$ only to the System Operator when the local triggering event takes place. Similarly, the System Operator, communicates the value $z_i(t+1)$ to bus i only if it has changed significantly enough from the local value $z_i^{\text{local}}(t)$.

e.g., $\text{ET}(t) = \rho^t$ for $\rho \in (0, 1)$, then $\bar{q}(t)$ converges to q^* at the rate $\mathcal{O}(1/\sqrt{t})$. Similarly, if $\text{ET}(t)$ decreases as $\mathcal{O}(1/t)$ then $A(t) = \mathcal{O}(\log(t))$ and $\bar{q}(t)$ converges to q^* at the rate $\mathcal{O}(\log(t)/\sqrt{t})$. We note that the convergence analysis in case c) is based on the inexact proximal gradient methods in [21], which provides a more thorough discussion on how different choices of $\text{ET}(t)$ affect the convergence rate. Like before, we note that the numerical results in Section 6 show that $q(t)$ converges similarly as $\bar{q}(t)$.

We have the following guarantees on communication.

THEOREM 2 (COMMUNICATION). *Suppose that the relationship $v(q)$, between the voltage and reactive power, follows the linear Eq. (2). Let $q^* \in \mathbb{R}^n$ be the unique optimal solution and set $\text{ET}(t) = 0$. If $v_i^* := v_i(q^*) \in (v_i, \bar{v}_i)$ then there exists T such that bus i does not need to communicate after iteration T , i.e.,*

$$\underline{\lambda}_i(t) = \bar{\lambda}_i(t) = 0, \quad \text{for all } t \geq T.$$

In particular, after iteration T at most $|\{i \in \mathcal{I} : v_i^ = v_i \text{ or } v_i^* = \bar{v}_i\}|$ nodes need to communicate.*

PROOF. See Section 5.3. □

The theorem shows that when $\text{ET}(t) = 0$ then the nodes that have their optimal voltage v_i^* strictly inside $[v_i, \bar{v}_i]$ communicate only finite number of times. In our numerical experiments in Section 6, we show that this results holds in general for all choices of $\text{ET}(t)$. We further show in Section 6 that when $\text{ET}(t)$ is a constant function then every bus needs to communicate only finite number of times. We leave it for future work to prove analytically these communication limits for more general cases than given in Theorem 2.

We now study variants of the algorithm that can provide better convergence behavior in some cases.

4.2 Improved Convergence

The **DIST-ETVC Algorithm** is essentially a gradient method for solving the dual problem. Gradient methods can be accelerated by using, e.g., Nesterov’s algorithm which gives better convergence behavior [19]. We now illustrate how such ideas can accelerate the convergence of the **DIST-ETVC Algorithm**.

THEOREM 3 (ACCELERATED GRADIENT METHOD). *Suppose that the relationship $v(q)$, between the voltage and reactive power, follows the linear Eq. (2). Consider the **DIST-ETVC Algorithm** with $\text{ET}(t) = 0$. Moreover, the updates in Eq.s (8a) and (8b) are changed as follows for each i . Set $\underline{\mu}_i(1) = \bar{\mu}_i(1) = 0$ and $\alpha(1) = 1$ and for $t \geq 1$ do*

$$\begin{aligned} \underline{\mu}_i(t) &= \lceil \underline{\lambda}_i(t) + \gamma(v_i - v_i(t)) \rceil + \\ \bar{\mu}_i(t) &= \lceil \bar{\lambda}_i(t) + \gamma(v_i(t) - v_i) \rceil + \\ \alpha(t+1) &= \frac{1 + \sqrt{1 + 4\alpha(t)^2}}{2} \\ \underline{\lambda}_i(t+1) &= \underline{\mu}_i(t) + \left(\frac{\alpha(t) - 1}{\alpha(t+1)} \right) (\underline{\mu}_i(t) - \underline{\mu}_i(t-1)) \\ \bar{\lambda}_i(t+1) &= \bar{\mu}_i(t) + \left(\frac{\alpha(t) - 1}{\alpha(t+1)} \right) (\bar{\mu}_i(t) - \bar{\mu}_i(t-1)), \end{aligned}$$

where $\gamma \in (0, 1/L)$. Then

$$\|q^\mu(t) - q^*\| \leq \frac{2\|\lambda^*\|}{\sqrt{a_{\min}\gamma}} \frac{1}{1+t},$$

where q^* is the optimal solution, $\lambda^* \in \mathbb{R}_+^{2n}$ is a constant given in Section 5.2, and q^μ is defined component-wise as follows

$$q_i^\mu(t) = \left[\frac{1}{a_i} \left(\sum_{j \in \mathcal{I}} X_{ij} (\underline{\mu}_i(t) - \bar{\mu}_i(t)) - b_i \right) \right]_{q_i}^{\bar{q}_i}.$$

PROOF. See Section 5.4. \square

The theorem shows that the acceleration can generate iterates $q^\mu(t)$ that converges at the rate $\mathcal{O}(1/t)$ to the optimal solution q^* , hence improving the $\mathcal{O}(1/\sqrt{t})$ convergence rate of the **DIST-ETVC Algorithm**. This means that we need $\mathcal{O}(1/\epsilon)$ iterations to obtain an $\epsilon > 0$ accurate solution, instead of $\mathcal{O}(1/\epsilon^2)$ iterations as in the standard **DIST-ETVC Algorithm**. Note that here we are measuring the convergence in the iterates $q^\mu(t)$ but not in $q(t)$. However, as the algorithm converges then $q(t) - q^\mu(t)$ converges to zero. We show in simulations that the accelerated version can improve the convergence in practice. We now prove these theorems.

5 CONVERGENCE ANALYSIS

In this section we prove the main convergence results from Section 4.1. First we provide some preliminary results in Section 5.1 and then prove Theorems 1, 2, and 3, respectively, in Sections 5.2, 5.3, and 5.4.

5.1 Duality Theory

The dual of the optimization problem in Eq. (4) is given by

$$\begin{aligned} & \underset{\lambda = (\underline{\lambda}, \bar{\lambda})}{\text{maximize}} && D(\lambda) := \min_{q \in [q, \bar{q}]} \mathcal{L}(q, \lambda) \\ & \text{subject to} && \lambda \geq 0, \end{aligned} \quad (12)$$

where $\underline{\lambda}$ and $\bar{\lambda}$ are, respectively, the dual variable associated to the lower and upper bounds in the constraint $[\underline{v}, \bar{v}]$, $D: \mathbb{R}^{2I} \rightarrow \mathbb{R}$ is the dual function and $\mathcal{L}(\cdot)$ is the Lagrangian function defined as

$$\mathcal{L}(q, \lambda) = \sum_{i=1}^I C_i(q_i) + \underline{\lambda}^\top (\underline{v} - v(q)) + \bar{\lambda}^\top (v(q) - \bar{v}), \quad (13)$$

where $\lambda = (\underline{\lambda}, \bar{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^n$, see Chapter 5 in [5] for details. We have the following result (proved in Appendix A).

LEMMA 1. *The dual gradient is*

$$\nabla D(\lambda) = \begin{bmatrix} \underline{v} - v(q(\lambda)) \\ v(q(\lambda)) - \bar{v} \end{bmatrix} \quad (14)$$

where

$$q(\lambda) = \underset{q \in [q, \bar{q}]}{\text{argmin}} \mathcal{L}(q, \lambda) = [\Lambda^{-1} X (\underline{\lambda} - \bar{\lambda}) - b]_{q, \bar{q}}, \quad (15)$$

where $\Lambda = \text{diag}(a_1, \dots, a_I)$ and $b = [b_1, \dots, b_n]^T$. Moreover, $\nabla D(\lambda)$ is L -Lipschitz continuous where $L = 4\|X\|^2/a_{\min}$ and $a_{\min} = \min_{i=1, \dots, n} a_i$.

From the lemma we can write the update of

$$\lambda^{\text{Local}} := (\underline{\lambda}^{\text{local}}, \bar{\lambda}^{\text{local}})$$

in Eq. (8) in the **DIST-ETVC Algorithm** as

$$\lambda^{\text{Local}}(t+1) = \lceil \lambda^{\text{Local}}(t) + \gamma \nabla D(\lambda(t)) \rceil_+, \quad (16)$$

where $\lambda(t) = (\underline{\lambda}(t), \bar{\lambda}(t))$. When $\text{ET}(t) = 0$ then $\lambda(t) = \lambda^{\text{Local}}(t)$ and the algorithm is just a standard dual gradient method

$$\lambda^{\text{Local}}(t+1) = \lceil \lambda^{\text{Local}}(t) + \gamma \nabla D(\lambda^{\text{Local}}(t)) \rceil_+. \quad (17)$$

When $\text{ET}(t) \neq 0$ then it can happen that $\lambda(t) \neq \lambda^{\text{Local}}(t)$ meaning that the steps in Eq. (16) do not follow a standard gradient method. However, the following lemma shows that Eq. (16) is an inexact gradient method, which has been well studied recently [10, 21].

LEMMA 2. *The steps of Eq. (8) in the **DIST-ETVC Algorithm** can be written equivalently as*

$$\lambda^{\text{Local}}(t+1) = \lceil \lambda^{\text{Local}}(t) + \gamma (\nabla D(\lambda^{\text{Local}}(t)) + e(t)) \rceil_+, \quad (18)$$

where $e(t) = \nabla D(\lambda(t)) - \nabla D(\lambda^{\text{Local}}(t)) \in \mathbb{R}^{2I}$ and

$$\|e(t)\| \leq \frac{4\|X\|^2 \sqrt{I}}{a_{\min}} \text{ET}(t).$$

PROOF. By using that $\nabla D(\cdot)$ is L -Lipschitz continuous with $L = 4\|X\|^2/a_{\min}$ (Lemma 1) we get

$$\begin{aligned} \|e(t)\| &= \|\nabla D(\lambda(t)) - \nabla D(\lambda^{\text{Local}}(t))\| \\ &\leq L \|\lambda(t) - \lambda^{\text{Local}}(t)\| \leq L \sqrt{I} \text{ET}(t) \end{aligned}$$

where we have used that $\lambda(t) = \underline{\lambda}(t) - \bar{\lambda}(t)$, $\lambda^{\text{Local}}(t) = \underline{\lambda}^{\text{Local}}(t) - \bar{\lambda}^{\text{Local}}(t)$, and the equivalence between the 2-norm and ∞ -norm, i.e., $\|x\| \leq \sqrt{I} \|x\|_\infty$ for $x \in \mathbb{R}^I$. \square

We will now use the fact that the algorithm is exact or inexact gradient method to prove its convergence. However, first we need one more result, connecting the convergence of the dual variables $\lambda^{\text{Local}}(t)$ to the primal variables $q(t)$.

LEMMA 3. *For any $\lambda \in \mathbb{R}_+^{2n}$ we have*

$$\|q(\lambda) - q^*\| \leq \sqrt{\frac{2}{a_{\min}} (D^* - D(\lambda))},$$

where D^* is the optimal value of the dual problem.

PROOF. The proof is similar to the proof of Theorem 1 in [3], but we include it here for completeness. For a fixed $\lambda \in \mathbb{R}_+^{2n}$ the Lagrangian function $\mathcal{L}(q, \lambda)$ in Eq. (13) is strongly-convex in q with convexity parameter a_{\min} , since $a_i > 0$ for $i = 1, \dots, n$. Therefore, by Definition 2.1.2 of strongly convex function in [20] and Theorem 2.1.2 in [5] we have for any $q \in [q, \bar{q}]$, $\lambda \in \mathbb{R}_+^{2n}$, and $q(\lambda)$ defined in Eq. (15) that

$$\frac{a_{\min}}{2} \|q - q(\lambda)\|^2 \leq \mathcal{L}(q, \lambda) - \mathcal{L}(q(\lambda), \lambda).$$

Let $q^* \in [q, \bar{q}]$ and $\lambda^* \in \mathbb{R}_+^{2n}$, respectively, be optimal solutions to the primal and dual problems. Then by Eq. (13) we have (setting $\lambda = (\underline{\lambda}, \bar{\lambda})$)

$$\begin{aligned} \mathcal{L}(q^*, \lambda) - \mathcal{L}(q(\lambda), \lambda) &= \sum_{i=1}^n C_i(q_i^*) + \underline{\lambda}^\top (\underline{v} - v(q^*)) \\ &\quad + \bar{\lambda}^\top (v(q^*) - \bar{v}) - D(\lambda) \\ &\leq D(\lambda^*) - D(\lambda), \end{aligned}$$

where we have used that $D(\lambda^*) = \sum_{i=1}^n C_i(q_i^*)$ by the strong duality and that $\underline{\lambda}^\top (\underline{v} - v(q^*))$, $\bar{\lambda}^\top (v(q^*) - \bar{v}) \geq 0$ since

$\lambda \geq 0$ and q^* is a primal feasible point. Combining the two inequalities above yields the result. \square

5.2 Proof of Theorem 1

We have shown that the update of $\lambda^{\text{Local}}(t)$ in Eq. (8) of the **DIST-ETVC Algorithm** is a standard gradient method when $\text{ET}(t) = 0$ (cf. Eq. (17)) and an inexact gradient method otherwise (cf. Eq. (18)). Because of the projection $[\cdot]_+$, it will be helpful to write the steps of the algorithm equivalently as a proximal gradient method. We can then directly use the convergence results in [4] and [21], respectively, for exact and inexact proximal gradient methods. To that end, we write the dual problem in Eq. (19) equivalently in the form of the problem in Eq. (M) on page 14 in [4]

$$\underset{\lambda \in \mathbb{R}^{2n}}{\text{minimize}} \quad -D(\lambda) + g(\lambda) \quad (19)$$

where g is the indicator function

$$g(\lambda) = \begin{cases} 0 & \text{if } \lambda \in \mathbb{R}_+^{2n} \\ \infty & \text{otherwise.} \end{cases}$$

We can now write the algorithm equivalently as

$$\begin{aligned} \lambda^{\text{Local}}(t+1) &= [\lambda^{\text{Local}}(t) + \gamma(\nabla D(\lambda^{\text{Local}}(t)) + e(t))]_+, \\ &= \text{prox}_\gamma(g)(\lambda^{\text{Local}}(t) + \gamma(\nabla D(\lambda^{\text{Local}}(t)) + e(t))), \end{aligned}$$

where $e(t)$ is from Lemma 2. We now prove a), b) and c).

Case a): From Theorem 1.1 in [4] we have that

$$D^* - D(\lambda^{\text{Local}}(t)) \leq \frac{\|\lambda^*\|^2}{2\gamma t} \quad \text{for } t \geq 1,$$

where D^* is the optimal value of the dual problem. Combining this inequality with Lemma 2 we get the result.

Case b): Using the concavity of $D(\cdot)$ and properties of the proximal operator we get the inequality (for details see, e.g., the proof of Eq. (10) in the extended version of [21])

$$\begin{aligned} D^* - D(\lambda^{\text{Local}}(t)) &\leq -\frac{1}{2\gamma} \|\lambda^{\text{Local}}(t) - \lambda^*\|^2 \\ &\quad + \frac{1}{2\gamma} \|\lambda^{\text{Local}}(t-1) - \lambda^*\|^2 + \|e(t)\|, \end{aligned}$$

where λ^* is some dual optimal solution. By summing over t we get

$$\begin{aligned} \sum_{\tau=1}^t D^* - D(\lambda^{\text{Local}}(\tau)) &\leq \frac{1}{2\gamma} \|\lambda^{\text{Local}}(0) - \lambda^*\|^2 + \sum_{\tau=1}^t \|e(\tau)\|, \\ &\leq \frac{1}{2\gamma} \|\lambda^*\|^2 + t \frac{4\|X\|^2 I}{a_{\min}} \kappa, \end{aligned}$$

where we have used that $\lambda^{\text{Local}}(0) = 0$ in the second inequality. By taking the average we get

$$\frac{1}{t} \sum_{\tau=1}^t D^* - D(\lambda^{\text{Local}}(\tau)) \leq \frac{1}{2\gamma t} \|\lambda^*\|^2 + \frac{4\|X\|^2 I}{a_{\min}} \kappa.$$

Therefore, by Lemma 3 we have

$$\begin{aligned} \|\bar{q}(t) - q^*\| &\leq \frac{1}{t} \sum_{\tau=1}^t \|q(\tau) - q^*\| \\ &\leq \frac{1}{t} \sum_{\tau=1}^t \sqrt{\frac{2}{a_{\min}} (D^* - D(\lambda^{\text{Local}}(\tau)))} \\ &\leq \frac{\sqrt{t}}{t} \sqrt{\sum_{\tau=1}^t \frac{2}{a_{\min}} (D^* - D(\lambda^{\text{Local}}(\tau)))} \\ &\leq \frac{1}{\sqrt{t}} \frac{\|\lambda^*\|}{\sqrt{\gamma a_{\min}}} + \sqrt{\frac{4\|X\|^2 I}{a_{\min}}} \kappa, \end{aligned}$$

which yields the result.

Case c): From Proposition 1 in [21] we have that

$$\begin{aligned} D^* - D \left(\frac{1}{t} \sum_{\tau=1}^t \lambda^{\text{Local}}(\tau) \right) &\leq \frac{1}{t} \sum_{\tau=1}^t D^* - D(\lambda^{\text{Local}}(\tau)) \\ &\leq \frac{1}{2\gamma t} (\|\lambda^*\| + 2A(t))^2, \end{aligned}$$

where $A(t)$ is defined in Eq. (11). We can now derive the result by combining this with Lemma 3 and following similar steps as in Case b).

5.3 Proof of Theorem 2

Suppose that $v_i^* := v_i(p^*) \in (v_i, \bar{v}_i)$. From Theorem 1 and the fact that $v(\cdot)$ is a continuous function, there exists $\delta > 0$ and $T_0 \in \mathbb{N}$ such that $v_i(t) \in (v_i^* - \delta, v_i^* + \delta) \subseteq (v_i, \bar{v}_i)$. Then $\eta := \min\{(v_i^* - \delta) - v_i, \bar{v}_i - (v_i^* + \delta)\} > 0$. It follows that

$$v_i - v_i(t) < -\eta \quad \text{and} \quad v_i(t) - \bar{v}_i < -\eta \quad \text{for } t \geq T_0.$$

Hence, from Eq.s (8a) and (8b) we have for $t \geq T_0$ that

$$\underline{\lambda}_i(t) \leq [\underline{\lambda}_i(T_0) - (t - T_0)\gamma\eta]_+, \quad \bar{\lambda}_i(t) \leq [\bar{\lambda}_i(T_0) - (t - T_0)\gamma\eta]_+.$$

Therefore, if we set $T = T_0 + \max\{\underline{\lambda}_i(T_0), \bar{\lambda}_i(T_0)\}/(\gamma\eta)$ then $\underline{\lambda}_i(t) = \bar{\lambda}_i(t) = 0$ for all $t \geq T$.

5.4 Proof of Theorem 3

Here we have changed the proximal gradient method with constant step-size, in Eq. (17), to fast proximal gradient method with constant step size, see page 23 in [4]. From Theorem 1.4 in [4] we have that

$$D^* - D(\lambda^{\text{Local}}(t)) \leq \frac{2\|\lambda^*\|^2}{\gamma(t+1)^2} \quad \text{for } t \geq 1.$$

The convergence rate in Theorem 3 is now obtained by combining this lemma with Lemma 3.

6 NUMERICAL EXPERIMENTS

We test the algorithms on the 56 bus network in [14]. In Subsection 6.1 we illustrate the performance of the algorithms on the full nonlinear AC power flow model. In Subsection 6.2 we test the algorithms on the linearized power flows in Equation (1) and compare the results to the theory in Section 4.

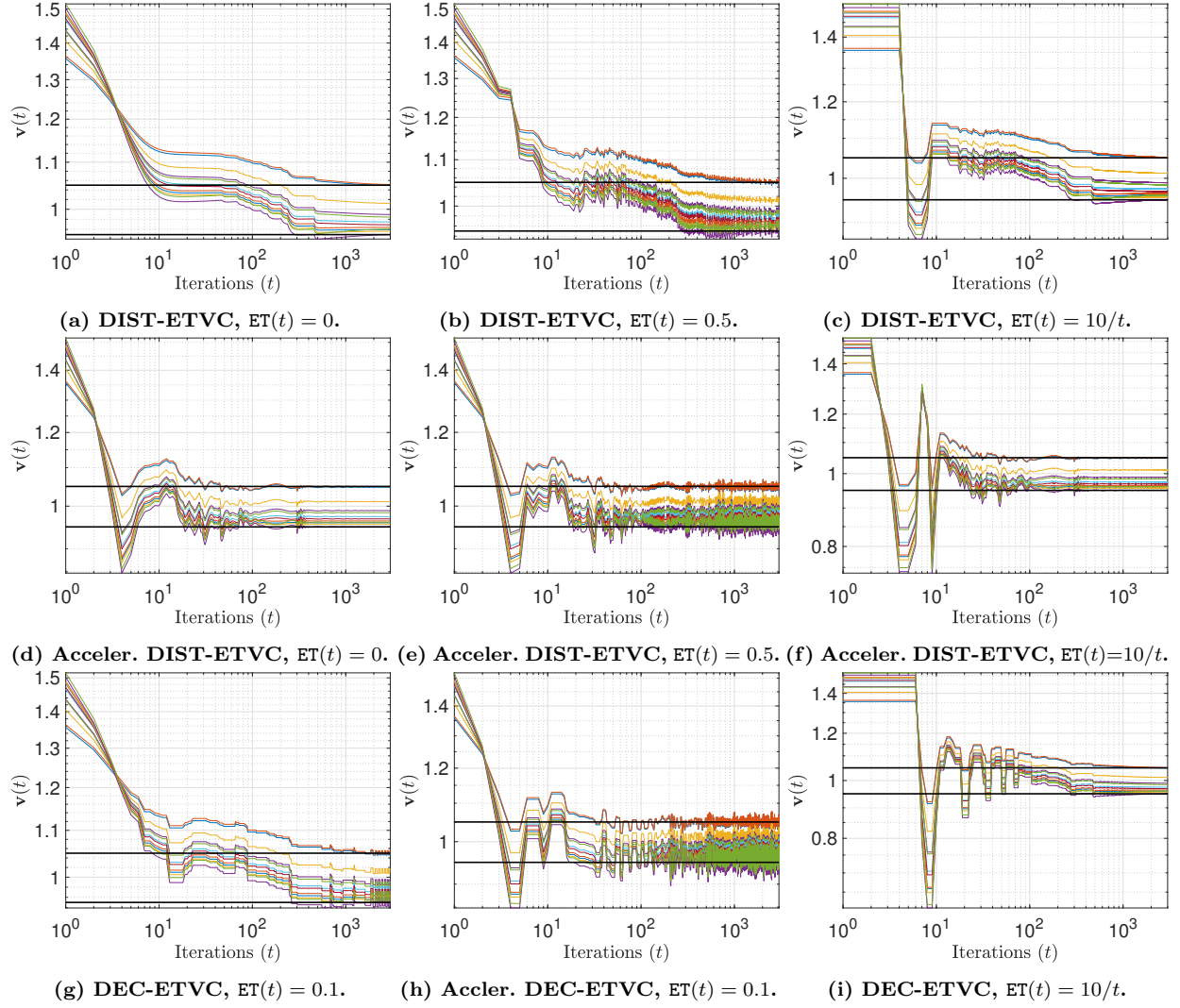


Figure 2: Convergence of the algorithms on the full nonlinear AC power flow.

6.1 Full Nonlinear AC Power Flow

We now illustrate the algorithms' convergence on the full nonlinear AC power flow grids. To solve the nonlinear power flow equations we use the MATLAB extension package MATPOWER [27]. In these simulations Bus 1 is the feeder bus and we assume that there is an inverter that can inject reactive power on buses 7, 8, ..., 18, so there are $I = 12$ inverter buses in total. The voltage and reactive power limits are $\underline{v}_i = 0.95$, $\bar{v}_i = 1.05$, $\underline{q}_i = -2$ and $\bar{q}_i = 2$. The real power injections are chosen at uniform random from the interval $[-0.5, 1.5]$. We set $a_i = 1$ and $b_i = 0$ in $C_i(q_i)$.

Figure 2 depicts the convergence of the voltages for the standard and accelerated versions of the **DIST-ETVC** and **DEC-ETVC** algorithms. We test the algorithms for $ET(t) =$

0 , $ET(t) = \kappa$, $ET(t) = 10/t$, where $\kappa = 0.5$ in the **DIST-ETVC Algorithm** and $\kappa = 0.1$ in the **DEC-ETVC Algorithm**. All of the algorithms converge to a solution or fluctuate around an approximate solution. As could be expected, the accelerated methods generally converge much faster. However, when $ET(t) = \kappa$ then the accelerated methods induce more fluctuations than the non-accelerated methods. This can be expected, since when $ET(t) = \kappa$ then the algorithms are equivalent to inexact dual-gradient methods (see Section 5) and accelerated gradient methods are generally less robust to noise [10, 21].

Table 1 summarizes the average number of communications per bus during the 3000 iterations depicted in Figure 2. The number in the parentheses indicates the percentages of iterations when each bus communicates. In all cases the event-triggered algorithms significantly reduce the communications. In the worst case buses are communicating during less than

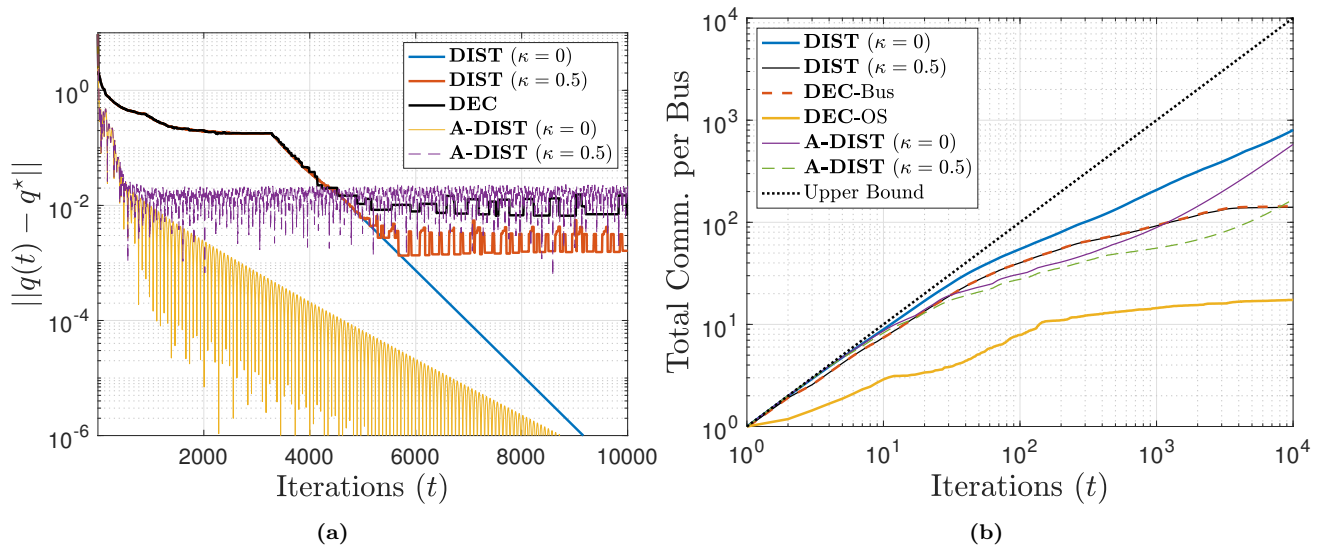


Figure 3: Convergence of the algorithms on the linear (approximated) power flow.

Algorithm	ET(t)	# Comm. Bus	# Comm. SO
DIST	0	851 (28.37%)	-
DIST	0.5	28 (0.93%)	-
DIST	10/ t	417 (13.90%)	-
A-DIST	0	562 (18.73%)	-
A-DIST	0.5	107 (3.57%)	-
A-DIST	10/ t	442 (14.73%)	-
DEC	0.1	137 (4.57%)	12 (0.40%)
DEC	10/ t	857 (14.27%)	175 (5.83 %)
A-DEC	0.1	428 (29.17%)	24 (0.80%)

Table 1: The average number of communication per bus during the 3000 iterations depicted in Figure 2.

30% of the iterations and for the **DIST Algorithm** with $ET(t) = 0.5$ each bus communicates during less than 1% of iterations without a significant loss in solution accuracy. Even when $ET(t) = 0$ then the event-triggered algorithms save a large amount of communications. In the **DEC Algorithm** with $\kappa = 0.1$ the OS communicates only 12 times to each bus (at 0.4% of the iterations). Because of the complexity of the nonlinear model we cannot find the true optimal solution q^* . However, in the next subsection we illustrate the convergence of $\|q(t) - q^*\|$ for the linear power flows.

6.2 Linearized Power Flow

We now evaluate the algorithm on the linearized power flow model in Equation (2). All the parameters are the same as in the previous section, except here we assume that every bus except bus 1 is an inverter bus, i.e., $\mathcal{I} = \{2, 3, \dots, 56\}$ and $I = 55$. In the **DEC-ETVC Algorithm** we let $ET(t) = 0.5$ at the buses and $ET(t) = 0.1$ at the system operator. The results are depicted in Figure 3, where (a) depicts the convergence of the reactive powers to the optimal solution and

(b) depicts the average cumulative communication for each bus. The results show that when $ET(t) = 0$ then the algorithm converges to the optimal solution. However, when $ET(t) = \kappa$ then the algorithm converges only to an approximate solution. Both of these results coincide with what we would expect from Theorem 1. In all cases the event triggered algorithm drastically reduces the communication compared to if the buses and the SO would communicate at every iteration.

7 CONCLUSIONS

This paper explored optimal smart grid control algorithms that use event triggered communication. In these algorithms the network's buses only communicate at critical times when their local voltage constraint is violated and if the difference between local and global versions of the variable that they communicate is significant enough. We proved the algorithms' convergence to optimal solutions assuming linearized power flows. We illustrate the good performance on the full nonlinear power flows in simulations. The results show significant communication reductions of the suggested algorithms.

8 ACKNOWLEDGMENT

We are grateful to Prof. Hartmut Schmeck, our shepherd, and to the anonymous reviewers for their constructive comments that have helped us to increase the quality of the manuscript. This work was supported by NSF EPCN 1608509, NSF CAREER 1553407, ARPA-E NODES, and the Harvard Climate Change Solution Funds.

REFERENCES

- [1] M. E. Baran and F. F. Wu. 1989. Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Transactions on Power Delivery* 4, 2 (Apr 1989), 1401–1407. <https://doi.org/10.1109/61.25627>

- [2] M. E Baran and F. F Wu. 1989. Optimal capacitor placement on radial distribution systems. *IEEE Transactions on power Delivery* 4, 1 (1989), 725–734.
- [3] A. Beck, A. Nedic, A. Ozdaglar, and M. Teboulle. 2014. An O(1/k) Gradient Method for Network Resource Allocation Problems. *Control of Network Systems, IEEE Transactions on* 1, 1 (March 2014), 64–73.
- [4] Amir Beck and Marc Teboulle. 2009. Gradient-based algorithms with applications to signal recovery. *Convex optimization in signal processing and communications* (2009), 42–88.
- [5] D. P. Bertsekas. 1999. *Nonlinear Programming: 2nd Edition*. Athena Scientific.
- [6] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri. 2015. Distributed Reactive Power Feedback Control for Voltage Regulation and Loss Minimization. *IEEE Trans. Automat. Control* 60, 4 (April 2015), 966–981. <https://doi.org/10.1109/TAC.2014.2363931>
- [7] S. Bolognani and S. Zampieri. 2013. A Distributed Control Strategy for Reactive Power Compensation in Smart Microgrids. *IEEE Trans. Automat. Control* 58, 11 (Nov 2013), 2818–2833. <https://doi.org/10.1109/TAC.2013.2270317>
- [8] Pedro MS Carvalho, Pedro F Correia, and Luís AFM Ferreira. 2008. Distributed reactive power generation control for voltage rise mitigation in distribution networks. *IEEE Transactions on Power Systems* 23, 2 (2008), 766–772.
- [9] G. Cavraro, S. Bolognani, R. Carli, and S. Zampieri. 2016. The value of communication in the voltage regulation problem. In *2016 54th IEEE Conference on Decision and Control (CDC)*.
- [10] Olivier Devolder, François Glineur, and Yurii Nesterov. 2014. First-order methods of smooth convex optimization with inexact oracle. *Mathematical Programming* 146, 1-2 (2014), 37–75.
- [11] Dimos V Dimarogonas, Emilio Frazzoli, and Karl H Johansson. 2012. Distributed event-triggered control for multi-agent systems. *IEEE Trans. Automat. Control* 57, 5 (2012), 1291–1297.
- [12] Yuan Fan, Guoqiang Hu, and Magnus Egerstedt. 2017. Distributed reactive power sharing control for microgrids with event-triggered communication. *IEEE Transactions on Control Systems Technology* 25, 1 (2017), 118–128.
- [13] Masoud Farivar, Lijun Chen, and Steven Low. 2013. Equilibrium and dynamics of local voltage control in distribution systems. In *52nd IEEE Conference on Decision and Control*. IEEE, 4329–4334.
- [14] Masoud Farivar, Russell Neal, Christopher Clarke, and Steven Low. 2012. Optimal inverter var control in distribution systems with high pv penetration. In *2012 IEEE Power and Energy Society General Meeting*. IEEE, 1–7.
- [15] Chaojie Li, Xinghuo Yu, Wenwu Yu, Tingwen Huang, and Zhi-Wei Liu. 2016. Distributed event-triggered scheme for economic dispatch in smart grids. *IEEE Transactions on Industrial Informatics* 12, 5 (2016), 1775–1785.
- [16] N. Li, G. Qu, and M. Dahleh. 2014. Real-time decentralized voltage control in distribution networks. In *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*. 582–588. <https://doi.org/10.1109/ALLERTON.2014.7028508>
- [17] Sindri Magnússon, Guannan Qu, Carlo Fischione, and Na Li. 2019. Voltage control using limited communication. *IEEE Transactions on Control of Network Systems* (2019).
- [18] Sindri Magnússon, Guannan Qu, and Na Li. 2019. Distributed Voltage Control with Communication Delays. *arXiv preprint arXiv:1903.01065* (2019).
- [19] Yurii Nesterov. 1983. A method of solving a convex programming problem with convergence rate O(1/k²). In *Soviet Mathematics Doklady*, Vol. 27. 372–376.
- [20] Yurii Nesterov. 2004. *Introductory Lectures on Convex Optimization*. Springer. 1177–1189 pages.
- [21] Mark Schmidt, Nicolas L Roux, and Francis R Bach. 2011. Convergence rates of inexact proximal-gradient methods for convex optimization. In *Advances in neural information processing systems*. 1458–1466.
- [22] Petr vSulc, Scott Backhaus, and Michael Chertkov. 2014. Optimal distributed control of reactive power via the alternating direction method of multipliers. *IEEE Transactions on Energy Conversion* 29, 4 (2014), 968–977.
- [23] Paulo Tabuada. 2007. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Trans. Automat. Control* 52, 9 (2007), 1680–1685.
- [24] Pu Wan and Michael D Lemmon. 2010. Optimal power flow in microgrids using event-triggered optimization. In *American Conference (ACC), 2010*. IEEE, 2521–2526.
- [25] B. Zhang, A. Y. S. Lam, A. D. Domínguez-García, and D. Tse. 2015. An Optimal and Distributed Method for Voltage Regulation in Power Distribution Systems. *IEEE Transactions on Power Systems* 30, 4 (July 2015), 1714–1726. <https://doi.org/10.1109/TPWRS.2014.2347281>
- [26] H. Zhu and H. J. Liu. 2016. Fast Local Voltage Control Under Limited Reactive Power: Optimality and Stability Analysis. *IEEE Transactions on Power Systems* 31, 5 (Sept 2016), 3794–3803. <https://doi.org/10.1109/TPWRS.2015.2504419>
- [27] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, Robert John Thomas, et al. 2011. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on power systems* 26, 1 (2011), 12–19.

A PROOF OF LEMMA 1

The gradient of the Lagrangian function in Equation (13) with respect to q is

$$\nabla \mathcal{L}(q, \lambda) = \Lambda q + b + X(\bar{\lambda} - \underline{\lambda}).$$

From Proposition 6.1.1 in [5] the dual gradient is

$$\nabla D(\lambda) = \left[\begin{array}{c} v - v(q^*(\lambda)) \\ v(q^*(\lambda)) - \bar{v} \end{array} \right] \quad (20)$$

where

$$q^*(\lambda) = \underset{q \in [\underline{q}, \bar{q}]}{\operatorname{argmin}} \mathcal{L}(q, \lambda). \quad (21)$$

Therefore, to prove Equation (14) it suffices to show that $q^*(\lambda) = q(\lambda)$. We prove that by showing that $q(\lambda)$ is the optimal solution to the optimization problem in Equation (21) by using Proposition 2.1.2. in [5]. That is we show that

$$\nabla \mathcal{L}(q(\lambda), \lambda)^T (q - q(\lambda)) \geq 0 \quad \text{for all } q \in [\underline{q}, \bar{q}].$$

We have that

$$\nabla \mathcal{L}(q(\lambda), \lambda)^T (q - q(\lambda)) = \sum_{i=1}^n \nabla_i \mathcal{L}(q(\lambda), \lambda) (q_i - q_i(\lambda)),$$

and hence it suffices to show that a) $\nabla_i \mathcal{L}(q(\lambda), \lambda) = 0$ if $q_i(\lambda) \in (\underline{q}_i, \bar{q}_i)$, b) $\nabla_i \mathcal{L}(q(\lambda), \lambda) \geq 0$ if $q_i(\lambda) = \underline{q}_i$, and c) $\nabla_i \mathcal{L}(q(\lambda), \lambda) \leq 0$ if $q_i(\lambda) = \bar{q}_i$. We now conclude the proof by proving a), b), and c) below.

Proof of a): Note that $q(\lambda) = [q_{\text{UC}}^*(\lambda)]_{\underline{q}}^{\bar{q}}$ where

$$q_{\text{UC}}^*(\lambda) := \underset{q \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{L}(q, \lambda) = \Lambda^{-1} X(\bar{\lambda} - \underline{\lambda}) - b$$

is the unconstrained optimizer of $\mathcal{L}(\cdot, \lambda)$. If $q_i(\lambda) \in (\underline{q}_i, \bar{q}_i)$ then $q_i(\lambda) = [q_{\text{UC}}^*(\lambda)]_i$. Using that Λ is a diagonal matrix we also have that

$$\begin{aligned} \nabla_i \mathcal{L}(q(\lambda), \lambda) &= a_i q_i(\lambda) + b_i + [X(\bar{\lambda} - \underline{\lambda})]_i \\ &= a_i [q_{\text{UC}}^*(\lambda)]_i + b_i + [X(\bar{\lambda} - \underline{\lambda})]_i \\ &= \nabla_i \mathcal{L}(q_{\text{UC}}^*(\lambda), \lambda) = 0, \end{aligned}$$

where we used that $q_{\text{UC}}^*(\lambda)$ is the global optimizer in the final equality.

Proof of b) and c): If $q_i(\lambda) = \underline{q}_i$ then $[q_{\text{UC}}^*(\lambda)]_i \leq q(\lambda)$. Therefore, since $a_i > 0$ we have

$$\begin{aligned} 0 &= \nabla_i \mathcal{L}(q_{\text{UC}}^*(\lambda), \lambda) = a_i [q_{\text{UC}}^*(\lambda)]_i + b_i + [X(\bar{\lambda} - \underline{\lambda})]_i, \\ &\leq a_i q_i(\lambda) + b_i + [X(\bar{\lambda} - \underline{\lambda})]_i = \nabla_i \mathcal{L}(q(\lambda), \lambda) \end{aligned}$$

Condition c) follows from similar arguments as condition b).

Finally we show that $\nabla D(\lambda)$ is L -Lipschitz continuous with $L = 4\|X\|^2/a_{\min}$. Take $\lambda_1 = (\underline{\lambda}_1, \bar{\lambda}_1), \lambda_2 = (\underline{\lambda}_2, \bar{\lambda}_2) \in \mathbb{R}_+^{2n}$,

then from Equations (14) and (15) we have

$$\begin{aligned} \|\nabla D(\lambda_1) - \nabla D(\lambda_2)\| &\leq 2\|v(q(\lambda_1)) - v(q(\lambda_2))\| \\ &\leq 2\|X\| \|q(\lambda_1) - q(\lambda_2)\| \\ &\leq 2\|X\|^2 \|\Lambda^{-1}\| \|\underline{\lambda}_1 - \underline{\lambda}_2 + \bar{\lambda}_2 - \bar{\lambda}_1\| \\ &\leq 4 \frac{\|X\|^2}{a_{\min}} \|\lambda_1 - \lambda_2\|, \end{aligned}$$

where we have used the triangle inequality in the first and last inequality and the fact that $\|\Lambda^{-1}\| = 1/a_{\min}$.